FTOS Configuration Guide
FTOS 8.4.2.7
E-Series TeraScale, C-Series, S-Series (S50/S25)

Notes, Cautions, and Warnings

NOTE: A NOTE indicates important information that helps you make better use of your computer.

CAUTION: A CAUTION indicates potential damage to hardware or loss of data if instructions are not followed.

WARNING: A WARNING indicates a potential for property damage, personal injury, or death.

October 2012

1

About this Guide

Objectives

This guide describes the protocols and features supported by the Dell Force10 Operating System (FTOS) and provides configuration instructions and examples for implementing them. It supports the system platforms E-Series, C-Series, and S-Series.

The E-Series ExaScale platform is supported with FTOS version 8.1.1.0. and later.

Though this guide contains information on protocols, it is not intended to be a complete reference. This guide is a reference for configuring protocols on Dell Force10 systems. For complete information on protocols, refer to other documentation including IETF Requests for Comment (RFCs). The instructions in this guide cite relevant RFCs, and Appendix 63, Standards Compliance contains a complete list of the supported RFCs and Management Information Base files (MIBs).

Audience

This document is intended for system administrators who are responsible for configuring and maintaining networks and assumes you are knowledgeable in Layer 2 and Layer 3 networking technologies.

Conventions

This document uses the following conventions to describe command syntax:

Convention	Description
keyword	Keywords are in bold and should be entered in the CLI as listed.
parameter	Parameters are in italics and require a number or word to be entered in the CLI.
{X}	Keywords and parameters within braces must be entered in the CLI.
[X]	Keywords and parameters within brackets are optional.
x | y	Keywords and parameters separated by bar require you to choose one.

Information Symbols

Table 1-1 describes symbols contained in this guide.

Table 1-1. Information Symbols

Symbol	Warning	Description
	FTOS Behavior	This symbol informs you of an FTOS behavior. These behaviors are inherent to the Dell Force10 system or FTOS feature and are non-configurable.
	Platform Specific Feature	This symbol informs you of a feature that supported on one or two platforms only: E is for E-Series, C is for C-Series, S is for S-Series.
	E-Series Specific Feature/Command	If a feature or command applies to only one of the E-Series platforms, a separate symbol calls this to attention: E T for the TeraScale or E X for the ExaScale.
	Exception	This symbol is a note associated with some other text on the page that is marked with an asterisk.

Related Documents

For more information about the Dell Force10 E-Series, C-Series, and S-Series refer to the following documents:

- *FTOS Command Reference*
- *Installing and Maintaining the <Dell Force10 chassis> System*
- *FTOS Release Notes*

2

Configuration Fundamentals

The FTOS Command Line Interface (CLI) is a text-based interface through which you can configure interfaces and protocols. The CLI is largely the same for the E-Series, C-Series, and S-Series with the exception of some commands and command outputs. The CLI is structured in modes for security and management purposes. Different sets of commands are available in each mode, and you can limit user access to modes using privilege levels.

In FTOS, after a command is enabled, it is entered into the running configuration file. You can view the current configuration for the whole system or for a particular CLI mode. To save the current configuration copy the running configuration to another location.

Note: Due to a differences in hardware architecture and the continued system development, features may occasionally differ between the platforms. These differences are identified by the information symbols shown on Table 1-1 on page 34.

Accessing the Command Line

Access the command line through a serial console port or a Telnet session (Figure 2-1). When the system successfully boots, you enter the command line in the EXEC mode.

Note: You must have a password configured on a virtual terminal line before you can Telnet into the system. Therefore, you must use a console connection when connecting to the system for the first time.

Figure 2-1. Logging into the System using Telnet

CLI Modes

Different sets of commands are available in each mode. A command found in one mode cannot be executed from another mode (with the exception of EXEC mode commands preceded by the command **do**; see The do Command on page 40). You can set user access rights to commands and command modes using privilege levels; for more information on privilege levels and security options, refer to Chapter 9, Security, on page 627.

The FTOS CLI is divided into three major mode levels:

- **EXEC mode** is the default mode and has a privilege level of 1, which is the most restricted level. Only a limited selection of commands is available, notably **show** commands, which allow you to view system information.
- **EXEC Privilege mode** has commands to view configurations, clear counters, manage configuration files, run diagnostics, and enable or disable debug operations. The privilege level is 15, which is unrestricted. You can configure a password for this mode; see Configure the Enable Password on page 50.
- **CONFIGURATION mode** enables you to configure security features, time settings, set logging and SNMP functions, configure static ARP and MAC addresses, and set line cards on the system.

Beneath CONFIGURATION mode are sub-modes that apply to interfaces, protocols, and features. Figure 2-2 illustrates this sub-mode command structure. Two sub-CONFIGURATION modes are important when configuring the chassis for the first time:

- **INTERFACE sub-mode** is the mode in which you configure Layer 2 and Layer 3 protocols and IP services specific to an interface. An interface can be physical (Management interface, 1-Gigabit Ethernet, or 10-Gigabit Ethernet, or SONET) or logical (Loopback, Null, port channel, or VLAN).
- **LINE sub-mode** is the mode in which you to configure the console and virtual terminal lines.

Note: At any time, entering a question mark (?) will display the available command options. For example, when you are in CONFIGURATION mode, entering the question mark first will list all available commands, including the possible sub-modes.

Figure 2-2. CLI Modes in FTOS

EXEC
EXEC Privilege
CONFIGURATION
- ARCHIVE
- AS-PATH ACL
- INTERFACE
  - GIGABIT ETHERNET
  - 10 GIGABIT ETHERNET
  - INTERFACE RANGE
  - LOOPBACK
  - MANAGEMENT ETHERNET
  - NULL
  - PORT-CHANNEL
  - SONET
  - VLAN
  - VRRP
- IP
- IPv6
- IP COMMUNITY-LIST
- IP ACCESS-LIST
  - STANDARD ACCESS-LIST
  - EXTENDED ACCESS-LIST
- LINE
  - AUXILIARY
  - CONSOLE
  - VIRTUAL TERMINAL
- MAC ACCESS-LIST
- MONITOR SESSION
- MULTIPLE SPANNING TREE
- Per-VLAN SPANNING TREE
- PREFIX-LIST
- RAPID SPANNING TREE
- REDIRECT
- ROUTE-MAP
- ROUTER BGP
- ROUTER ISIS
- ROUTER OSPF
- ROUTER RIP
- SPANNING TREE
- TRACE-LIST

Navigating CLI Modes

The FTOS prompt changes to indicate the CLI mode. Table 2-1 lists the CLI mode, its prompt, and information on how to access and exit this CLI mode. You must move linearly through the command modes, with the exception of the **end** command which takes you directly to EXEC Privilege mode; the **exit** command moves you up one command mode level.

Note: Sub-CONFIGURATION modes all have the letters “conf” in the prompt with additional modifiers to identify the mode and slot/port information. These are shown in Table 2-1.

Table 2-1. FTOS Command Modes

CLI Command Mode		Prompt	Access Command
EXEC		FTOS>	Access the router through the console or Telnet.
EXEC Privilege		FTOS#	• From EXEC mode, enter the command **enable**. • From any other mode, use the command **end**.
CONFIGURATION		FTOS(conf)#	• From EXEC privilege mode, enter the command **configure**. • From every mode except EXEC and EXEC Privilege, enter the command **exit**.
Note: Access all of the following modes from CONFIGURATION mode.			
	ARCHIVE	FTOS(conf-archive)	**archive**
	AS-PATH ACL	FTOS(config-as-path)#	**ip as-path access-list**
INTERFACE modes	Gigabit Ethernet Interface	FTOS(conf-if-gi-0/0)#	**interface**
	10 Gigabit Ethernet Interface	FTOS(conf-if-te-0/0)#	
	Interface Range	FTOS(conf-if-range)#	
	Loopback Interface	FTOS(conf-if-lo-0)#	
	Management Ethernet Interface	FTOS(conf-if-ma-0/0)#	
	Null Interface	FTOS(conf-if-nu-0)#	
	Port-channel Interface	FTOS(conf-if-po-0)#	
	SONET Interface	FTOS(conf-if-so-0/0)#	
	VLAN Interface	FTOS(conf-if-vl-0)#	
IP ACCESS-LIST	STANDARD ACCESS-LIST	FTOS(config-std-nacl)#	**ip access-list standard**
	EXTENDED ACCESS-LIST	FTOS(config-ext-nacl)#	**ip access-list extended**
	IP COMMUNITY-LIST	FTOS(config-community-list)#	**ip community-list**
LINE	AUXILIARY	FTOS(config-line-aux)#	**line**
	CONSOLE	FTOS(config-line-console)#	
	VIRTUAL TERMINAL	FTOS(config-line-vty)#	

Table 2-1. FTOS Command Modes

CLI Command Mode		Prompt	Access Command
MAC ACCESS-LIST	STANDARD ACCESS-LIST	FTOS(config-std-macl)#	**mac access-list standard**
	EXTENDED ACCESS-LIST	FTOS(config-ext-macl)#	**mac access-list extended**
	MULTIPLE SPANNING TREE	FTOS(config-mstp)#	**protocol spanning-tree mstp**
	Per-VLAN SPANNING TREE Plus	FTOS(config-pvst)#	**protocol spanning-tree pvst**
	PREFIX-LIST	FTOS(conf-nprefixl)#	**ip prefix-list**
	RAPID SPANNING TREE	FTOS(config-rstp)#	**protocol spanning-tree rstp**
	REDIRECT	FTOS(conf-redirect-list)#	**ip redirect-list**
	ROUTE-MAP	FTOS(config-route-map)#	**route-map**
	ROUTER BGP	FTOS(conf-router_bgp)#	**router bgp**
	ROUTER ISIS	FTOS(conf-router_isis)#	**router isis**
	ROUTER OSPF	FTOS(conf-router_ospf)#	**router ospf**
	ROUTER RIP	FTOS(conf-router_rip)#	**router rip**
	SPANNING TREE	FTOS(config-span)#	**protocol spanning-tree 0**
	TRACE-LIST	FTOS(conf-trace-acl)#	**ip trace-list**

Figure 2-3 illustrates how to change the command mode from CONFIGURATION mode to PROTOCOL SPANNING TREE.

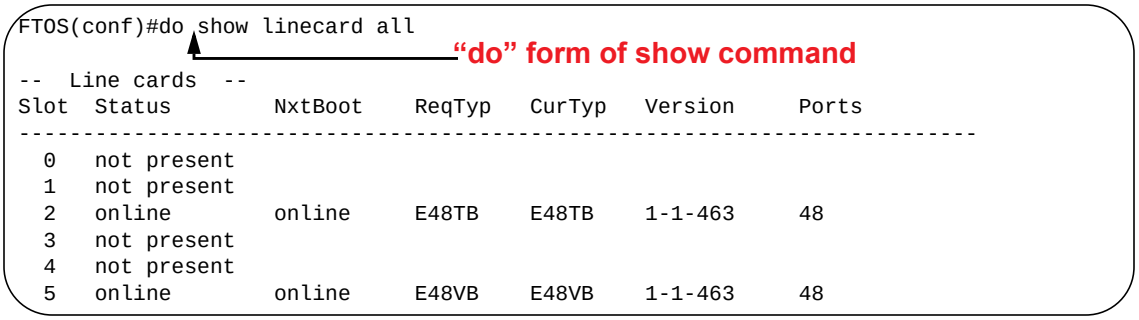
Figure 2-3. Changing CLI Modes

The do Command

Enter an EXEC mode command from any CONFIGURATION mode (CONFIGURATION, INTERFACE, SPANNING TREE, etc.) without returning to EXEC mode by preceding the EXEC mode command with the command **do**. Figure 2-4 illustrates the **do** command.

Note: The following commands cannot be modified by the **do** command: **enable, disable, exit**, and **configure**.

Figure 2-4. Using the do Command

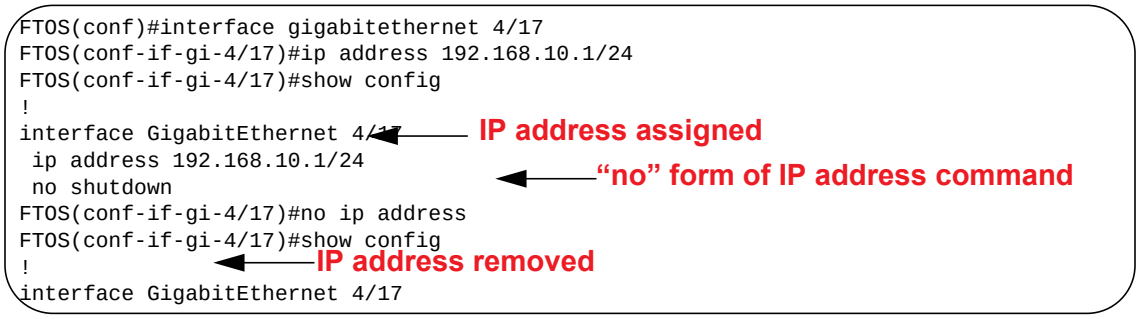


Undoing Commands

When you enter a command, the command line is added to the running configuration file. Disable a command and remove it from the running-config by entering the original command preceded by the command **no**. For example, to delete an ip address configured on an interface, use the **no ip address** *ip-address* command, as shown in Figure 2-5.

Note: Use the **help** or **?** command as discussed in Obtaining Help command to help you construct the "no" form of a command.

Figure 2-5. Undoing a command with the no Command



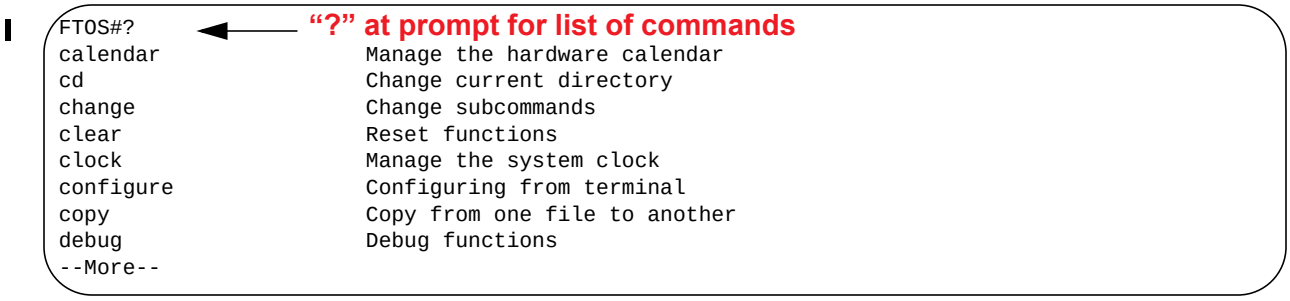
Layer 2 protocols are disabled by default. Enable them using the **no disable** command. For example, in PROTOCOL SPANNING TREE mode, enter **no disable** to enable Spanning Tree.

Obtaining Help

Obtain a list of keywords and a brief functional description of those keywords at any CLI mode using the **?** or **help** command:

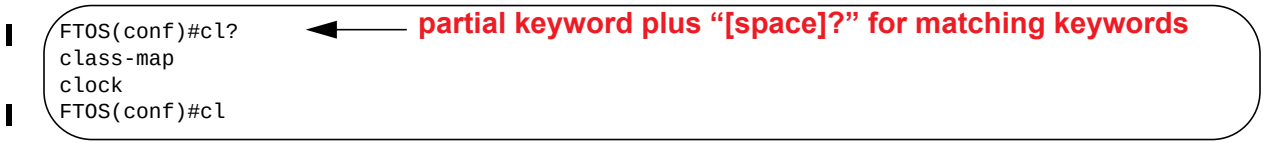
- Enter **?** at the prompt or after a keyword to list the keywords available in the current mode.
  - **?** after a prompt lists all of the available keywords. The output of this command is the same for the **help** command.

Figure 2-6. ? Command Example



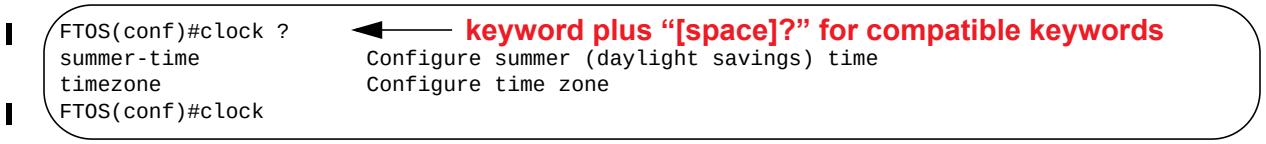
  - **?** after a partial keyword lists all of the keywords that begin with the specified letters.

Figure 2-7. Keyword? Command Example



  - A keyword followed by [space]**?** lists all of the keywords that can follow the specified keyword.

Figure 2-8. Keyword ? Command Example



Entering and Editing Commands

When entering commands:

- The CLI is not case sensitive.
- You can enter partial CLI keywords.
  - You must enter the minimum number of letters to uniquely identify a command. For example, **cl** cannot be entered as a partial keyword because both the **clock** and **class-map** commands begin with the letters "cl." **clo**, however, can be entered as a partial keyword because only one command begins with those three letters.
- The TAB key auto-completes keywords in commands. You must enter the minimum number of letters to uniquely identify a command.

- The UP and DOWN arrow keys display previously entered commands (see Command History).
- The BACKSPACE and DELETE keys erase the previous letter.
- Key combinations are available to move quickly across the command line, as described in Table 2-2.

Table 2-2. Short-Cut Keys and their Actions

Key Combination	Action
CNTL-A	Moves the cursor to the beginning of the command line.
CNTL-B	Moves the cursor back one character.
CNTL-D	Deletes character at cursor.
CNTL-E	Moves the cursor to the end of the line.
CNTL-F	Moves the cursor forward one character.
CNTL-I	Completes a keyword.
CNTL-K	Deletes all characters from the cursor to the end of the command line.
CNTL-L	Re-enters the previous command.
CNTL-N	Return to more recent commands in the history buffer after recalling commands with CTRL-P or the UP arrow key.
CNTL-P	Recalls commands, beginning with the last command
CNTL-R	Re-enters the previous command.
CNTL-U	Deletes the line.
CNTL-W	Deletes the previous word.
CNTL-X	Deletes the line.
CNTL-Z	Ends continuous scrolling of command outputs.
Esc B	Moves the cursor back one word.
Esc F	Moves the cursor forward one word.
Esc D	Deletes all characters from the cursor to the end of the word.

Command History

FTOS maintains a history of previously-entered commands for each mode. For example:

- When you are in EXEC mode, the UP and DOWN arrow keys display the previously-entered EXEC mode commands.
- When you are in CONFIGURATION mode, the UP or DOWN arrows keys recall the previously-entered CONFIGURATION mode commands.

Filtering show Command Outputs

Filter the output of a **show** command to display specific information by adding **|** [**except** | **find** | **grep | no-more | save**] *specified_text* after the command. The variable *specified_text* is the text for which you are filtering and it IS case sensitive unless the **ignore-case** sub-option is implemented.

Starting with FTOS 7.8.1.0, the **grep** command accepts an **ignore-case** sub-option that forces the search to case-*in*sensitive. For example, the commands:

- **show run | grep Ethernet** returns a search result with instances containing a capitalized "Ethernet," such as `interface GigabitEthernet 0/0.`
- **show run | grep ethernet** would not return that search result because it only searches for instances containing a non-capitalized "ethernet."

Executing the command **show run | grep Ethernet ignore-case** would return instances containing both "Ethernet" and "ethernet."

- **grep** displays only the lines containing specified text. Figure 2-9 shows this command used in combination with the command **show linecard all**.

Figure 2-9. Filtering Command Outputs with the grep Command

```
FTOS(conf)#do show linecard all | grep 0
  0   not present
```

Note: FTOS accepts a space or no space before and after the pipe. To filter on a phrase with spaces, underscores, or ranges, enclose the phrase with double quotation marks.

- **except** displays text that does not match the specified text. Figure 2-10 shows this command used in combination with the command **show linecard all**.

Figure 2-10. Filtering Command Outputs with the except Command

```
FTOS#show linecard all | except 0

--  Line cards  --
Slot  Status        NxtBoot    ReqTyp   CurTyp   Version     Ports
------------------------------------------------------------------------------
  2   not present
  3   not present
  4   not present
  5   not present
  6   not present
```

- **find** displays the output of the show command beginning from the first occurrence of specified text Figure 2-11 shows this command used in combination with the command **show linecard all**.

Figure 2-11. Filtering Command Outputs with the find Command

```
FTOS(conf)#do show linecard all | find 0
  0   not present
  1   not present
  2   online        online     E48TB    E48TB    1-1-463     48
  3   not present
  4   not present
  5   online        online     E48VB    E48VB    1-1-463     48
  6   not present
  7   not present
```

- **display** displays additional configuration information.
- **no-more** displays the output all at once rather than one screen at a time. This is similar to the command **terminal length** except that the **no-more** option affects the output of the specified command only.
- **save** copies the output to a file for future reference.

Note: You can filter a single command output multiple times. The save option should be the last option entered. For example:

FTOS# *command* **| grep** *regular-expression* **| except** *regular-expression* **| grep** *other-regular-expression* **| find** *regular-expression* **| save**

Multiple Users in Configuration mode

FTOS notifies all users in the event that there are multiple users logged into CONFIGURATION mode. A warning message indicates the username, type of connection (console or vty), and in the case of a vty connection, the IP address of the terminal on which the connection was established. For example:

- On the system that telnets into the switch, Message 1 appears:

Message 1 Multiple Users in Configuration mode Telnet Message

```
% Warning: The following users are currently configuring the system:
User "<username>" on line console0
```

- On the system that is connected over the console, Message 2 appears:

Message 2 Multiple Users in Configuration mode Telnet Message

```
% Warning: User "<username>" on line vty0 "10.11.130.2" is in configuration mode
```

If either of these messages appears, Dell Force10 recommends that you coordinate with the users listed in the message so that you do not unintentionally overwrite each other's configuration changes.

3

Getting Started

This chapter contains the following major sections:

When you power up the chassis, the system performs a Power-On Self Test (POST) during which Route Processor Module (RPM), Switch Fabric Module (SFM), and line card status LEDs blink green.The system then loads FTOS and boot messages scroll up the terminal window during this process. No user interaction is required if the boot process proceeds without interruption.

When the boot process is complete, the RPM and line card status LEDs remain online (green), and the console monitor displays the Force10 banner and EXEC mode prompt, as shown in Figure 3-1.

For details on using the Command Line Interface (CLI), see the Accessing the Command Line section in Chapter 1, Configuration Fundamentals, on page 47.

Figure 3-1. Completed Boot Process

```
                                                          .*************.
                                                        .#  ####   #######.
 ########  #######   #########     ######## ########   .#. ###### ############.
 ###     ###     ##  ###    ###  ####       ###      .##. ## ### ####       ###.
 ###    ###      ### ###    ### ###        ###      *#.    ### ###           #*
 ###    ###       ## ###   #### ###        ######## *#    -## ###           #*
 ###### ###       ## #########  ###        ######## *#    ### ##            #*
 ###    ###       ## ###  ####  ###        ###      *#   #### ###           #*
 ###     ###     ### ###   #### ####       ###      *#. ####  ###          ###*
 ###      ###   ###  ###     ### #####   ## ######## .#.#####   ####      #### .
 ###        #####    ###      ###   ######  ########   .######  ############ .
                                                        .#      ######### .
                                                          `*************'


                  Copyright 1999-2006 Force10 Networks, Inc.

 + Force10 Networks, Inc.
 + CPU: DB-MV64460-BP/IBM750Fx (2.3)
 + Version: VxWorks5.5.1
 + Memory Size: 1038876672 bytes.
 + BSP Version: 1.2/1.3.6
 + Creation Date : Jan  2 2007

nvDrvInit: nvDrvErase passed
-> 00:00:10: %RPM0-U:CP %RAM-6-ELECTION_ROLE: RPM0 is transitioning to Primary RPM.
00:00:11: %RPM0-P:CP %CHMGR-2-FAN_BAD: Minor alarm: some fans in fan tray 0 are down
00:00:11: %RPM0-P:CP %CHMGR-5-CARDDETECTED: Line card 1 present

 DSA Card Init
00:00:11: %RPM0-P:CP POEMGR-4-POE_POWER_USAGE_ABOVE_THRESHOLD: Inline power used is exceeded 90% o
available inline power
00:00:12: %RPM0-P:CP %CHMGR-5-CARDDETECTED: Line card 2 present
00:00:12: %RPM0-P:CP %TSM-6-SFM_SWITCHFAB_STATE: Switch Fabric: UP
00:00:12: %RPM0-P:CP %TSM-6-SFM_FULL_PARTIAL_STATE: SW_FAB_UP_1 SFM in the system
00:00:13: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Ma 0/0

00:01:27: %RPM0-P:CP %CHMGR-5-CHECKIN: Checkin from line card 1 (type E48TB, 48 ports)
00:01:27: %RPM0-P:CP %CHMGR-5-CHECKIN: Checkin from line card 2 (type E48TB, 48 ports)
00:01:28: %RPM0-P:CP %CHMGR-5-LINECARDUP: Line card 1 is up
00:01:28: %RPM0-P:CP %CHMGR-5-LINECARDUP: Line card 2 is up
00:01:36: %RPM0-P:CP %RAM-5-RPM_STATE: RPM0 is in Active State.
00:01:36: %RPM0-P:CP %CHMGR-5-CHAS_READY: Chassis ready

00:01:37: %RPM0-P:CP %SEC-5-LOGIN_SUCCESS: Login successful for user  on line console
FTOS>
```

Default Configuration

A version of FTOS is pre-loaded onto the chassis, however the system is not configured when you power up for the first time (except for the default hostname, which is Force10). You must configure the system using the CLI.

Configure a Host Name

The host name appears in the prompt. The default host name is **force10**.

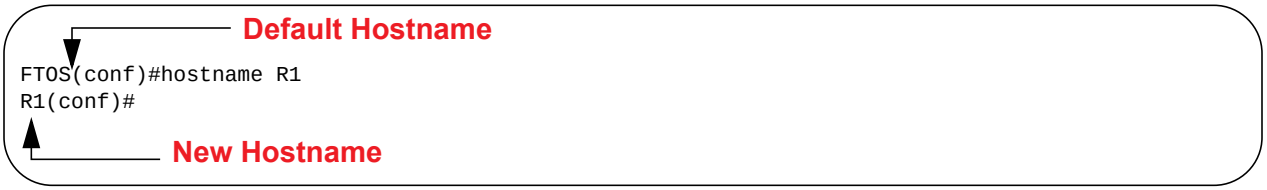
- Host names must start with a letter and end with a letter or digit.
- Characters within the string can be letters, digits, and hyphens.

To configure a host name:

Step	Task	Command Syntax	Command Mode
1	Create a new host name.	**hostname** *name*	CONFIGURATION

Figure 3-2 illustrates the **hostname** command.

Figure 3-2. Configuring a Hostname



Access the System Remotely

You can configure the system to access it remotely by Telnet. The method for configuring the C-Series and E-Series for Telnet access is different from S-Series.

- The C-Series and E-Series have a dedicated management port and a management routing table that is separate from the IP routing table.
- The S-Series does not have a dedicated management port, but is managed from any port. It does not have a separate management routing table.

Access the C-Series and E-Series Remotely

Note: Use this process for the S60 system.

Configuring the system for Telnet is a three-step process:

1. Configure an IP address for the management port. See Configure the Management Port IP Address.
2. Configure a management route with a default gateway. See Configure a Management Route.
3. Configure a username and password. See Configure a Username and Password.

Configure the Management Port IP Address

Assign IP addresses to the management ports in order to access the system remotely.

Note: Assign different IP addresses to each RPM's management port.

To configure the management port IP address:

Step	Task	Command Syntax	Command Mode
1	Enter INTERFACE mode for the Management port.	**interface ManagementEthernet** *slot/port* • *slot* range: 0 to 1 • *port* range: 0	CONFIGURATION
2	Assign an IPv4 or IPv6 address to the interface.	**ip address** {*ipv4-address* | *ipv6-address*}*/mask* • *ipv4-address:* an address in dotted-decimal format (A.B.C.D). • *ipv6-address:* an address in hexadecimal format (X:X:X:X::X). • *mask:* a subnet mask in /prefix-length format (/xx).	INTERFACE
3	Enable the interface.	**no shutdown**	INTERFACE

Configure a Management Route

Define a path from the system to the network from which you are accessing the system remotely. Management routes are separate from IP routes and are only used to manage the system through the management port.

To configure a management route:

Step	Task	Command Syntax	Command Mode
1	Configure an IPv4 or IPv6 management route to the network from which you are accessing the system.	**management route** {*ipv4-address* | *ipv6-address*}/*mask gateway* • *ip-address:* the network address in dotted-decimal format (A.B.C.D). • *mask:* a subnet mask in /prefix-length format (/xx). • *gateway*: the next hop for network traffic originating from the management port.	CONFIGURATION

Configure a Username and Password

Configure a system username and password to access the system remotely.

To configure a username and password:

Step	Task	Command Syntax	Command Mode
1	Configure a username and password to access the system remotely.	**username** *username* **password** [*encryption-type*] *password* *encryption-type* specifies how you are inputting the password, is 0 by default, and is not required. • 0 is for inputting the password in clear text. • 7 is for inputting a password that is already encrypted using a Type 7 hash. Obtaining the encrypted password from the configuration of another Dell Force10 system.	CONFIGURATION

Access the S-Series Remotely

The S-Series does not have a dedicated management port nor a separate management routing table. Configure any port on the S-Series to be the port through which you manage the system and configure an IP route to that gateway.

Note: The S60 system uses management ports and should be configured similar to the C-Series and E-Series systems. Refer to Access the C-Series and E-Series Remotely

Configuring the system for Telnet access is a three-step process:

1. Configure an IP address for the port through which you will manage the system using the command **ip address** from INTERFACE mode, as shown in Figure 3-3.
2. Configure a IP route with a default gateway using the command **ip route** from CONFIGURATION mode, as shown in Figure 3-3.
3. Configure a username and password using the command **username** from CONFIGURATION mode, as shown in Figure 3-3.

Figure 3-3. Configuring the S-Series for Remote Access

```
R5(conf)#int gig 0/48
R5(conf-if-gi-0/48)#ip address 10.11.131.240
R5(conf-if-gi-0/48)#show config
!
interface GigabitEthernet 0/48
 ip address 10.11.131.240/24
 no shutdown
R5(conf-if-gi-0/48)#exit
R5(conf)#ip route 10.11.32.0/23 10.11.131.254
R5(conf)#username admin pass force10
```

Configure the Enable Password

The EXEC Privilege mode is accessed by the **enable** command. Configure a password as a basic security measure. When using a console connection, EXEC Privilege mode is unrestricted by default; it cannot be reached by a VTY connection if no password is configured. There are two types of **enable** passwords:

- **enable password** stores the password in the running/startup configuration using a DES encryption method.
- **enable secret** is stored in the running/startup configuration in using a stronger, MD5 encryption method.

Dell Force10 recommends using the **enable secret** password.

To configure an enable password:

Task	Command Syntax	Command Mode
Create a password to access EXEC Privilege mode.	**enable** [**password | secret**] [**level** *level*] [*encryption-type*] *password* *level* is the privilege level, is 15 by default, and is not required. *encryption-type* specifies how you are inputting the password, is 0 by default, and is not required. • 0 is for inputting the password in clear text. • 7 is for inputting a password that is already encrypted using a DES hash. Obtain the encrypted password from the configuration file of another Dell Force10 system. • 5 is for inputting a password that is already encrypted using an MD5 hash. Obtain the encrypted password from the configuration file of another Dell Force10 system.	CONFIGURATION

Configuration File Management

Files can be stored on and accessed from various storage media. Rename, delete, and copy files on the system from the EXEC Privilege mode.

The E-Series EtherScale platform architecture uses MMC cards for both the internal and external Flash memory. MMC cards support a maximum of 100 files. The E-Series TeraScale and ExaScale platforms architecture use Compact Flash for the internal and external Flash memory. It has a space limitation but does not limit the number of files it can contain.

Note: Using flash memory cards in the system that have not been approved by Dell Force10 can cause unexpected system behavior, including a reboot.

Copy Files to and from the System

The command syntax for copying files is similar to UNIX. The **copy** command uses the format **copy** *source-file-url destination-file-url*.

Note: See the *FTOS Command Reference* for a detailed description of the **copy** command.

- To copy a local file to a remote system, combine the *file-origin* syntax for a local file location with the *file-destination* syntax for a remote file location shown in Table 3-1.
- To copy a remote file to Dell Force10 system, combine the *file-origin* syntax for a remote file location with the *file-destination* syntax for a local file location shown in Table 3-1.

Table 3-1. Forming a copy Command

	source-file-url **Syntax**	*destination-file-url* **Syntax**
Local File Location		
Internal flash:		
primary RPM	**copy flash://***filename*	**flash://***filename*
standby RPM	**copy rpm{0|1}flash://***filename*	**rpm{0|1}flash://***filename*
External flash:		
primary RPM	**copy rpm{0|1}slot0://***filename*	**rpm{0|1}slot0://***filename*
standby RPM	**copy rpm{0|1}slot0://***filename*	**rpm{0|1}slot0://***filename*
USB Drive (E-Series ExaScale only)		
USB drive on RPM0	**copy rpm0usbflash://***filepath*	**rpm0usbflash://***filename*
External USB drive	**copy usbflash://***filepath*	**usbflash://***filename*
Remote File Location		
Note: FTOS supports IPv4 and IPv6 addressing for FTP, TFTP, and SCP (in the *hostip* field).		
FTP server	**copy ftp://***username***:***password*@{*hostip* | *hostname*}/*filepath*/*filename*	**ftp://***username***:***password*@{*hostip* | *hostname*}/*filepath*/*filename*
TFTP server	**copy tftp:**//{*hostip* | *hostname*}*/filepath/filename*	**tftp:**//{*hostip* | *hostname*}*/filepath/filename*
SCP server	**copy scp://**{*hostip* | *hostname*}*/filepath/filename*	**scp://**{*hostip* | *hostname*}*/filepath/filename*

Important Points to Remember

- You may not copy a file from one remote system to another.
- You may not copy a file from one location to the same location.
- The internal flash memories on the RPMs are synchronized whenever there is a change, but only if both RPMs are running the same version of FTOS.
- When copying to a server, a hostname can only be used if a DNS server is configured.

- The **usbflash** and **rpm0usbflash** commands are supported on E-Series ExaScale platform only. Refer to the FTOS Release Notes for a list of approved USB vendors.

Figure 3-4 shows an example of using the **copy** command to save a file to an FTP server.

Figure 3-4. Saving a file to a Remote System

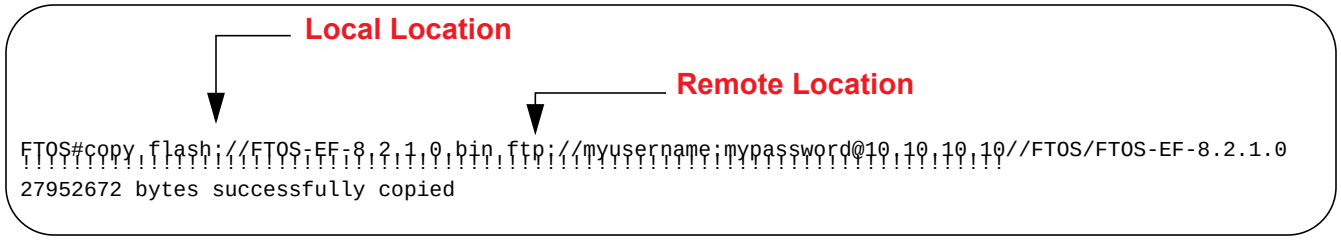
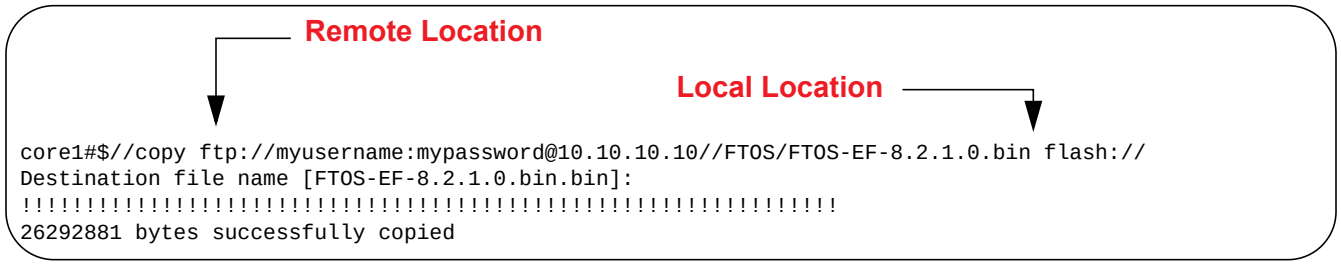


Figure 3-5 shows an example of using the **copy** command to import a file to the Dell Force10 system from an FTP server.

Figure 3-5. Saving a file to a Remote System



Save the Running-configuration

The running-configuration contains the current system configuration. Dell Force10 recommends that you copy your running-configuration to the startup-configuration. The system uses the startup-configuration during boot-up to configure the system. The startup-configuration is stored in the internal flash on the primary RPM by default, but it can be saved onto an external flash (on an RPM) or a remote server.

To save the running-configuration:

Note: The commands in this section follow the same format as those in Copy Files to and from the System on page 51 but use the filenames *startup-configuration* and *running-configuration*. These commands assume that current directory is the internal flash, which is the system default.

Task	Command Syntax	Command Mode
Save the running-configuration to:		
the startup-configuration on the internal flash of the primary RPM	**copy running-config startup-config**	EXEC Privilege
the internal flash on an RPM	**copy running-config rpm**{**0|1**}**flash://***filename*	
the external flash of an RPM	**copy running-config rpm**{**0|1**}**slot0://***filename*	
an FTP server	**copy running-config ftp://***username***:***password*@{*hostip* | *hostname*}*/filepath/filename*	
a TFTP server	**copy running-config tftp://**{*hostip* | *hostname*}*/filepath/filename*	
an SCP server	**copy running-config scp://**{*hostip* | *hostname*}*/filepath/filename*	
Save the running-configuration to the startup-configuration on the internal flash of the primary RPM. Then copy the new startup-config file to the external flash of the primary RPM.	**copy running-config startup-config duplicate**	EXEC Privilege

Note: The internal flash memories on the RPMs are synchronized whenever there is a change, but only if the RPMs are running the same version of FTOS.
Note: FTOS supports IPv4 and IPv6 addressing for FTP, TFTP, and SCP (in the *hostip* field).

Note: When copying to a server, a hostname can only be used if a DNS server is configured.

FTOS Behavior: If you create a startup-configuration on an RPM and then move the RPM to another chassis, the startup-configuration is stored as a backup file (with the extension *.bak*), and a new, empty startup-configuration file is created. To restore your original startup-configuration in this situation, overwrite the new startup-configuration with the original one using the command **copy** *startup-config.bak startup-config*.

View Files

File information and content can only be viewed on local file systems.

To view a list of files on the internal or external Flash:

Step	Task	Command Syntax	Command Mode
1	View a list of files on:		
	the internal flash of an RPM	**dir flash:**	EXEC Privilege
	the external flash of an RPM	**dir slot:**	

The output of the command **dir** also shows the read/write privileges, size (in bytes), and date of modification for each file, as shown in Figure 3-6.

Figure 3-6. Viewing a List of Files in the Internal Flash

```
FTOS#dir
Directory of flash:

  1  drw-      32768   Jan 01 1980 00:00:00  .
  2  drwx        512   Jul 23 2007 00:38:44  ..
  3  drw-       8192   Mar 30 1919 10:31:04  TRACE_LOG_DIR
  4  drw-       8192   Mar 30 1919 10:31:04  CRASH_LOG_DIR
  5  drw-       8192   Mar 30 1919 10:31:04  NVTRACE_LOG_DIR
  6  drw-       8192   Mar 30 1919 10:31:04  CORE_DUMP_DIR
  7  d---       8192   Mar 30 1919 10:31:04  ADMIN_DIR
  8  -rw-   33059550   Jul 11 2007 17:49:46  FTOS-EF-7.4.2.0.bin
  9  -rw-   27674906   Jul 06 2007 00:20:24  FTOS-EF-4.7.4.302.bin
 10  -rw-   27674906   Jul 06 2007 19:54:52  boot-image-FILE
 11  drw-       8192   Jan 01 1980 00:18:28  diag
 12  -rw-       7276   Jul 20 2007 01:52:40  startup-config.bak
 13  -rw-       7341   Jul 20 2007 15:34:46  startup-config
 14  -rw-   27674906   Jul 06 2007 19:52:22  boot-image
 15  -rw-   27674906   Jul 06 2007 02:23:22  boot-flash
--More--
```

To view the contents of a file:

Step	Task	Command Syntax	Command Mode
1	View the:		
	contents of a file in the internal flash of an RPM	**show file rpm**{**0**|**1**}**flash:**//*filename*	
	contents of a file in the external flash of an RPM	**show file rpm**{**0**|**1**}**slot0:**//*filename*	EXEC Privilege
	running-configuration	**show running-config**	
	startup-configuration	**show startup-config**	

View Configuration Files

Configuration files have three commented lines at the beginning of the file, as shown in Figure 3-7, to help you track the last time any user made a change to the file, which user made the changes, and when the file was last saved to the startup-configuration.

In the running-configuration file, if there is a difference between the timestamp on the "Last configuration change," and "Startup-config last updated," then you have made changes that have not been saved and will not be preserved upon a system reboot.

Figure 3-7. Tracking Changes with Configuration Comments

```
FTOS#show running-config
Current Configuration ...
! Version 8.2.1.0
! Last configuration change at Thu Apr 3 23:06:28 2008 by admin
! Startup-config last updated at Thu Apr 3 23:06:55 2008 by admin
!
boot system rpm0 primary flash://FTOS-EF-8.2.1.0.bin
boot system rpm0 secondary flash://FTOS-EF-7.8.1.0.bin
boot system rpm0 default flash://FTOS-EF-7.7.1.1.bin
boot system rpm1 primary flash://FTOS-EF-7.8.1.0.bin
boot system gateway 10.10.10.100
--More--
```

File System Management

The Dell Force10 system can use the internal Flash, external Flash, or remote devices to store files. It stores files on the internal Flash by default but can be configured to store files elsewhere.

To view file system information:

Task	Command Syntax	Command Mode
View information about each file system.	**show file-systems**	EXEC Privilege

The output of the command **show file-systems** (Figure 3-8) shows the total capacity, amount of free memory, file structure, media type, read/write privileges for each storage device in use.

Figure 3-8. show file-systems Command Example

```
FTOS#show file-systems
Size(b)     Free(b)      Feature      Type   Flags  Prefixes
    520962048   213778432     dosFs2.0 USERFLASH      rw  flash:
    127772672    21936128     dosFs2.0 USERFLASH      rw  slot0:
            -           -            -   network      rw  ftp:
            -           -            -   network      rw  tftp:
            -           -            -   network      rw  scp:
```

You can change the default file system so that file management commands apply to a particular device or memory.

To change the default storage location:

Task	Command Syntax	Command Mode
Change the default directory.	**cd** *directory*	EXEC Privilege

In Figure 3-9, the default storage location is changed to the external Flash of the primary RPM. File management commands then apply to the external Flash rather than the internal Flash.

Figure 3-9. Alternative Storage Location

```
FTOS#cd slot0:
FTOS#copy running-config test
FTOS#copy run test                                    <-- No File System Specified
!
7419 bytes successfully copied
FTOS#dir
Directory of slot0:

  1  drw-      32768   Jan 01 1980 00:00:00  .
  2  drwx        512   Jul 23 2007 00:38:44  ..
  3  ----          0   Jan 01 1970 00:00:00  DCIM
  4  -rw-       7419   Jul 23 2007 20:44:40  test   <-- File Saved to External Flash
  5  ----          0   Jan 01 1970 00:00:00  BT
  6  ----          0   Jan 01 1970 00:00:00  200702~1VSN
  7  ----          0   Jan 01 1970 00:00:00  G
  8  ----          0   Jan 01 1970 00:00:00  F
  9  ----          0   Jan 01 1970 00:00:00  F

slot0: 127772672 bytes total (21927936 bytes free)
```

View command history

The command-history trace feature captures all commands entered by all users of the system with a time stamp and writes these messages to a dedicated trace log buffer. The system generates a trace message for each executed command. No password information is saved to the file.

To view the command-history trace, use the **show command-history** command, as shown in Figure 487.

Figure 3-10. Command Example show command-history

```
FTOS#show command-history
[12/5 10:57:8]: CMD-(CLI):service password-encryption
[12/5 10:57:12]: CMD-(CLI):hostname FTOS
[12/5 10:57:12]: CMD-(CLI):ip telnet server enable
[12/5 10:57:12]: CMD-(CLI):line console 0
[12/5 10:57:12]: CMD-(CLI):line vty 0 9
[12/5 10:57:13]: CMD-(CLI):boot system rpm0 primary flash://FTOS-CB-1.1.1.2E2.bin
```

Upgrading and Downgrading FTOS

Note: To upgrade or downgrade FTOS, see the release notes for the version you want to load on the system.

4

System Management

System Management is supported on platforms: C E S

This chapter explains the different protocols or services used to manage the Dell Force10 system including:

Configure Privilege Levels

Privilege levels restrict access to commands based on user or terminal line. There are 16 privilege levels, of which three are pre-defined. The default privilege level is 1.

- **Level 0—**Access to the system begins at EXEC mode, and EXEC mode commands are limited to **enable**, **disable**, and **exit**.
- **Level 1**—Access to the system begins at EXEC mode, and all commands are available.
- **Level 15**—Access to the system begins at EXEC Privilege mode, and all commands are available.

Create a Custom Privilege Level

Custom privilege levels start with the default EXEC mode command set. You can then customize privilege levels 2-14 by:

- restricting access to an EXEC mode command
- moving commands from EXEC Privilege to EXEC mode
- restricting access

A user can access all commands at his privilege level and below.

Removing a command from EXEC mode

Remove a command from the list of available commands in EXEC mode for a specific privilege level using the command **privilege exec** from CONFIGURATION mode. In the command, specify a level *greater* than the level given to a user or terminal line, followed by the first keyword of each command to be restricted.

Move a command from EXEC Privilege mode to EXEC mode

Move a command from EXEC Privilege to EXEC mode for a privilege level using the command **privilege exec** from CONFIGURATION mode. In the command, specify the privilege level of the user or terminal line, and specify *all* keywords in the command to which you want to allow access.

Allow Access to CONFIGURATION mode commands

Allow access to CONFIGURATION mode using the command **privilege exec level** *level* **configure** from CONFIGURATION mode. A user that enters CONFIGURATION mode remains at his privilege level, and has access to only two commands, **end** and **exit**. You must individually specify each CONFIGURATION mode command to which you want to allow access using the command **privilege configure level** *level*. In the command, specify the privilege level of the user or terminal line, and specify *all* keywords in the command to which you want to allow access.

Allow Access to INTERFACE, LINE, ROUTE-MAP, and ROUTER mode

1. Similar to allowing access to CONFIGURATION mode, to allow access to INTERFACE, LINE, ROUTE-MAP, and ROUTER modes, you must first allow access to the command that enters you into the mode. For example, allow a user to enter INTERFACE mode using the command **privilege configure level** *level* **interface gigabitethernet**
2. Then, individually identify the INTERFACE, LINE, ROUTE-MAP or ROUTER commands to which you want to allow access using the command **privilege** {**interface** | **line** | **route-map** | **router**} **level** *level*. In the command, specify the privilege level of the user or terminal line, and specify *all* keywords in the command to which you want to allow access.

The following table lists the configuration tasks you can use to customize a privilege level:

Task	Command Syntax	Command Mode
Remove a command from the list of available commands in EXEC mode.	**privilege exec level** *level* {*command* ||...|| *command*}	CONFIGURATION
Move a command from EXEC Privilege to EXEC mode.	**privilege exec level** *level* {*command* ||...|| *command*}	CONFIGURATION
Allow access to CONFIGURATION mode.	**privilege exec level** *level* **configure**	CONFIGURATION

Task	Command Syntax	Command Mode
Allow access to INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode. Specify *all* keywords in the command.	**privilege configure level** *level* {**interface** | **line** | **route-map** | **router**} {*command-keyword* ||...|| *command-keyword*}	CONFIGURATION
Allow access to a CONFIGURATION, INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode command.	**privilege** {**configure** |**interface** | **line** | **route-map** | **router**} **level** *level* {*command* ||...|| *command*}	CONFIGURATION

The configuration in Figure 4-1 creates privilege level 3. This level:

- removes the **resequence** command from EXEC mode by requiring a minimum of privilege level 4,
- moves the command **capture bgp-pdu max-buffer-size** from EXEC Privilege to EXEC mode by, requiring a minimum privilege level 3, which is the configured level for VTY 0,
- allows access to CONFIGURATION mode with the **banner** command, and
- allows access to INTERFACE and LINE modes are allowed with no commands.

Figure 4-1. Create a Custom Privilege Level

```
FTOS(conf)#do show run priv
!
privilege exec level 3 capture
privilege exec level 3 configure
privilege exec level 4 resequence
privilege exec level 3 capture bgp-pdu
privilege exec level 3 capture bgp-pdu max-buffer-size
privilege configure level 3 line
privilege configure level 3 interface
FTOS(conf)#do telnet 10.11.80.201
[telnet output omitted]
FTOS#show priv
Current privilege level is 3.
FTOS#?
capture                 Capture packet
configure               Configuring from terminal
disable                 Turn off privileged commands
enable                  Turn on privileged commands
exit                    Exit from the EXEC
ip                      Global IP subcommands
monitor                 Monitoring feature
mtrace                  Trace reverse multicast path from destination to source
ping                    Send echo messages
quit                    Exit from the EXEC
show                    Show running system information
[output omitted]
FTOS#config
[output omitted]
FTOS(conf)#do show priv
Current privilege level is 3.
FTOS(conf)#?
end                     Exit from configuration mode
exit                    Exit from configuration mode
interface               Select an interface to configure
line                    Configure a terminal line
linecard                Set line card type
FTOS(conf)#interface ?
fastethernet            Fast Ethernet interface
gigabitethernet         Gigabit Ethernet interface
loopback                Loopback interface
managementethernet      Management Ethernet interface
null                    Null interface
port-channel            Port-channel interface
range                   Configure interface range
sonet                   SONET interface
tengigabitethernet      TenGigabit Ethernet interface
vlan                    VLAN interface
FTOS(conf)#interface gigabitethernet 1/1
FTOS(conf-if-gi-1/1)#?
end                             Exit from configuration mode
exit                            Exit from interface configuration mode
FTOS(conf-if-gi-1/1)#exit
FTOS(conf)#line ?
aux                             Auxiliary line
console                         Primary terminal line
vty                             Virtual terminal
FTOS(conf)#line vty 0
FTOS(config-line-vty)#?
exit                            Exit from line configuration mode
FTOS(config-line-vty)#
```

Apply a Privilege Level to a Username

To set a privilege level for a user:

Task	Command Syntax	Command Mode
Configure a privilege level for a user.	**username** *username* **privilege** *level*	CONFIGURATION

Apply a Privilege Level to a Terminal Line

To set a privilege level for a terminal line:

Task	Command Syntax	Command Mode
Configure a privilege level for a terminal line.	**privilege level** *level*	LINE

Note: When you assign a privilege level between 2 and 15, access to the system begins at EXEC mode, but the prompt is *hostname*#, rather than *hostname*>.

Configure Logging

FTOS tracks changes in the system using event and error messages. By default, FTOS logs these messages on:

- the internal buffer
- console and terminal lines, and
- any configured syslog servers

Disable Logging

To disable logging:

Task	Command Syntax	Command Mode
Disable all logging except on the console.	**no logging on**	CONFIGURATION
Disable logging to the logging buffer.	**no logging buffer**	CONFIGURATION
Disable logging to terminal lines.	**no logging monitor**	CONFIGURATION
Disable console logging.	**no logging console**	CONFIGURATION

Log Messages in the Logging Buffer

All error messages, except those beginning with %BOOTUP (Message 1), are log in the internal buffer.

Message 1 BootUp Events

```
%BOOTUP:RPM0:CP %PORTPIPE-INIT-SUCCESS: Portpipe 0 enabled
```

Configuration Task List for System Log Management

The following list includes the configuration tasks for system log management:

- Disable System Logging on page 62
- Send System Messages to a Syslog Server on page 63

Disable System Logging

By default, logging is enabled and log messages are sent to the logging buffer, all terminal lines, console, and syslog servers.

Enable and disable system logging using the following commands:

Task	Command Syntax	Command Mode
Disable all logging except on the console.	**no logging on**	CONFIGURATION
Disable logging to the logging buffer.	**no logging buffer**	CONFIGURATION
Disable logging to terminal lines.	**no logging monitor**	CONFIGURATION
Disable console logging.	**no logging console**	CONFIGURATION

Send System Messages to a Syslog Server

Send system messages to a syslog server by specifying a server:

Task	Command Syntax	Command Mode
Specify the server to which you want to send system messages. You can configure up to eight syslog servers, which may be IPv4 and/or IPv6 addressed.	**logging** {*ip-address* | *ipv6-address* | *hostname*}	CONFIGURATION

Configure a Unix System as a Syslog Server

Configure a UNIX system as a syslog server by adding the following lines to */etc/syslog.conf* on the Unix system and assigning write permissions to the file.

- on a 4.1 BSD UNIX system, add the line: **local7.debugging /var/log/force10.log**
- on a 5.7 SunOS UNIX system, add the line: **local7.debugging /var/adm/force10.log**

In the lines above, **local7** is the logging facility level and debugging is the severity level.

Change System Logging Settings

You can change the default system logging settings (severity level and the storage location). The default is to log all messages up to debug level.

Task	Command Syntax	Command Mode
Specify the minimum severity level for logging to the logging buffer.	**logging buffered** *level*	CONFIGURATION
Specify the minimum severity level for logging to the console.	**logging console** *level*	CONFIGURATION
Specify the minimum severity level for logging to terminal lines.	**logging monitor** *level*	CONFIGURATION
Specifying the minimum severity level for logging to a syslog server.	**logging trap** *level*	CONFIGURATION
Specify the minimum severity level for logging to the syslog history table.	**logging history** *level*	CONFIGURATION

Task	Command Syntax	Command Mode
Specify the size of the logging buffer. **Note**: When you decrease the buffer size, FTOS deletes all messages stored in the buffer. Increasing the buffer size does not affect messages in the buffer.	**logging buffered** *size*	CONFIGURATION
Specify the number of messages that FTOS saves to its logging history table.	**logging history size** *size*	CONFIGURATION

Display the logging buffer and configuration using the **show logging** command from EXEC Privilege mode, as shown in Figure 4-2.

Display the logging configuration using the **show running-config logging** command from EXEC Privilege mode, as shown in Figure 4-3.

Display the Logging Buffer and the Logging Configuration

Display the current contents of the logging buffer and the logging settings for the system using the **show logging** command from EXEC Privilege mode, as shown in Figure 4-2.

Figure 4-2. show logging Command Example

```
FTOS#show logging
syslog logging: enabled
    Console logging: level Debugging
    Monitor logging: level Debugging
    Buffer logging: level Debugging, 40 Messages Logged, Size (40960 bytes)
    Trap logging: level Informational
%IRC-6-IRC_COMMUP: Link to peer RPM is up
%RAM-6-RAM_TASK: RPM1 is transitioning to Primary RPM.
%RPM-2-MSG:CP1 %POLLMGR-2-MMC_STATE: External flash disk missing in 'slot0:'
%CHMGR-5-CARDDETECTED: Line card 0 present
%CHMGR-5-CARDDETECTED: Line card 2 present
%CHMGR-5-CARDDETECTED: Line card 4 present
%CHMGR-5-CARDDETECTED: Line card 5 present
%CHMGR-5-CARDDETECTED: Line card 8 present
%CHMGR-5-CARDDETECTED: Line card 10 present
%CHMGR-5-CARDDETECTED: Line card 12 present
%TSM-6-SFM_DISCOVERY: Found SFM 0
%TSM-6-SFM_DISCOVERY: Found SFM 1
%TSM-6-SFM_DISCOVERY: Found SFM 2
%TSM-6-SFM_DISCOVERY: Found SFM 3
%TSM-6-SFM_DISCOVERY: Found SFM 4
%TSM-6-SFM_DISCOVERY: Found SFM 5
%TSM-6-SFM_DISCOVERY: Found SFM 6
%TSM-6-SFM_DISCOVERY: Found SFM 7
%TSM-6-SFM_SWITCHFAB_STATE: Switch Fabric: UP
%TSM-6-SFM_DISCOVERY: Found SFM 8
%TSM-6-SFM_DISCOVERY: Found 9 SFMs
%CHMGR-5-CHECKIN: Checkin from line card 5 (type EX1YB, 1 ports)
%TSM-6-PORT_CONFIG: Port link status for LC 5 => portpipe 0: OK portpipe 1: N/A
%CHMGR-5-LINECARDUP: Line card 5 is up
%CHMGR-5-CHECKIN: Checkin from line card 12 (type S12YC12, 12 ports)
%TSM-6-PORT_CONFIG: Port link status for LC 12 => portpipe 0: OK portpipe 1: N/A
%CHMGR-5-LINECARDUP: Line card 12 is up
%IFMGR-5-CSTATE_UP: changed interface Physical state to up: So 12/8
%IFMGR-5-CSTATE_DN: changed interface Physical state to down: So 12/8
```

Configure a UNIX Logging Facility Level

Facility is a message tag used to describe the application or process that submitted the log message. You can save system log messages with a UNIX system logging facility:

Command Syntax	Command Mode	Purpose
logging facility [*facility-type*]	CONFIGURATION	Specify one of the following parameters. • auth (for authorization messages) • cron (for system scheduler messages) • daemon (for system daemons) • kern (for kernel messages) • local0 (for local use) • local1 (for local use) • local2 (for local use) • local3 (for local use) • local4 (for local use) • local5 (for local use) • local6 (for local use) • local7 (for local use). This is the default. • lpr (for line printer system messages) • mail (for mail system messages) • news (for USENET news messages) • sys9 (system use) • sys10 (system use) • sys11 (system use) • sys12 (system use) • sys13 (system use) • sys14 (system use) • syslog (for syslog messages) • user (for user programs) • uucp (UNIX to UNIX copy protocol) The default is local7.

Display non-default settings using the **show running-config logging** command from EXEC mode, as shown in Figure 4-3.

Figure 4-3. show running-config logging Command Example

```
FTOS#show running-config logging
!
logging buffered 524288 debugging
service timestamps log datetime msec
service timestamps debug datetime msec
!
logging trap debugging
logging facility user
logging source-interface Loopback 0
logging 10.10.10.4
FTOS#
```

Synchronize Log Messages

You can configure a terminal line to hold all logs until all command inputs and outputs are complete so that log printing does not interfere when you are performing management tasks. Log synchronization also filters system messages for a specific line based on severity level and limits number of messages that are printed at once.

Step	Task	Command Syntax	Command Mode
1	Enter the LINE mode. Configure the following parameters for the virtual terminal lines: • *number* range: zero (0) to 8. • *end-number* range: 1 to 8. You can configure multiple virtual terminals at one time by entering a *number* followed by an *end-number.*	**line** {**console 0** | **vty** *number* [*end-number*] | **aux 0**}	CONFIGURATION
2	Set a level and the maximum number of messages to be printed. The following parameters are optional: • **level** *severity-level* range: 0 to 7. Default is 2. Use the **all** keyword to include all messages. • *limit* range: 20 to 300. Default is 20.	**logging synchronous** [**level** *severity-level* | **all**] [*limit*]	LINE

Display the logging synchronous configuration using the **show config** command from LINE mode.

Enable Timestamp on Syslog Messages

Syslog messages, by default, do not include a time/date stamp stating when the error or message was created. To have FTOS include a timestamp with the syslog message:

Purpose	Command Syntax	Command Mode
Add timestamp to syslog messages. Specify the following optional parameters: • **datetime**: You can add the keyword localtime to include the **localtime, msec,** and **show-timezone**. If you do not add the keyword **localtime**, the time is UTC. • **uptime**. To view time since the last boot.	**service timestamps** [**log | debug**] [**datetime** [**localtime**] [**msec**] [**show-timezone**] | **uptime**] Default: **uptime**	CONFIGURATION

Display your configuration using the command **show running-config logging** from EXEC Privilege mode, as shown in Figure 4-3.

File Transfer Services

You can configure the system to transfer files over the network using File Transfer Protocol (FTP).

Configuration Task List for File Transfer Services

The following list includes the configuration tasks for file transfer services:

Enable FTP server

To make the system an FTP server:

Task	Command Syntax	Command Mode
Make the system an FTP server.	**ftp-server enable**	CONFIGURATION

Display your FTP configuration using the command **show running-config ftp** from EXEC Privilege mode, as shown in Figure 4-4.

Figure 4-4. show running-config ftp Command Example

```
FTOS#show running ftp
!
ftp-server enable
ftp-server username nairobi password 0 zanzibar
FTOS#
```

Configure FTP server parameters

To configure FTP server parameters:

Task	Command Syntax	Command Mode
Specify the directory for users using FTP to reach the system. The default is the internal flash.	**ftp-server topdir** *dir*	CONFIGURATION
Specify a user name for all FTP users and configure either a plain text or encrypted password. Configure the following optional and required parameters: • *username*: Enter a text string • *encryption-type:* Enter 0 for plain text or 7 for encrypted text. • *password:* Enter a text string.	**ftp-server username** *username* **password** [*encryption-type*] *password*	CONFIGURATION

Note: You cannot use the change directory (**cd**) command until **ftp-server topdir** is configured.

Display your FTP configuration using the command **show running-config ftp** from EXEC Privilege mode, as shown in Figure 4-4.

Configure FTP client parameters

When the system will be an FTP client, configure FTP client parameters:

Task	Command Syntax	Command Mode
Specify a source interface.	**ip ftp source-interface** *interface*	CONFIGURATION
Configure a password.	**ip ftp password** *password*	CONFIGURATION
Enter username to use on FTP client.	**ip ftp username** *name*	CONFIGURATION

Display the FTP configuration using the command **show running-config ftp** from EXEC Privilege mode, Figure 4-4

Terminal Lines

You can access the system remotely and restrict access to the system by creating user profiles. The terminal lines on the system provide different means of accessing the system. The console line (console) connects you through the Console port in the RPMs. The virtual terminal lines (VTY) connect you through Telnet to the system. The auxiliary line (aux) connects secondary devices such as modems.

Deny and Permit Access to a Terminal Line

Force 10 recommends applying only standard ACLs to deny and permit access to VTY lines.

- Layer 3 ACL deny all traffic that is not explicitly permitted, but in the case of VTY lines, an ACL with no rules does not deny any traffic.
- You cannot use **show ip accounting access-list** to display the contents of an ACL that is applied only to a VTY line.

To apply an IP ACL to a line:

Task	Command Syntax	Command Mode
Apply an ACL to a VTY line.	**ip access-class** *access-list*	LINE

To view the configuration, enter the **show config** command in the LINE mode, as shown in Figure 4-5.

Figure 4-5. Applying an Access List to a VTY Line

```
FTOS(config-std-nacl)#show config
!
ip access-list standard myvtyacl
 seq 5 permit host 10.11.0.1
FTOS(config-std-nacl)#line vty 0
FTOS(config-line-vty)#show config
line vty 0
 access-class myvtyacl
```

FTOS Behavior: Prior to FTOS version 7.4.2.0, in order to deny access on a VTY line, you must apply an ACL and AAA authentication to the line. Then users are denied access only *after* they enter a username and password. Beginning in FTOS version 7.4.2.0, only an ACL is required, and users are denied access *before* they are prompted for a username and password.

Configure Login Authentication for Terminal Lines

You can use any combination of up to 6 authentication methods to authenticate a user on a terminal line. A combination of authentication methods is called a method list. If the user fails the first authentication method, FTOS prompts the next method until all methods are exhausted, at which point the connection is terminated. The available authentication methods are:

- **enable**—Prompt for the enable password.
- **line**—Prompt for the e password you assigned to the terminal line. You must configure a password for the terminal line to which you assign a method list that contains the **line** authentication method. Configure a password using the command password from LINE mode.
- **local**—Prompt for the the system username and password.
- **none**—Do not authenticate the user.
- **radius**—Prompt for a username and password and use a RADIUS server to authenticate.
- **tacacs+**—Prompt for a username and password and use a TACACS+ server to authenticate.

To configure authentication for a terminal line:

Step	Task	Command Syntax	Command Mode
1	Create an authentication method list. You may use a mnemonic name or use the keyword **default**. The default authentication method for terminal lines is **local**, and the **default** method list is empty.	**aaa authentication login** {*method-list-name* | **default**} [*method-1*] [*method-2*] [*method-3*] [*method-4*] [*method-5*] [*method-6*]	CONFIGURATION
2	Apply the method list from Step 1 to a terminal line.	**login authentication** {*method-list-name* | **default**}	CONFIGURATION

Step	Task	Command Syntax	Command Mode
3	If you used the line authentication method in the method list you applied to the terminal line, configure a password for the terminal line.	**password**	LINE

In Figure 4-6 VTY lines 0-2 use a single authentication method, **line**.

Figure 4-6. Configuring Login Authentication on a Terminal Line

```
FTOS(conf)#aaa authentication login myvtymethodlist line
FTOS(conf)#line vty 0 2
FTOS(config-line-vty)#login authentication myvtymethodlist
FTOS(config-line-vty)#password myvtypassword
FTOS(config-line-vty)#show config
line vty 0
 password myvtypassword
login authentication myvtymethodlist
line vty 1
 password myvtypassword
login authentication myvtymethodlist
line vty 2
 password myvtypassword
login authentication myvtymethodlist
FTOS(config-line-vty)#
```

Time out of EXEC Privilege Mode

EXEC timeout is a basic security feature that returns FTOS to the EXEC mode after a period of inactivity on terminal lines.

To change the timeout period or disable EXEC timeout.

Task	Command Syntax	Command Mode
Set the number of minutes and seconds. Default: 10 minutes on console, 30 minutes on VTY. Disable EXEC timeout by setting the timeout period to 0.	**exec-timeout** *minutes* [*seconds*]	LINE
Return to the default timeout values.	**no exec-timeout**	LINE

View the configuration using the command **show config** from LINE mode.

Figure 4-7. Configuring EXEC Timeout

```
FTOS(conf)#line con 0
FTOS(config-line-console)#exec-timeout 0
FTOS(config-line-console)#show config
line console 0
 exec-timeout 0 0
FTOS(config-line-console)#
```

Telnet to Another Network Device

To telnet to another device:

Task	Command Syntax	Command Mode
Telnet to the peer RPM. You do not need to configure the management port on the peer RPM to be able to telnet to it.	**telnet-peer-rpm**	EXEC Privilege
Telnet to a device with an IPv4 or IPv6 address. If you do not enter an IP address, FTOS enters a Telnet dialog that prompts you for one. • Enter an IPv4 address in dotted decimal format (A.B.C.D). • Enter an IPv6 address in the format 0000:0000:0000:0000:0000:0000:0000:0000. Elision of zeros is supported. **Note:** Telnet to link-local addresses is not supported.	**telnet** [*ipv4-address* | *ipv6-address*]	EXEC Privilege

Figure 4-8. Telnet to Another Network Device

```
FTOS# telnet 10.11.80.203
Trying 10.11.80.203...
Connected to 10.11.80.203.
Exit character is '^]'.
Login:
Login: admin
Password:
FTOS>exit
FTOS#telnet 2200:2200:2200:2200:2200::2201
Trying 2200:2200:2200:2200:2200::2201...
Connected to 2200:2200:2200:2200:2200::2201.
Exit character is '^]'.
FreeBSD/i386 (freebsd2.force10networks.com) (ttyp1)
login: admin
FTOS#
```

Lock CONFIGURATION mode

FTOS allows multiple users to make configurations at the same time. You can lock CONFIGURATION mode so that only one user can be in CONFIGURATION mode at any time (Message 2).

A two types of locks can be set: auto and manual.

- Set an auto-lock using the command **configuration mode exclusive auto** from CONFIGURATION mode. When you set an auto-lock, every time a user is in CONFIGURATION mode all other users are denied access. This means that you can exit to EXEC Privilege mode, and re-enter CONFIGURATION mode without having to set the lock again.
- Set a manual lock using the command **configure terminal lock** from CONFIGURATION mode. When you configure a manual lock, which is the default, you must enter this command time you want to enter CONFIGURATION mode and deny access to others.

Figure 4-9. Locking CONFIGURATION mode

```
R1(conf)#configuration mode exclusive auto
BATMAN(conf)#exit
3d23h35m: %RPM0-P:CP %SYS-5-CONFIG_I: Configured from console by  console

R1#config
! Locks configuration mode exclusively.
R1(conf)#
```

If another user attempts to enter CONFIGURATION mode while a lock is in place, Message 1 appears on their terminal.

Message 1 CONFIGURATION mode Locked Error

```
% Error: User "" on line console0 is in exclusive configuration mode
```

If *any* user is already in CONFIGURATION mode when while a lock is in place, Message 2 appears on their terminal.

Message 2 Cannot Lock CONFIGURATION mode Error

```
% Error: Can't lock configuration mode exclusively since the following users are currently
configuring the system:
User "admin" on line vty1 ( 10.1.1.1 )
```

Note: The CONFIGURATION mode lock corresponds to a VTY session, not a user. Therefore, if you configure a lock and then exit CONFIGURATION mode, and another user enters CONFIGURATION mode, when you attempt to re-enter CONFIGURATION mode, you are denied access even though *you* are the one that configured the lock.

Note: If your session times out and you return to EXEC mode, the CONFIGURATION mode lock is unconfigured.

Viewing the Configuration Lock Status

If you attempt to enter CONFIGURATION mode when another user has locked it, you may view which user has control of CONFIGURATION mode using the command **show configuration lock** from EXEC Privilege mode.

You can then send any user a message using the **send command** from EXEC Privilege mode. Alternatively you can clear any line using the command **clear** from EXEC Privilege mode. If you clear a console session, the user is returned to EXEC mode.

Recovering from a Forgotten Password

If you configure authentication for the console and you exit out of EXEC mode or your console session times out, you are prompted for a password to re-enter.

If you forget your password:

Step	Task	Command Syntax	Command Mode
1	Log onto the system via console.		
2	Power-cycle the chassis by switching off all of the power modules and then switching them back on.		
3	Abort bootup by sending the break signal when prompted.	**Ctrl+Shift+6**	
4	Enter BOOT_ADMIN mode using the command **enable admin**. Enter **ncorerulz** when prompted for a password.	**enable admin**	BOOT_USER
5	Rename the startup-config so it does not load on the next system reload.	**rename :flash://startup-config flash://startup-config.bak**	BOOT_ADMIN
6	Verify that startup-config is renamed.	**dir flash:**	BOOT_ADMIN

Figure 4-10. Entering BOOT_USER mode

```
Type "go 0x00040004" to enter the FTOS BLI shell
You can use U-boot native networking facilities
============================================================


Hit any key to stop autoboot:  0
Starting F10 BLI Shell ...

BOOT_USER # enable admin
Password : XXXXXXXXX
RPM0-CP BOOT_ADMIN #
```

Figure 4-11. Entering BOOT_ADMIN mode

```
***** Welcome to FTOS Boot Interface *****
Use "help" or "?" for more information.
BOOT_USER # enable admin
Password : XXXXXXXXX
RPM0-CP BOOT_ADMIN #
```

Step	Task	Command Syntax	Command Mode

Figure 4-12. Renaming the startup-config

```
RPM0-CP BOOT_ADMIN # dir flash:
Directory of flash:
1 -rwx 11407411 Jun 09 2004 09:38:40 FTOS-EE3-5.3.1.1.bin
2 -rwx 4977 Jun 09 2004 09:38:38 startup-config.bak
```

7	Reload the system.	**reload**	BOOT_ADMIN
8	Copy startup-config.bak to the running config.	**copy flash://startup-config.bak running-config**	EXEC Privilege
9	Remove all authentication statements you might have for the console.	**no authentication login** **no password**	LINE
10	Save the running-config.	**copy running-config startup-config**	EXEC Privilege

Recovering from a Forgotten Enable Password

If you forget the enable password:

Step	Task	Command Syntax	Command Mode
1	Log onto the system via console.		
2	Eject the secondary RPM if there is one.		
3	Power-cycle the chassis by switching off all of the power modules and then switching them back on.		
4	Abort bootup by sending the break signal when prompted. See Figure 4-10.	**Ctrl+Shift+6**	
5	Configure the system to ignore the enable password on bootup. **Note:** This command only bypasses the enable password once. You must repeat this procedure to bypass it again.	**ignore enable-password**	BOOT_USER

Figure 4-13. Ignoring the Enable Password

```
***** Welcome to FTOS Boot Interface *****
Use "help" or "?" for more information.
BOOT_USER # ignore enable-password
```

6	Reload the system.	**reload**	BOOT_USER
7	Configure a new enable password.	**enable** {**secret** | **password**}	CONFIGURATION
8	Insert the secondary RPM.		

Step	Task	Command Syntax	Command Mode
9	Save the running-config to the startup-config. The startup-config files on both RPMs will be synchronized.	**copy running-config startup-config**	EXEC Privilege

Recovering from a Forgotten Password on S-Series

If you configure authentication for the console and you exit out of EXEC mode or your console session times out, you are prompted for a password to re-enter.

If you forget your password:

Step	Task	Command Syntax	Command Mode
1	Log onto the system via console.		
2	Power-cycle the chassis by unplugging the power cord.		
3	Abort bootup by sending the break signal when prompted.	(any key)	

Figure 4-14. Entering BOOT_USER mode

```
Type "go 0x00040004" to enter the FTOS BLI shell
You can use U-boot native networking facilities
============================================================


***** Welcome to FTOS Boot Interface *****
Use "help" or "?" for more information.
BOOT_USER #
```

4	Configure the system to ignore the startup-config, which prevents the system from prompting you for a password to enter EXEC mode. **Note:** This command only bypasses the password once. You must repeat this procedure to bypass it again.	**ignore startup-config**	BOOT_USER
5	Remove all authentication statements you might have for the console.	**no authentication login**	CONFIGURATION
6	Reload the system.	**reload**	BOOT_USER

Recovering from a Failed Start

A system that does not start correctly might be attempting to boot from a corrupted FTOS image or from a incorrect location. To resolve the problem, you can restart the system and interrupt the boot process to point the system to another boot location by using the **boot change** command, as described below. For details on the **boot change** command, its supporting commands, and other commands that can help recover from a failed start, refer to the BOOT_USER chapter in the *FTOS Command Reference*.

Step	Task	Command Syntax	Command Mode
1	Power-cycle the chassis (pull the power cord and reinsert it).		
2	Abort bootup by sending the break signal when prompted.	**Ctrl-Shift 6** (**Ctrl-**^)—C-Series and E-Series (On the S-Series, hit any key)	(during bootup)
3	Tell the system where to access the FTOS image used to boot the system: • Enter **primary** to configure the boot parameters used in the first attempt to boot the system. • Enter **secondary** for when the primary operating system boot selection is not available. • Enter **default** to configure boot parameters used if the secondary operating system boot parameter selection is not available. The default location should always be the internal flash device (flash:), and a verified image should be stored there.	**boot change** {**primary** | **secondary** | **default**} After entering the keywords and desired option, press **Enter**. The software prompts you to enter the following: • boot device (ftp, tftp, flash, slot0) **Note**: S-Series can only use a TFTP location. • image file name • IP address of the server with the image • username and password (only for FTP)	BOOT_USER
4	On S-Series systems only, assign a port to be the Management Ethernet interface.	**interface management ethernet port** *portID*	BOOT_USER
5	Assign an IP address to the Management Ethernet interface.	[**no**] **interface management ethernet ip address** *ip-address mask*	BOOT_USER
6	(OPTIONAL) On C- and E-Series systems only, configure speed, duplex, and negotiation settings for the management interface.	**interface management port config** {**half-duplex** | **full-duplex** | **10m** | **100m** | **auto-negotiation** | **no auto-negotiation** | **show**}	BOOT_USER
7	Assign an IP address as the default gateway for the system.	[**no**] **default-gateway** *ip-address*	BOOT_USER
8	Reload the system.	**reload**	BOOT_USER

Very similar to the options of the **boot change** command, the **boot system** command is available in CONFIGURATION mode on the C-Series and E-Series to set the boot parameters that, when saved to the startup configuration file, are stored in NVRAM and are then used routinely:

Task	Command Syntax	Command Mode
Configure the system to routinely boot from the designated location. After entering **rpm0** or **rpm1**, enter one of the three keywords and then the *file-url.* You can use the command for each of the combinations of RPM and option.	**boot system** {**rpm0** | **rpm1**} (**default** | **primary** | **secondary**} *file-url* For *file-url, to* boot from a file: • on the internal Flash, enter **flash://** followed by the filename. • on an FTP server, enter **ftp://** *user:password@hostip/filepath* • on the external Flash, enter **slot0://** followed by the filename. • on a TFTP server, enter **tftp://***hostip/ filepath*	CONFIGURATION

Also, because the C-Series and E-Series can boot from an external flash, you can recover from a failed boot image on the flash by simply fixing that source. For details on boot code and FTOS setup, see the *FTOS Release Notes* for the specific FTOS versions that you want to use.

The network boot facility has only become available on the S-Series with FTOS 7.8.1.0 and its accompanying boot code. In addition to installing FTOS 7.8.1.0, you must separately install that new boot code. For installation details, see the *S-Series and FTOS Release Notes* for Version 7.8.1.0.

5

802.1ag

802.1ag is available only on platform: S

Ethernet Operations, Administration, and Maintenance (OAM) is a set of tools used to install, monitor, troubleshoot and manage Ethernet infrastructure deployments. Ethernet OAM consists of three main areas:

1. Service Layer OAM: IEEE 802.1ag Connectivity Fault Management (CFM)
2. Link Layer OAM: IEEE 802.3ah OAM
3. Ethernet Local management Interface (MEF-16 E-LMI)

Ethernet CFM

Ethernet CFM is an end-to-end, per-service-instance Ethernet OAM scheme which enables: proactive connectivity monitoring, fault verification, and fault isolation.

The service-instance in the OAM for Metro/Carrier Ethernet context is a VLAN. This service is sold to an end-customer by a network service provider. Typically the service provider contracts with multiple network operators to provide end-to-end service between customers. For end-to-end service between customer switches, connectivity must be present across the service provider through multiple network operators.

Layer 2 Ethernet networks usually cannot be managed with IP tools such as ICMP Ping and IP Traceroute. Traditional IP tools often fail because:

- there are complex interactions between various Layer 2 and Layer 3 protocols such as STP, LAG, VRRP and ECMP configurations.
- Ping and traceroute are not designed to verify data connectivity in the network and within each node in the network (such as in the switching fabric and hardware forwarding tables).
- when networks are built from different operational domains, access controls impose restrictions that cannot be overcome at the IP level, resulting in poor fault visibility. There is a need for hierarchical domains that can be monitored and maintained independently by each provider or operator.
- routing protocols choose a subset of the total network topology for forwarding, making it hard to detect faults in links and nodes that are not included in the active routing topology. This is made more complex when using some form of Traffic Engineering (TE) based routing.
- network and element discovery and cataloging is not clearly defined using IP troubleshooting tools.

There is a need for Layer 2 equivalents to manage and troubleshoot native Layer 2 Ethernet networks. With these tools, you can identify, isolate, and repair faults quickly and easily, which reduces operational cost of running the network. OAM also increases availability and reduces mean time to recovery, which allows for tighter service level agreements, resulting in increased revenue for the service provider.

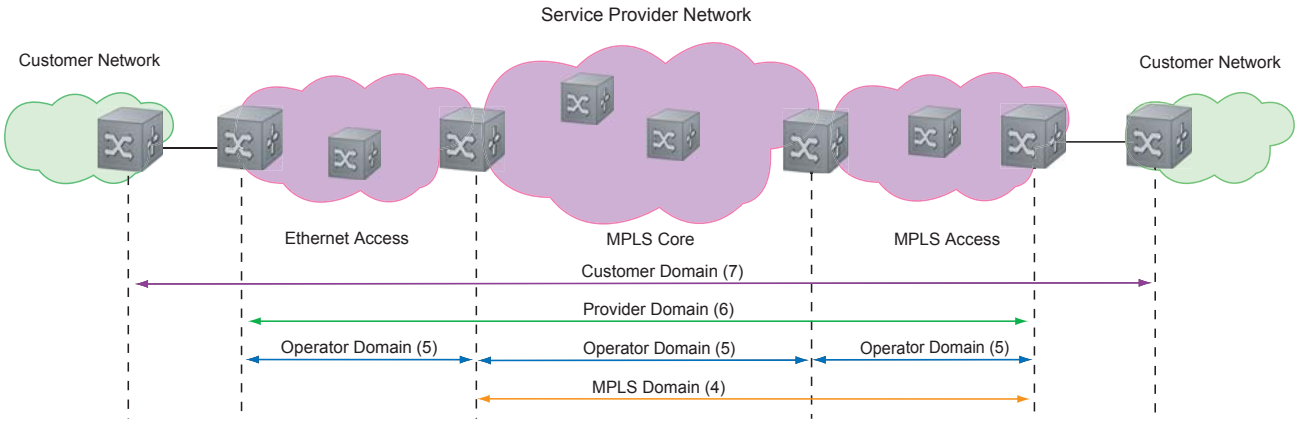
In addition to providing end-to-end OAM in native Layer 2 Ethernet Service Provider/Metro networks, you can also use CFM to manage and troubleshoot any Layer 2 network including enterprise, datacenter, and cluster networks.

Maintenance Domains

Connectivity Fault Management (CFM) divides a network into hierarchical maintenance domains, as shown in Figure 5-1.

A CFM maintenance domain is a management space on a network that is owned and operated by a single management entity. The network administrator assigns a unique maintenance level (0 to 7) to each domain to define the hierarchical relationship between domains. Domains can touch or nest but cannot overlap or intersect as that would require management by multiple entities.

Figure 5-1. OAM Domains



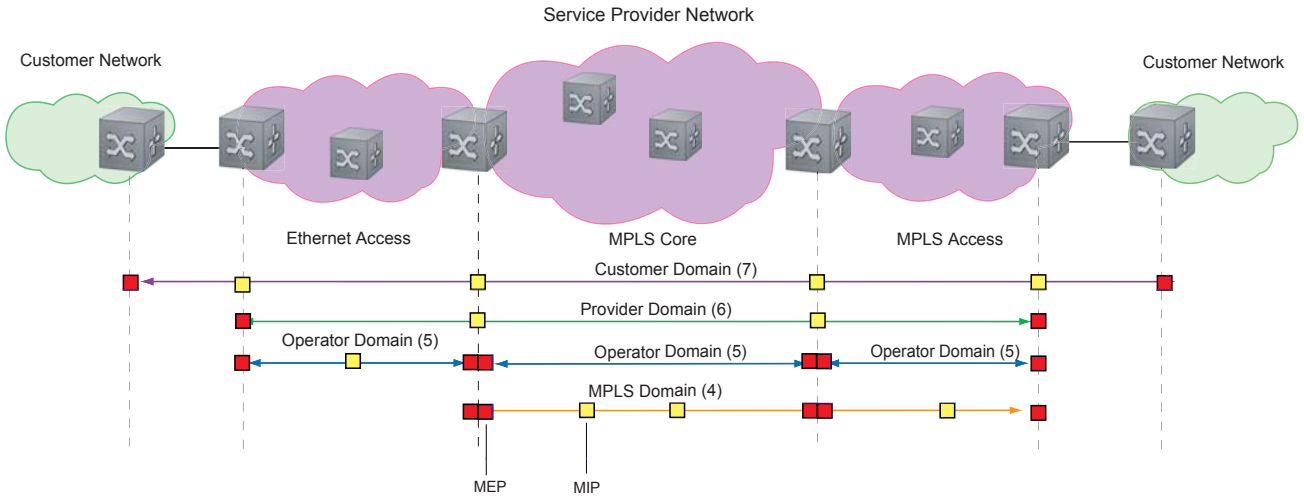
Maintenance Points

Domains are comprised of logical entities called Maintenance Points. A maintenance point is an interface demarcation that confines CFM frames to a domain. There are two types of maintenance points:

- **Maintenance End Points (MEPs)**: a logical entity that marks the end-point of a domain
- **Maintenance Intermediate Points (MIPs)**: a logical entity configured at a port of a switch that is an intermediate point of a Maintenance Entity (ME). An ME is a point-to-point relationship between two MEPs within a single domain. MIPs are internal to a domain, not at the boundary, and respond to CFM only when triggered by linktrace and loopback messages. MIPs can be configured to snoop Continuity Check Messages (CCMs) to build a MIP CCM database.

These roles define the relationships between all devices so that each device can monitor the layers under its responsibility. Maintenance points drop all lower-level frames and forward all higher-level frames.

Figure 5-2. Maintenance Points



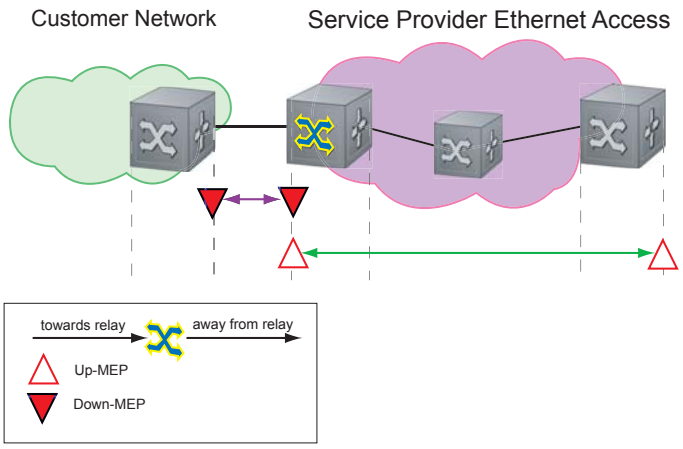
Maintenance End Points

A Maintenance End Point (MEP) is a logical entity that marks the end-point of a domain. There are two types of MEPs defined in 802.1ag for an 802.1 bridge:

- **Up-MEP**: monitors the forwarding path internal to an bridge on the customer or provider edge; on Dell Force10 systems the internal forwarding path is effectively the switch fabric and forwarding engine.
- **Down-MEP**: monitors the forwarding path external another bridge.

Configure Up- MEPs on ingress ports, ports that send traffic towards the bridge relay. Configure Down-MEPs on egress ports, ports that send traffic away from the bridge relay.

Figure 5-3. Up-MEP versus Down-MEP



Implementation Information

- Since the S-Series has a single MAC address for all physical/LAG interfaces, only one MEP is allowed per MA (per VLAN or per MD level).

Configure CFM

Configuring CFM is a five-step process:

1. Configure the ecfmacl CAM region using the **cam-acl** command. See Configure Ingress Layer 2 ACL Sub-partitions.
2. Enable Ethernet CFM. See page 83.
3. Create a Maintenance Domain. See page 83.
4. Create a Maintenance Association. See page 84.
5. Create Maintenance Points. See page 84.
6. Use CFM tools:
   a Continuity Check Messages on page 87
   b Loopback Message and Response on page 88
   c Linktrace Message and Response on page 88

Related Configuration Tasks

- Enable CFM SNMP Traps. on page 90
- Display Ethernet CFM Statistics on page 91

Enable Ethernet CFM

Task	Command Syntax	Command Mode
Spawn the CFM process. No CFM configuration is allowed until the CFM process is spawned.	**ethernet cfm**	CONFIGURATION
Disable Ethernet CFM without stopping the CFM process.	**disable**	ETHERNET CFM

Create a Maintenance Domain

Connectivity Fault Management (CFM) divides a network into hierarchical maintenance domains, as shown in Figure 5-1.

Step	Task	Command Syntax	Command Mode
1	Create maintenance domain.	**domain** *name* **md-level** *number* Range: 0-7	ETHERNET CFM
2	Display maintenance domain information.	**show ethernet cfm domain** [*name* | **brief**]	EXEC Privilege

```
FTOS# show ethernet cfm domain

Domain Name: customer
Level: 7
Total Service: 1
    Services
              MA-Name          VLAN        CC-Int        X-CHK Status

              My_MA            200           10s           enabled

Domain Name: praveen
Level: 6
Total Service: 1
    Services
              MA-Name          VLAN        CC-Int        X-CHK Status

              Your_MA          100           10s           enabled
```

Create a Maintenance Association

A Maintenance Association MA is a subdivision of an MD that contains all managed entities corresponding to a single end-to-end service, typically a VLAN. An MA is associated with a VLAN ID.

Task	Command Syntax	Command Mode
Create maintenance association.	**service** *name* **vlan** *vlan-id*	ECFM DOMAIN

Create Maintenance Points

Domains are comprised of logical entities called Maintenance Points. A maintenance point is a interface demarcation that confines CFM frames to a domain. There are two types of maintenance points:

- **Maintenance End Points (MEPs)**: a logical entity that marks the end-point of a domain
- **Maintenance Intermediate Points (MIPs)**: a logical entity configured at a port of a switch that constitutes intermediate points of an Maintenance Entity (ME). An ME is a point-to-point relationship between two MEPs within a single domain.

These roles define the relationships between all devices so that each device can monitor the layers under its responsibility.

Create a Maintenance End Point

A Maintenance End Point (MEP) is a logical entity that marks the end-point of a domain. There are two types of MEPs defined in 802.1ag for an 802.1 bridge:

- **Up-MEP**: monitors the forwarding path internal to an bridge on the customer or provider edge; on Dell Force10 systems the internal forwarding path is effectively the switch fabric and forwarding engine.
- **Down-MEP**: monitors the forwarding path external another bridge.

Configure Up- MEPs on ingress ports, ports that send traffic towards the bridge relay. Configure Down-MEPs on egress ports, ports that send traffic away from the bridge relay.

Task	Command Syntax	Command Mode
Create an MEP.	**ethernet cfm mep** {**up-mep | down-mep**} **domain** {*name* | *level* } **ma-name** *name* **mepid** *mep-id* Range: 1-8191	INTERFACE
Display configured MEPs and MIPs.	**show ethernet cfm maintenance-points local** [**mep** | **mip**]	EXEC Privilege

Task	Command Syntax	Command Mode

```
FTOS#show ethernet cfm maintenance-points local mep
 -------------------------------------------------------------------------------
MPID          Domain Name      Level   Type          Port           CCM-Status
                MA Name        VLAN    Dir            MAC
-------------------------------------------------------------------------------

 100                   cfm0     7      MEP         Gi 4/10            Enabled
                      test0     10     DOWN     00:01:e8:59:23:45

 200                   cfm1     6      MEP         Gi 4/10            Enabled
                      test1     20     DOWN     00:01:e8:59:23:45

 300                   cfm2     5      MEP         Gi 4/10            Enabled
                      test2     30     DOWN     00:01:e8:59:23:45
```

Create a Maintenance Intermediate Point

Maintenance Intermediate Point (MIP) is a logical entity configured at a port of a switch that constitutes intermediate points of an Maintenance Entity (ME). An ME is a point-to-point relationship between two MEPs within a single domain. An MIP is not associated with any MA or service instance, and it belongs to the entire MD.

Task	Command Syntax	Command Mode
Create an MIP.	**ethernet cfm mip domain** {*name* | *level*} **ma-name** *name*	INTERFACE
Display configured MEPs and MIPs.	**show ethernet cfm maintenance-points local** [**mep** | **mip**]	EXEC Privilege

```
FTOS#show ethernet cfm maintenance-points local mip
 -------------------------------------------------------------------------------
MPID          Domain Name      Level   Type          Port           CCM-Status
                MA Name        VLAN    Dir            MAC
-------------------------------------------------------------------------------

   0                service1     4      MIP          Gi 0/5           Disabled
                    My_MA       3333    DOWN     00:01:e8:0b:c6:36

   0                service1     4      MIP          Gi 0/5           Disabled
                    Your_MA     3333      UP     00:01:e8:0b:c6:36
```

MP Databases

CFM maintains two MP databases:

- **MEP Database (MEP-DB)**: Every MEP must maintain a database of all other MEPs in the MA that have announced their presence via CCM.

- **MIP Database (MIP-DB)**: Every MIP must maintain a database of all other MEPs in the MA that have announced their presence via CCM

Task	Command Syntax	Command Mode
Display the MEP Database.	**show ethernet cfm maintenance-points remote detail** [**active** | **domain** {*level* | *name*} | **expired** | **waiting**]	EXEC Privilege

```
FTOS#show ethernet cfm maintenance-points remote detail

MAC Address: 00:01:e8:58:68:78
Domain Name: cfm0
MA Name: test0
Level: 7
VLAN: 10
MP ID: 900
Sender Chassis ID: FTOS
MEP Interface status: Up
MEP Port status: Forwarding
Receive RDI: FALSE
MP Status: Active
```

Display the MIP Database.	**show ethernet cfm mipdb**	EXEC Privilege

MP Database Persistence

Task	Command Syntax	Command Mode
Set the amount of time that data from a missing MEP is kept in the Continuity Check Database.	**database hold-time** *minutes* Default: 100 minutes Range: 100-65535 minutes	ECFM DOMAIN

Continuity Check Messages

Continuity Check Messages (CCM) are periodic hellos used to:

- discover MEPs and MIPs within a maintenance domain
- detect loss of connectivity between MEPs
- detect misconfiguration, such as VLAN ID mismatch between MEPs
- to detect unauthorized MEPs in a maintenance domain

Continuity Check Messages (CCM) are multicast Ethernet frames sent at regular intervals from each MEP. They have a destination address based on the MD level (01:80:C2:00:00:3X where X is the MD level of the transmitting MEP from 0 to 7). All MEPs must listen to these multicast MAC addresses and process these messages. MIPs may optionally processes the CCM messages originated by MEPs and construct a MIP CCM database.

MEPs and MIPs filter CCMs from higher and lower domain levels as described in Table 5-1.

Table 5-1. Continuity Check Message Processing

Frames at	Frames from	UP-MEP Action	Down-MEP Action	MIP Action
Less than my level	Bridge-relay side or Wire side	Drop	Drop	Drop
My level	Bridge-relay side	Consume	Drop	Add to MIP-DB and forward
	Wire side	Drop	Consume	
Greater than my level	Bridge-relay side or Wire side	Forward	Forward	Forward

All the remote MEPs in the maintenance domain are defined on each MEP. Each MEP then expects a periodic CCM from the configured list of MEPs. A connectivity failure is then defined as:

1. Loss of 3 consecutive CCMs from any of the remote MEP, which indicates a network failure
2. Reception of a CCM with an incorrect CCM transmission interval, which indicates a configuration error.
3. Reception of CCM with an incorrect MEP ID or MAID, which indicates a configuration or cross-connect error. This could happen when different VLANs are cross-connected due to a configuration error.
4. Reception of a CCM with an MD level lower than that of the receiving MEP, which indicates a configuration or cross-connect error.
5. Reception of a CCM containing a port status/interface status TLV, which indicates a failed bridge or aggregated port.

The Continuity Check protocol sends fault notifications (Syslogs, and SNMP traps if enabled) whenever any of the above errors are encountered.

Enable CCM

Step	Task	Command Syntax	Command Mode
1	Enable CCM.	**no ccm disable** Default: Disabled	ECFM DOMAIN
2	Configure the transmit interval (mandatory). The interval specified applies to all MEPs in the domain.	**ccm transmit-interval** *seconds* Default: 10 seconds	ECFM DOMAIN

Enable Cross-checking

Task	Command Syntax	Command Mode
Enable cross-checking.	**mep cross-check enable** Default: Disabled	ETHERNET CFM
Start the cross-check operation for an MEP.	**mep cross-check** *mep-id*	ETHERNET CFM
Configure the amount of time the system waits for a remote MEP to come up before the cross-check operation is started.	**mep cross-check start-delay** *number*	ETHERNET CFM

Loopback Message and Response

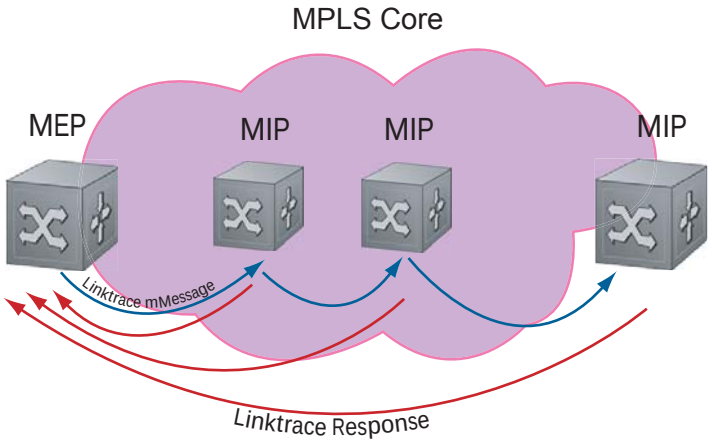
Loopback Message and Response (LBM, LBR), also called Layer 2 Ping, is an administrative echo transmitted by MEPs to verify reachability to another MEP or MIP within the maintenance domain. LBM and LBR are unicast frames.

Task	Command Syntax	Command Mode
Send a Loopback message.	**ping ethernet domain** *name* **ma-name** *ma-name* **remote** {*mep-id* | **mac-addr** *mac-address*} **source** {*mep-id* | **port** *interface*}	EXEC Privilege

Linktrace Message and Response

Linktrace Message and Response (LTM, LTR), also called Layer 2 Traceroute, is an administratively sent multicast frames transmitted by MEPs to track, hop-by-hop, the path to another MEP or MIP within the maintenance domain. All MEPs and MIPs in the same domain respond to an LTM with a unicast LTR. Intermediate MIPs forward the LTM toward the target MEP.

Figure 5-4. Linktrace Message and Response



Link trace messages carry a unicast target address (the MAC address of an MIP or MEP) inside a multicast frame. The destination group address is based on the MD level of the transmitting MEP (01:80:C2:00:00:3[8 to F]). The MPs on the path to the target MAC address reply to the LTM with an LTR, and relays the LTM towards the target MAC until the target MAC is reached or TTL equals 0.

Task	Command Syntax	Command Mode
Send a Linktrace message. Since the LTM is a Multicast message sent to the entire ME, there is no need to specify a destination.	**traceroute ethernet domain**	EXEC Privilege

Link Trace Cache

After a Link Trace command is executed, the trace information can be cached so that you can view it later without retracing.

Task	Command Syntax	Command Mode
Enable Link Trace caching.	**traceroute cache**	CONFIGURATION
Set the amount of time a trace result is cached.	**traceroute cache hold-time** *minutes* Default: 100 minutes Range: 10-65535 minutes	ETHERNET CFM
Set the size of the Link Trace Cache.	**traceroute cache size** *entries* Default: 100 Range: 1 - 4095 entries	ETHERNET CFM
Display the Link Trace Cache.	**show ethernet cfm traceroute-cache**	EXEC Privilege

Task	Command Syntax	Command Mode

```
FTOS#show ethernet cfm traceroute-cache

Traceroute to 00:01:e8:52:4a:f8 on Domain Customer2, Level 7, MA name Test2 with
VLAN 2

 ------------------------------------------------------------------------------
  Hops              Host              IngressMAC     Ingr Action   Relay Action
                 Next Host            Egress MAC     Egress Action  FWD Status
------------------------------------------------------------------------------

   4      00:00:00:01:e8:53:4a:f8   00:01:e8:52:4a:f8  IngOK            RlyHit
         00:00:00:01:e8:52:4a:f8                                    Terminal MEP
```

Delete all Link Trace Cache entries.	**clear ethernet cfm traceroute-cache**	EXEC Privilege

Enable CFM SNMP Traps.

Task	Command Syntax	Command Mode
Enable SNMP trap messages for Ethernet CFM.	**snmp-server enable traps ecfm**	CONFIGURATION

A Trap is sent only when one of the five highest priority defects occur, as shown in Table 5-2.

Table 5-2. ECFM SNMP Traps

Cross-connect defect	`%ECFM-5-ECFM_XCON_ALARM: Cross connect fault detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000`
Error-CCM defect	`%ECFM-5-ECFM_ERROR_ALARM: Error CCM Defect detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000`
MAC Status defect	`%ECFM-5-ECFM_MAC_STATUS_ALARM: MAC Status Defect detected by MEP 1 in Domain provider at Level 4 VLAN 3000`
Remote CCM defect	`%ECFM-5-ECFM_REMOTE_ALARM: Remote CCM Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000`
RDI defect	`%ECFM-5-ECFM_RDI_ALARM: RDI Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000`

Three values are given within the trap messages: MD Index, MA Index, and MPID. You can reference these values against the output of **show ethernet cfm domain** and **show ethernet cfm maintenance-points local mep**.

```
FTOS#show ethernet cfm maintenance-points local mep
 -------------------------------------------------------------------------------
MPID          Domain Name      Level   Type          Port           CCM-Status
                MA Name        VLAN    Dir            MAC
-------------------------------------------------------------------------------

 100                    cfm0     7      MEP         Gi 4/10            Enabled
                       test0     10    DOWN     00:01:e8:59:23:45
```

```
FTOS(conf-if-gi-0/6)#do show ethernet cfm domain

Domain Name: My_Name
MD Index: 1
Level: 0
Total Service: 1
    Services
MA-Index        MA-Name           VLAN         CC-Int        X-CHK Status

      1         test               0             1s            enabled

Domain Name: Your_Name
MD Index: 2
Level: 2
Total Service: 1
    Services
MA-Index        MA-Name           VLAN         CC-Int        X-CHK Status

      1         test             100            1s            enabled
```

Display Ethernet CFM Statistics

Task	Command Syntax	Command Mode
Display MEP CCM statistics.	**show ethernet cfm statistics** [**domain** {*name* | *level*} **vlan-id** *vlan-id* **mpid** *mpid*	EXEC Privilege

```
FTOS#  show ethernet cfm statistics

Domain Name: Customer
Domain Level: 7
MA Name: My_MA
MPID: 300

     CCMs:
       Transmitted:                1503     RcvdSeqErrors:              0
     LTRs:
       Unexpected Rcvd:               0
     LBRs:
       Received:                      0     Rcvd Out Of Order:          0
       Received Bad MSDU:             0
       Transmitted:                   0
```

Task	Command Syntax	Command Mode
Display CFM statistics by port.	**show ethernet cfm port-statistics** [**interface**]	EXEC Privilege

```
FTOS#show ethernet cfm port-statistics interface gigabitethernet 0/5
Port statistics for port: Gi 0/5
================================

RX Statistics
=============
Total CFM Pkts 75394 CCM Pkts 75394
LBM Pkts 0 LTM Pkts 0
LBR Pkts 0 LTR Pkts 0
Bad CFM Pkts 0 CFM Pkts Discarded 0
CFM Pkts forwarded 102417

TX Statistics
=============
Total CFM Pkts 10303 CCM Pkts 0
LBM Pkts 0 LTM Pkts 3
LBR Pkts 0 LTR Pkts 0
```

6

802.3ah

802.3ah is available only on platform: S

A metropolitan area network (MAN) is a set of LANs, geographically separated but managed by a single entity. If the distance is large—across a city, for example—connectivity between LANs is managed by a service provider. While LANs use Ethernet, service providers networks use an array of protocols (PPP and ATM), and a variety access technologies. Implementing Ethernet from end to end, across the service provider network, simplifies design and management, increases scalability and bandwidth, and reduces costs.

Ethernet in a service provider environment introduces the concept of *Carrier-class Ethernet* and requires some basic management and diagnostic tools. Ethernet Operations, Administration, and Maintenance (OAM) is that toolset, which can be used to install, monitor, troubleshoot, and manage Ethernet infrastructure deployments. It consists of three main areas:

1. Service Layer OAM: IEEE 802.1ag, Connectivity Fault Management (CFM)
2. Link Layer OAM: IEEE 802.3ah, Ethernet in the First Mile (EFM) OAM
3. Ethernet Local management Interface (MEF-16 E-LMI)

Link Layer OAM Overview

Link Layer OAM introduces the toolset required to effectively monitor the link between the customer and service provider, which is called *the first mile*. Currently, service providers use a variety of access technologies including ISDN, DSL, and coax cable in the first mile. Implementing Ethernet here reduces the types of equipment in the subscriber access network, simplifying installation and management, and increasing bandwidth.

Link Layer OAM performs four primary operations for the purposes of link status, performance monitoring, and fault detection and isolation for Ethernet in the First Mile:

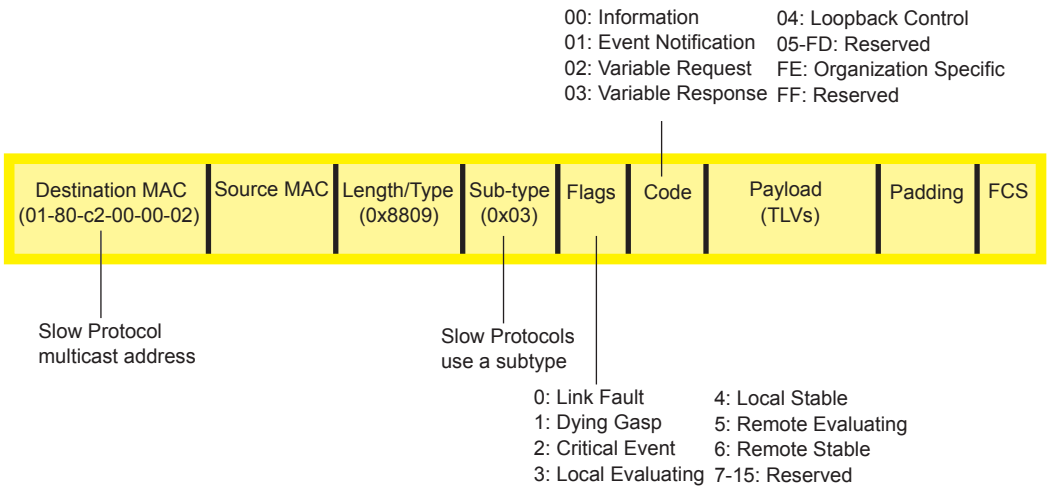
- **OAM Discovery—**detects whether the remote system is OAM capable, and negotiates OAM parameters.
- **Link Event Monitoring**—defines a set of events that may impact link operation, and monitors the link for those events.

- **Remote Loopback**—directs the remote system to reflects back frames that the local system transmits so that an administrator can isolate a fault.
- **Remote Failure Indication**—notifies a peer of a critical link event.

Link Layer OAMPDUs

Link Layer OAM is conducted using OAMPDUs, shown in Figure 6-1. OAM is a slow protocol and by requirement may transmit no more than 10 frames per second, transmits to a multicast destination MAC, and uses an Ethernet subtype.

Figure 6-1. OAMPFU Frame Format



There are six OAMPDU types, identified by the Code field:

- **Information**—carries state information and Local Information and/or Remote Information TLVs. Information OAMPDUs are used in discovery, and as keepalives.
  - **Local Information TLVs**—indicates support for variable retrieval, link performance events, and remote loopback, unidirectional support, and OAM mode
  - **Remote Information TLVs**—a copy of the peer's Local Information TLV.
- **Event Notification**—carries TLVs for each concurrent link fault.
- **Variable Request**—carries MIB object descriptors for which the remote peer should return values.
- **Variable Response**—carries the requested MIB object values.
- **Loopback Control**—carries the loopback control command (enable and disable).
- **Organization Specific**—contains and OUI followed by data, the format and function of which is defined by the organization.

OAMPDU Flags

1-bit flags are used it indicate OAM state and link state. During discovery, flags 3-6 are used to indicate the state of peership establishment. Flags 0-2 are used to indicate a local critical link event to the remote peer.

Link Layer OAM Operational Modes

When participating in EFM OAM, system may operate in active or passive mode.

- **Active mode**—Active mode systems initiate discovery. Once the Discovery process completes, they can send any OAMPDU while connected to a peer in Active mode, and a subset of OAMPDUs if the peer is in Passive mode (see Table 6-1).
- **Passive mode**—Passive mode systems wait for an active mode system to initiate discovery, and do not send Variable Request or Loopback Control OAMPDUs.

Taken from IEEE 802.3ah, Table 6-1 summarizes the permitted actions in each role.

Table 6-1. Active Mode and Passive Mode Behaviors

Capability	Active	Passive
Initiates OAM Discovery process	Yes	No
Reacts to OAM Discovery process initiation	Yes	Yes
Required to send Information OAMPDUs	Yes	Yes
Permitted to send Event Notification OAMPDUs	Yes	Yes
Permitted to send Variable Request OAMPDUs	Yes	No
Permitted to send Variable Response OAMPDUs (the peer must be in Active mode)	Yes	Yes
Permitted to send Loopback Control OAMPDUs	Yes	No
Reacts to Loopback Control OAMPDUs (the peer must be in Active mode)	Yes	Yes
Permitted to send Organization Specific OAMPDUs	Yes	Yes

Link Layer OAM Discovery

OAM Discovery is the mechanism a Link Layer OAM-capable system uses to determine if the remote system on the link has OAM functionality enabled. OAM Discovery ascertains OAM parameters, such as maximum allowable OAMPDU size, and supported functions such as OAM remote loopback.

The discovery process is as follows:

1. If the link is not in Fault state, Active mode systems send Information OAMPDUs that contain (only) the Local Information TLV.
2. Once a system receives an Information OAMPDU, it responds with an Information OAMPDU that contains the Local and Remote Information TLV. Negotiation is complete when both systems have received their peer's information and are satisfied with it; to be satisfied, both peers on the link must be have link performance event monitoring enabled.
3. When negotiation is complete, both peers may send any type of OAMPDU.

Link Layer OAM Events

Link Layer OAM defines a set of events that may impact link operation, and monitors the link for those events. If an event occurs, the detecting system notifies its peer. There are two types of events:

- **Critical Link Events**—There are three critical events; each has an associated flag which can be set in the OAMPDU when the event occurs. Critical link events are communicated to the peer using Remote Failure Indication.
  - **Link Fault**—A fault occurred in the receive direction of the local peer.
  - **Dying Gasp**—An unrecoverable local failure condition occurred. Dying Gasp notification is not supported on S-Series.
  - **Critical Event**—An unspecified critical event occurred. Critical Event notification is not supported on S-Series.
- **Link Performance Events**—Link events are either symbol errors or frame errors, and are communicated using Link Event TLVs.
  - **Symbol Errors**—a *symbol* is an (electrical or optical) pulse on the physical medium that represents one or more bits. A symbol error occurs when a symbol degrades in transit so that the receiver is not able to decode it. Gigabit and 10-Gigabit Ethernet have and expect symbol rate, also called *Baud*.
  - **Frame Error**s—frame errors are frames with a bad CRC.

Remote Loopback

An active-mode device can place a passive peer into loopback mode by sending a Loopback Control OAMPDU. When in loopback mode:

- the remote peer returns unaltered all non-OAMPDU frames sent by the local peer, and
- all outbound data frames are discarded (control frames are still forwarded).

Implementation Information

- Critical Link Events Dying Gasp and Critical Event are not supported.
- MIB retrieval is not supported.
- Both peers on a link must have Link Performance Monitoring Enabled, or else discovery does not complete.
- Control frames are still forwarded when an interface is in loopback mode.

Configure Link Layer OAM

Configuring Link Layer OAM is a two-step process:

1. Enable Link Layer OAM. See page 97.
2. Enable any or all of the following:
   a Link Performance Event Monitoring on page 99
   b Remote Failure Indication on page 102
   c Remote Loopback on page 103

Related Configuration Tasks

- Adjust the OAMPDU Transmission Parameters on page 99
- Display Link Layer OAM Configuration and Statistics on page 104
- Manage Link Layer OAM on page 106

Enable Link Layer OAM

Link Layer OAM is disabled by default. Enabling it places the system in Active mode and initiates OAM discovery.Both peers on the link must be have link performance event monitoring enabled for discovery to complete.

Task	Command Syntax	Command Mode
Enable Ethernet OAM.	**ethernet oam** Default: Disabled	INTERFACE
Display the OAM discovery status.	**show ethernet oam discovery interface** *interface*	EXEC Privilege

Task	Command Syntax	Command Mode

```
FTOS# show ethernet oam discovery interface <interface-name>
Output format:
<interface name>

Local client
____________
  Administrative configurations:
  Mode:active
  Unidirection:not supported
  Link monitor:supported (on)
  Remote loopback:not supported
      MIB retrieval:not supported
  Mtu size:1500
  Operational status:
Port status:operational
Loopback status:no loopback
PDU permission:any
PDU revision:1

Remote client
_____________
MAC address:0030.88fe.87de
Vendor(OUI):0x00 0x00 0x0C

Administrative configurations:
  Mode:active
  Unidirection:not supported
  Link monitor:supported
  Remote loopback:not supported
      MIB retrieval:not supported
  Mtu size:1500
```

Display Link Layer OAM sessions.	**show ethernet oam summary**	EXEC Privilege

```
FTOS# show ethernet oam summary

Output format :

Symbols:* - Master Loopback State, # - Slave Loopback State
Capability codes:L - Link Monitor, R - Remote Loopback
U - Unidirection,V - Variable Retrieval

  LocalRemote
InterfaceMAC AddressOUIModeCapability
 Gi6/1/10023.84ac.b8000000DactiveL R
```

Adjust the OAMPDU Transmission Parameters

Task	Command Syntax	Command Mode
Specify a the maximum or minimum number of OAMPDUs to be sent per second.	**ethernet oam** [**max-rate** *value* | **min-rate** *value*] Range: 1-10 Default: 10	INTERFACE
Set the transmission mode to active or passive.	**ethernet oam mode** {**active** | **passive**} Default: Active	INTERFACE
Specify the amount of time that the system waits to receive an OAMPDU from a peer before considering it non-operational.	**ethernet oam timeout** *value* Range: 2-30 seconds Default: 5 seconds	INTERFACE

Link Performance Event Monitoring

Link Performance Event Monitoring OAM monitors the receive side of a link for a set of pre-defined errors and executes an action when a threshold is exceeded; it is enabled by default. Both peers on the link must be have link performance event monitoring enabled for discovery to complete.

There is a high and low threshold for each pre-defined error; an *event* occurs when any threshold is exceeded. FTOS periodically polls hardware registers for the current frame and symbol error count. If an interface exceeds a threshold, a notification is sent to the peer and the interface is placed in error-disabled state.

- Enable Error Monitoring on page 99
- Execute an Action upon Exceeding the High Threshold on page 102

Enable Error Monitoring

The polling interval for Link Performance Monitoring is 100 milliseconds.

Task	Command Syntax	Command Mode
Start (or stop) Link Performance Monitoring on an interface.	**ethernet oam link-monitor on** **no ethernet oam link-monitor on** Default: Enabled	INTERFACE
Enable (or disable) support for Link Performance Monitoring on an interface.	**ethernet oam link-monitor supported** **no ethernet oam link-monitor supported** Default: Enabled	INTERFACE

Set Threshold Values

The available pre-defined errors fall under two categories:

- **Symbol Errors**—a *symbol* is an (electrical or optical) pulse on the physical medium that represents one or more bits. A symbol error occurs when a symbol degrades in transit so that the receiver is not able to decode it. Gigabit and 10-Gigabit Ethernet have and expect symbol rate, also called *Baud*.
- **Frame Errors**—frame errors are frames with a bad CRC.

The available pre-defined errors are:

- **Symbol Errors per Second**—the number of symbol errors during a specified period exceeds a threshold.
- **Frame Errors per Second**—the number of frame errors during a specified period exceeds a threshold.
- **Frame Errors per Frame Period**—the number of frame errors within the last *N* frames exceeds a threshold.
- **Frame Error Seconds per Time Period**—an *error second* is a 1-second period with at least one frame error. The Frame Error Seconds per Time Period error occurs when the number of error seconds within the last *M* seconds exceeds a threshold.

Symbol Errors per Second

Task	Command Syntax	Command Mode
Specify the high threshold value for symbol errors, or disable the high threshold.	**ethernet oam link-monitor symbol-period threshold high** {*symbols* | **none**} Range: 1-65535 Default: None	INTERFACE
Specify the low threshold for symbol errors.	**ethernet oam link-monitor symbol-period threshold low** *symbols* Range: 0-65535 Default: 10	INTERFACE
Specify the time period for symbol errors per second condition.	**ethernet oam link-monitor symbol-period window** *symbols* Range: 1-65535 (times 1,000,000 symbols) Default: 10 (10,000,000 symbols)	INTERFACE

Frame Errors per Second

Task	Command Syntax	Command Mode
Specify the high threshold value for frame errors, or disable the high threshold.	**ethernet oam link-monitor frame threshold high** {*frames* | **none**} Range: 1-65535 Default: None	INTERFACE
Specify the low threshold for frame errors.	**ethernet oam link-monitor frame threshold low** *frames* Range: 0-65535 Default: 1	INTERFACE
Specify the time period for frame errors per second condition.	**ethernet oam link-monitor frame window** *milliseconds* Range: 10-600 milliseconds Default: 100 milliseconds	INTERFACE

Frame Errors per Frame Period

Task	Command Syntax	Command Mode
Specify the high threshold value for frame errors per frame period, or disable the high threshold.	**ethernet oam link-monitor frame-period threshold high** {*frames* | **none**} Range: 1-65535 Default: None	INTERFACE
Specify the low threshold for frame errors per frame period.	**ethernet oam link-monitor frame-period threshold low** *frames* Range: 0-65535 Default: 1	INTERFACE
Specify the frame period for frame errors per frame period condition.	**ethernet oam link-monitor frame-period window** *milliseconds* Range: 1-65535 (times 10,000 frames) Default: 1000 (10 million frames)	INTERFACE

Error Seconds per Time Period

Task	Command Syntax	Command Mode
Specify the high threshold value for frame error seconds per time period, or disable the high threshold.	**ethernet oam link-monitor frame-seconds threshold high** {*milliseconds* | **none**} Range: 1-900 Default: None	INTERFACE
Specify the low threshold for frame error seconds per time period.	**ethernet oam link-monitor frame-seconds threshold low** *milliseconds* Range: 1-900 Default: 1	INTERFACE

Task	Command Syntax	Command Mode
Specify the time period for error second per time period condition.	**ethernet oam link-monitor frame-seconds window** *milliseconds* Range: 100-900, in multiples of 100 Default: 1000 milliseconds	INTERFACE

Execute an Action upon Exceeding the High Threshold

When an error exceeds the *low threshold*, an event notification is sent to the peer. When an error exceeds the *high threshold*, a pre-defined action is triggered such as disabling the interface.

Task	Command Syntax	Command Mode
Disable an interface when the high threshold is exceeded for any of the monitored error conditions.	**ethernet oam link-monitor high-threshold action error-disable-interface** Default: Enabled	INTERFACE

Remote Failure Indication

Remote Failure Indication is the mechanism a system uses to notify its peer of a local critical link event. There are three critical events; each has an associated flag which can be set in the OAMPDU when the event occurs.

- **Link Fault**—A fault occurred in the receive direction of the local peer.
- **Dying Gasp**—An unrecoverable local failure condition occurred. Dying Gasp notification is not supported on S-Series.
- **Critical Event**—An unspecified critical event occurred. Critical Event notification is not supported on S-Series.

When a link fault, dying gasp, or critical event occurs, the system sets an associated bit in subsequent OAMPDUs until the error is resolved (polling occurs every 100ms), and you can configure the system to take an additional action.

Task	Command Syntax	Command Mode
Block or disable an interface when a particular critical link event occurs.	**ethernet oam remote-failure** {**critical-event** | **dying-gasp** | **link-fault**} **action** {**error-block-interface** | **error-disable-interface**} Default: Disabled	INTERFACE

Remote Loopback

An active-mode device can place a passive peer into loopback mode by sending a Loopback Control OAMPDU. When in loopback mode:

- the remote peer returns unaltered all non-OAMPDU frames sent by the local peer, and
- all outbound data frames are discarded.

Note: Control traffic egresses from loopback initiator and from interface in loopback mode. You must explicitly disable L2/L3 protocols to stop control traffic.

Task	Command Syntax	Command Mode
Enable support for the OAM loopback capability on an interface so that it can exchange information with a remote peer.	**ethernet oam remote-loopback supported** Default: Enabled	INTERFACE
Configure the maximum amount of time the local peer waits for a frame to be returned before considering the remote peer to be non-operational.	**ethernet oam remote-loopback timeout** *seconds*	INTERFACE
Start or stop loopback operation on a local interface with a remote peer.	**ethernet oam remote-loopback** {**start** | **stop**} **interface** *interface*	EXEC Privilege

Display Link Layer OAM Configuration and Statistics

Task	Command Syntax	Command Mode
Display Link Layer OAM status per interface.	**show ethernet oam status interface** *interface*	EXEC Privilege

```
FTOS# show ethernet oam status interface <interface-name>

Output Format :

<interface-name>

General
_______
Mode:active
PDU max rate:10 packets per second
PDU min rate:1 packet per second
Link timeout:5 seconds
High threshold action:no action

Link Monitoring
_____________
Status supported (on)

Symbol Period Error
  Window:1 million symbols
  Low threshold:1 error symbol(s)
  High threshold:none
Frame Error
  Window:1 million symbols
  Low threshold:1 error symbol(s)
  High threshold:none
Frame Period Error
  Window:1 x 100,000 frames
  Low threshold:1 error symbol(s)
  High threshold:none
Frame Seconds Error
  Window:600 x 100 milliseconds
  Low threshold:1 error second(s)
  High threshold:none
```

Display Link Layer OAM statistics per interface.	**show ethernet oam statistics interface** *interface*	EXEC Privilege

Task	Command Syntax	Command Mode

```
FTOS# show ethernet oam statistics interface <interface-name>

<interface-name>
Counters:
_________
Information OAMPDU Tx: 3439489
Information OAMPDU Rx: 9489
Unique Event Notification OAMPDU Tx: 0
Unique Event Notification OAMPDU x: 0
Duplicate Event Notification OAMPDU Tx: 0
Duplicate Event Notification OAMPDU Rx: 0
Loopback Control OAMPDU Tx: 0
Loopback Control OAMPDU Rx: 2
Variable Request OAMPDU Tx: 0
Variable Request OAMPDU Rx: 0
Variable Response OAMPDU Tx: 0
Variable Response OAMPDU Rx: 0
Force10 OAMPDU Tx:: 10
Force10 OAMPDU Rx:: 21
Unsupported OAMPDU Tx:: 0
Unsupported OAMPDU Rx:0
Frame Lost due to OAM:0

Local Faults:
0 Link Fault Records
0 Dying Gasp Records
Total dying Gasps:: 2
Time Stamp: 00:40:23
Total dying Gasps:: 1
Time Stamp: 00:41:23
0 Critical Event Records

Remote Faults:
_________
0 Link Fault Records
0 Dying Gasp Records
0 Critical Event Records

Local Event Logs:
_____________
0 Errored Symbol Period Records
0 Errored Frame Records
0 Errored Frame Period Records
0 Errored Frame Second Records

Remote Event Logs:
_____________
0 Errored Symbol Period Records
0 Errored Frame Records
0 Errored Frame Period Records
0 Errored Frame Second Records
```

Clear Link Layer OAM statistics.	**clear ethernet oam statistics interface** *interface*	EXEC Privilege

Manage Link Layer OAM

Enable MIB Retrieval Support/Function

IEEE 802.3ah defines the Link OAM MIB in Sec 30A.20, “OAM entity managed object class”; all of the objects described there are supported. Note that 802.3ah does not include the ability to set/write remote MIB variables.

You must enable MIB retrieval support and the MIB retrieval function.

Task	Command Syntax	Command Mode
Enable MIB retrieval support and/or the MIB retrieval function.	**ethernet oam mib-retrieval** {**supported** | **on**} Default: Disabled	INTERFACE

Adjust the Size of the Link OAM Event Log

Task	Command Syntax	Command Mode
Configure the size of the OAM event log.	**ethernet oam event-log size** *entries* Range: 0 to 200. Default: 50.	CONFIGURATION

7

802.1X

802.1X is supported on platforms: C E S

This chapter has the following sections:

Protocol Overview

802.1X is a method of port security. A device connected to a port that is enabled with 802.1X is disallowed from sending or receiving traffic on the network until its identity can be verified (through a username and password, for example); all ingress frames, except those used for 802.1X authentication, are dropped. This feature is named for its IEEE specification.

802.1X employs Extensible Authentication Protocol (EAP)* to transfer a device's credentials to an authentication server (typically RADIUS) via a mandatory intermediary network access device, in this case, a Dell Force10 switch. The network access device mediates all communication between the end-user device and the authentication server so that the network remains secure. The network access device uses EAP over Ethernet (EAPOL) to communicate with the end-user device and EAP over RADIUS to communicate with the server.

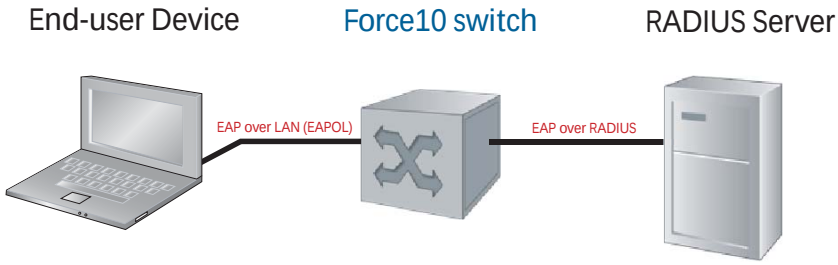
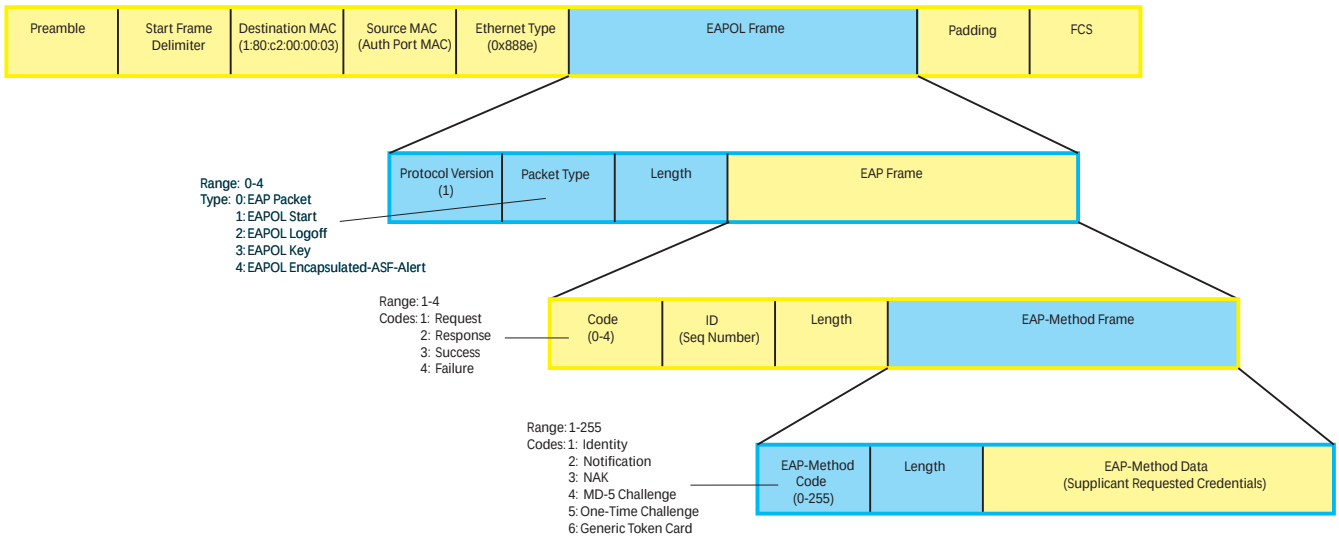


Figure 7-1 and Figure show how EAP frames are encapsulated in Ethernet and Radius frames.

Note: FTOS supports 802.1X with EAP-MD5, EAP-OTP, EAP-TLS, EAP-TTLS, PEAPv0, PEAPv1, and MS-CHAPv2 with PEAP.

Figure 7-1. EAPOL Frame Format



The authentication process involves three devices:

- The device attempting to access the network is the **supplicant**. The supplicant is not allowed to communicate on the network until the port is authorized by the authenticator. It can only communicate with the authenticator in response to 802.1X requests.
- The device with which the supplicant communicates is the **authenticator**. The authenicator is the gate keeper of the network. It translates and forwards requests and responses between the authentication server and the supplicant. The authenticator also changes the status of the port based on the results of the authentication process. The Dell Force10 switch is the authenticator.

- The **authentication-server** selects the authentication method, verifies the information provided by the supplicant, and grants it network access privileges.

Ports can be in one of two states:

- Ports are in an **unauthorized** state by default. In this state, non-802.1X traffic cannot be forwarded in or out of the port.
- The authenticator changes the port state to **authorized** if the server can authenticate the supplicant. In this state, network traffic can be forwarded normally.

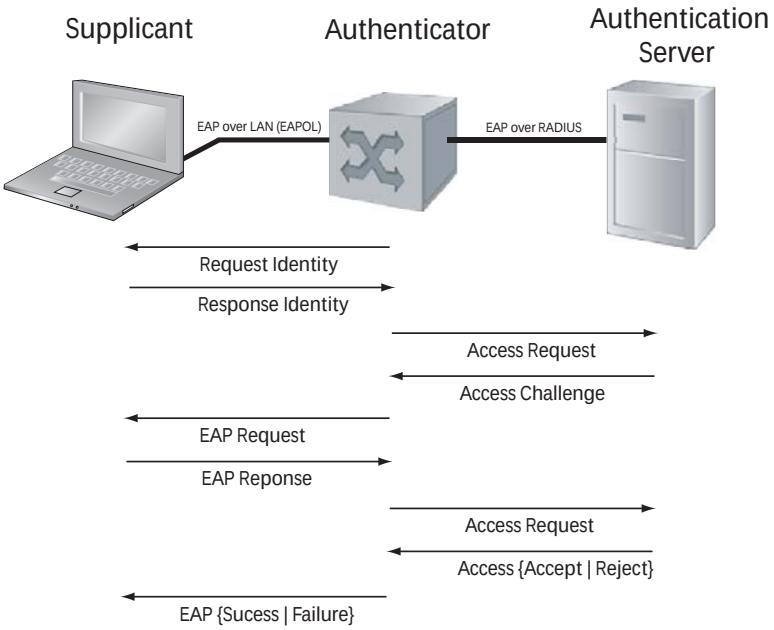
Note: The Dell Force10 switches place 802.1X-enabled ports in the unauthorized state by default.

The Port-authentication Process

The authentication process begins when the authenticator senses that a link status has changed from down to up:

1. When the authenticator senses a link state change, it requests that the supplicant identify itself using an EAP Identity Request Frame.
2. The supplicant responds with its identity in an EAP Response Identity frame.
3. The authenticator decapsulates the EAP Response from the EAPOL frame, encapsulates it in a RADIUS Access-Request frame, and forwards the frame to the authentication server.
4. The authentication server replies with an Access-Challenge. The Access-Challenge is request that the supplicant prove that it is who it claims to be, using a specified method (an EAP-Method). The challenge is translated and forwarded to the supplicant by the authenticator.
5. The supplicant can negotiate the authentication method, but if it is acceptable, the supplicant provides the requested challenge information in an EAP Response, which is translated and forwarded to the authentication server as another Access-Request.
6. If the identity information provided by the supplicant is valid, the authentication server sends an Access-Accept frame in which network privileges are specified. The authenticator changes the port state to authorized, and forwards an EAP Success frame. If the identity information is invalid, the server sends and Access-Reject frame. The port state remains unauthorized, and the authenticator forwards EAP Failure frame.

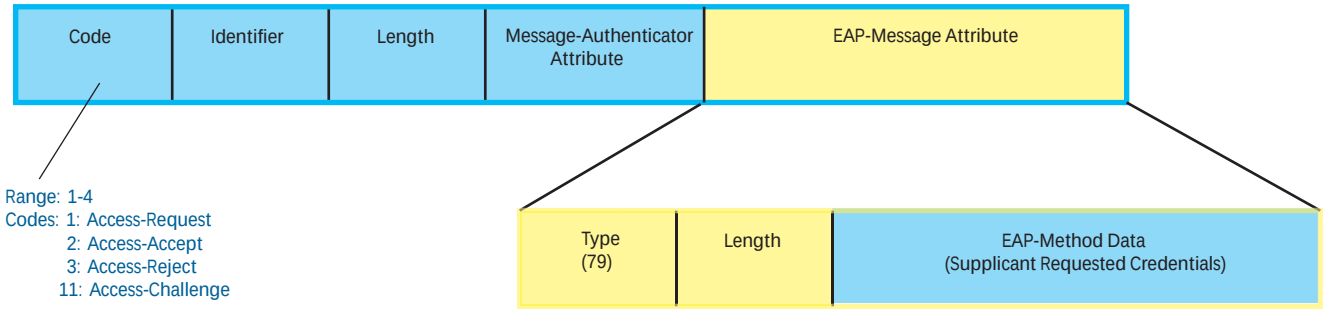
Figure 7-2. 802.1X Authentication Process



EAP over RADIUS

802.1X uses RADIUS to shuttle EAP packets between the authenticator and the authentication server, as defined in RFC 3579. EAP messages are encapsulated in RADIUS packets as a type of *attribute* in Type, Length, Value (TLV) format. The Type value for EAP messages is 79.

Figure 7-3. RADIUS Frame Format



RADIUS Attributes for 802.1 Support

Dell Force10 systems includes the following RADIUS attributes in all 802.1X-triggered Access-Request messages:

Table 7-1. 802.1X Supported RADIUS Attributes

Attribute	Name	Description
1	User-Name	the name of the supplicant to be authenticated.
4	NAS-IP-Address	
5	NAS-Port	the physical port number by which the authenticator is connected to the supplicant.
24	State	
30	Called-Station-Id	
31	Calling-Station-Id	relays the supplicant MAC address to the authentication server.
61	NAS-Port-Type	NAS-port physical port type. 5 indicates Ethernet.
64	Tunnel-Type	
65	Tunnel-Medium-Type	
79	EAP-Message	encapsulates EAP packets
80	Message-Authenticator	a calculated value included in Access-Requests to prevent spoofing.
81	Tunnel-Private-Group-ID	associate a tunneled session with a particular group of users.

Configuring 802.1X

Configuring 802.1X on a port is a two-step process:

1. Enable 802.1X globally. See page 112.
2. Enable 802.1X on an interface. See page 112.

Related Configuration Tasks

- Configuring Request Identity Re-transmissions on page 114
- Configuring Port-control on page 116
- Re-Authenticating a Port on page 116
- Configuring Timeouts on page 117
- Configuring a Guest VLAN on page 121
- Configuring an Authentication-Fail VLAN on page 122

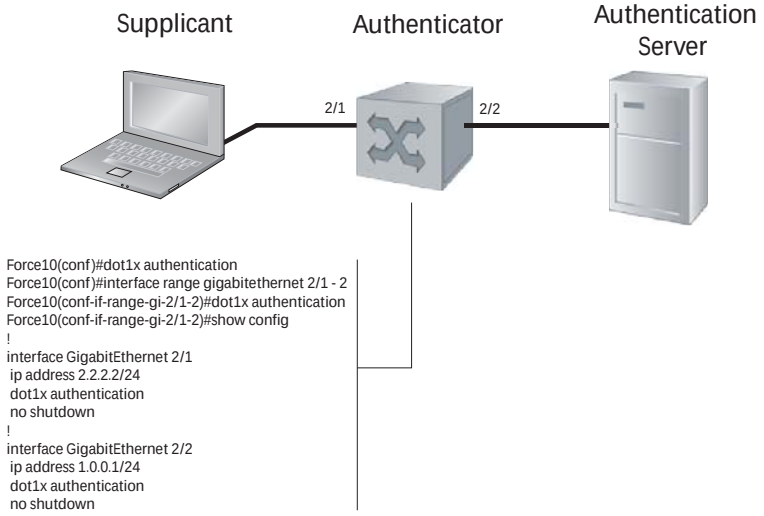
Important Points to Remember

- FTOS supports 802.1X with EAP-MD5, EAP-OTP, EAP-TLS, EAP-TTLS, PEAPv0, PEAPv1, and MS-CHAPv2 with PEAP.
- All platforms support only RADIUS as the authentication server.
- On E-Series ExaScale, if the primary RADIUS server becomes unresponsive, the authenticator begins using a secondary RADIUS server, if configured.
- 802.1X is not supported on port-channels or port-channel members.
- **On the C-series and S-Series platforms**:
  - Traffic may be forwarded on an 802.1X-enabled port that is in an unauthorized state and interoperates with a device through a MAC-authentication bypass (MAB) or the guest VLAN. 802.1X authentication on the port returns to normal operation only after a port flap or if you disable and then re-enable 802.1X authentication on the port.
  - If you enable multi-supplicant authorization on a port, configure a maximum number of supplicants that can be authenticated, and enable periodic re-authentication, if some of the supplicants fail re-authentication, these unauthorized supplicants are still counted in the total number of supplicants that can access the port.
  - Traffic may be transmitted on an 802.1X-enabled port before the port changes to an authorized state.
  - A MAB-authenticated port becomes unauthorized after an RPM failover.

Enabling 802.1X

802.1X must be enabled globally and at interface level.

Figure 7-4. Enabling 802.1X



To enable 802.1X:

Step	Task	Command Syntax	Command Mode
1	Enable 802.1X globally.	**dot1x authentication**	CONFIGURATION
2	Enter INTERFACE mode on an interface or a range of interfaces.	**interface** [**range**]	INTERFACE
3	Enable 802.1X on an interface or a range of interfaces.	**dot1x authentication**	INTERFACE

Verify that 802.1X is enabled globally and at interface level using the command **show running-config | find dot1x** from EXEC Privilege mode, as shown in Figure 7-5.

Figure 7-5. Verifying 802.1X Global Configuration

```
FTOS#show running-config | find dot1x
dot1x authentication                    <-- 802.1X Enabled
!
[output omitted]
!
interface GigabitEthernet 2/1
 ip address 2.2.2.2/24
 dot1x authentication                   <-- 802.1X Enabled on
 no shutdown
!
interface GigabitEthernet 2/2
 ip address 1.0.0.1/24
 dot1x authentication
 no shutdown
--More--
```

View 802.1X configuration information for an interface using the command **show dot1x interface**, as shown in Figure 7-6.

Figure 7-6. Verifying 802.1X Interface Configuration

```
FTOS#show dot1x interface gigabitethernet 2/1

802.1x information on Gi 2/1:
-----------------------------
Dot1x Status:         Enable        <-- 802.1X Enabled on Interface
Port Control:         AUTO
Port Auth Status:     UNAUTHORIZED  <-- All ports unauthorized by default
Re-Authentication:    Disable
Untagged VLAN id:     None
Tx Period:            30 seconds
Quiet Period:         60 seconds
ReAuth Max:           2
Supplicant Timeout:   30 seconds
Server Timeout:       30 seconds
Re-Auth Interval:     3600 seconds
Max-EAP-Req:          2
Auth Type:            SINGLE_HOST

Auth PAE State:       Initialize
Backend State:        Initialize
```

Configuring Request Identity Re-transmissions

If the authenticator sends a Request Identity frame, but the supplicant does not respond, the authenticator waits 30 seconds and then re-transmits the frame. The amount of time that the authenticator waits before re-transmitting and the maximum number of times that the authenticator re-transmits are configurable.

Note: There are several reasons why the supplicant might fail to respond; the supplicant might have been booting when the request arrived, there might be a physical layer problem, or the supplicant might not be 802.1x capable.

To configure the amount of time that the authenticator waits before re-transmitting an EAP Request Identity frame:

Step	Task	Command Syntax	Command Mode
1	Configure the amount of time that the authenticator waits before re-transmitting an EAP Request Identity frame.	**dot1x tx-period** *number* Range: 1-31536000 (1 year) Default: 30	INTERFACE

To configure a maximum number of Request Identity re-transmissions:

Step	Task	Command Syntax	Command Mode
1	Configure a maximum number of times that a Request Identity frame can be re-transmitted by the authenticator.	**dot1x max-eap-req** *number* Range: 1-10 Default: 2	INTERFACE

Figure 7-7 shows configuration information for a port for which the authenticator re-transmits an EAP Request Identity frame after 90 seconds and re-transmits a maximum of 10 times.

Configuring a Quiet Period after a Failed Authentication

If the supplicant fails the authentication process, the authenticator sends another Request Identity frame after 30 seconds by default, but this period can be configured.

Note: The quiet period (**dot1x quiet-period**) is an transmit interval for after a failed authentication where as the Request Identity Re-transmit interval (**dot1x tx-period**) is for an unresponsive supplicant.

To configure the quiet period after a failed authentication:

Step	Task	Command Syntax	Command Mode
1	Configure the amount of time that the authenticator waits to re-transmit a Request Identity frame after a failed authentication.	**dot1x quiet-period** *seconds* Range: 1-65535 Default: 60	INTERFACE

Figure 7-7 shows configuration information for a port for which the authenticator re-transmits an EAP Request Identity frame:

- After 90 seconds and a maximum of 10 times for an unresponsive supplicant
- Re-transmits an EAP Request Identity frame

Figure 7-7. Configuring a Request Identity Re-transmissions

```
FTOS(conf-if-range-gi-2/1)#dot1x tx-period 90
FTOS(conf-if-range-gi-2/1)#dot1x max-eap-req 10
FTOS(conf-if-range-gi-2/1)#dot1x quiet-period 120
FTOS#show dot1x interface gigabitethernet 2/1

802.1x information on Gi 2/1:
-----------------------------
Dot1x Status:          Enable
Port Control:          AUTO
Port Auth Status:      UNAUTHORIZED
Re-Authentication:     Disable      <-- New Re-transmit Interval
Untagged VLAN id:      None
Tx Period:             90 seconds
Quiet Period:          120 seconds  <-- New Quiet Period
ReAuth Max:            2
Supplicant Timeout:    30 seconds
Server Timeout:        30 seconds
Re-Auth Interval:      3600 seconds
Max-EAP-Req:           10           <-- New Maximum Re-transmissions
Auth Type:             SINGLE_HOST

Auth PAE State:        Initialize
Backend State:         Initialize
```

Forcibly Authorizing or Unauthorizing a Port

IEEE 802.1X requires that a port can be manually placed into any of three states:

- **ForceAuthorized** is an authorized state. A device connected to this port in this state is never subjected to the authentication process, but is allowed to communicate on the network. Placing the port in this state is same as disabling 802.1X on the port.
- **ForceUnauthorized** an unauthorized state. A device connected to a port in this state is never subjected to the authentication process and is not allowed to communicate on the network. Placing the port in this state is the same as shutting down the port. Any attempt by the supplicant to initiate authentication is ignored.

Note: On the C-Series, traffic may continue to be transmitted after an 802.1x-enabled port is configured as **force-unauthorized**.

- **Auto** is an unauthorized state by default. A device connected to this port is this state is subjected to the authentication process. If the process is successful, the port is authorized and the connected device can communicate on the network. All ports are placed in the **auto** state by default.

To place a port in one of these three states:

Step	Task	Command Syntax	Command Mode
1	Place a port in the ForceAuthorized, ForceUnauthorized, or Auto state.	**dot1x port-control** {**force-authorized** | **force-unauthorized** | **auto**} Default: auto	INTERFACE

Figure 7-8 shows configuration information for a port that has been force-authorized.

Figure 7-8. Configuring Port-control

```
FTOS(conf-if-gi-2/1)#dot1x port-control force-authorized
FTOS(conf-if-gi-2/1)#do show dot1x interface gigabitethernet 2/1

802.1x information on Gi 2/1:
-----------------------------
Dot1x Status:          Enable
Port Control:          FORCE_AUTHORIZED   <-- New Port-control State
Port Auth Status:      UNAUTHORIZED
Re-Authentication:     Disable
Untagged VLAN id:      None
Tx Period:             90 seconds
Quiet Period:          120 seconds
ReAuth Max:            2
Supplicant Timeout:    30 seconds
Server Timeout:        30 seconds
Re-Auth Interval:      3600 seconds
Max-EAP-Req:           10
Auth Type:             SINGLE_HOST

Auth PAE State:        Initialize
Backend State:         Initialize
Auth PAE State:        Initialize
```

Re-Authenticating a Port

Periodic Re-Authentication

After the supplicant has been authenticated and the port has been authorized, the authenticator can be configured to re-authenticate the supplicant periodically. If re-authentication is enabled, the supplicant is required to re-authenticate every 3600 seconds, but this interval can be configured. A maximum number of re-authentications can be configured as well.

To configure a re-authentication or a re-authentication period:

Step	Task	Command Syntax	Command Mode
1	Configure the authenticator to periodically re-authenticate the supplicant.	**dot1x reauthentication** [**interval**] *seconds* Range: 1-65535 Default: 60	INTERFACE

To configure a maximum number of re-authentications:

Step	Task	Command Syntax	Command Mode
1	Configure the maximum number of times that the supplicant can be reauthenticated.	**dot1x reauth-max** *number* Range: 1-10 Default: 2	INTERFACE

Figure 7-9. Configuring a Reauthentiction Period

```
FTOS(conf-if-gi-2/1)#dot1x reauthentication interval 7200
FTOS(conf-if-gi-2/1)#dot1x reauth-max 10
FTOS(conf-if-gi-2/1)#do show dot1x interface gigabitethernet 2/1

802.1x information on Gi 2/1:
-----------------------------
Dot1x Status:          Enable
Port Control:          FORCE_AUTHORIZED
Port Auth Status:      UNAUTHORIZED      <-- Re-authentication Enabled
Re-Authentication:     Enable
Untagged VLAN id:      None
Tx Period:             90 seconds
Quiet Period:          120 seconds
ReAuth Max:            10                <-- New Maximum Re-authentications
Supplicant Timeout:    30 seconds
Server Timeout:        30 seconds
Re-Auth Interval:      7200 seconds      <-- New Re-authentication Period
Max-EAP-Req:           10
Auth Type:             SINGLE_HOST

Auth PAE State:        Initialize
Backend State:         Initialize
Auth PAE State:        Initialize
Backend State:         Initialize
```

Configuring Timeouts

If the supplicant or the authentication server is unresponsive, the authenticator terminates the authentication process after 30 seconds by default. The amount of time that the authenticator waits for a response can be configured. The timeout for the supplicant applies to all EAP frames except for Request Identity frames which are governed by the **tx-period** and **max-eap-req** configurations.

To terminate the authentication process due to an unresponsive supplicant:

Step	Task	Command Syntax	Command Mode
1	Terminate the authentication process due to an unresponsive supplicant.	**dot1x supplicant-timeout** *seconds* Range: 1-300. Default: 30	INTERFACE

To terminate the authentication process due to an unresponsive authentication server:

Step	Task	Command Syntax	Command Mode
1	Terminate the authentication process due to an unresponsive authentication server.	**dot1x server-timeout** *seconds* Range: 1-300. Default: 30	INTERFACE

Note: When you configure the **dot1x server-timeout** value, you must take into account the communication medium used to communicate with an authentication server and the number of RADIUS servers configured. Ideally, the **dot1x server-timeout** value (in seconds) is based on the configured RADIUS-server timeout and retransmit values and calculated according to the following formula:
dot1x server-timeout *seconds* > (**radius-server retransmit** *seconds* + 1) * **radius-server timeout** *seconds*

Where the default values are as follows: **dot1x server-timeout** (30 seconds), **radius-server retransmit** (3 seconds), and **radius-server timeout** (5 seconds).

For example:

```
FTOS(conf)#radius-server host 10.11.197.105 timeout 6
FTOS(conf)#radius-server host 10.11.197.105 retransmit 4
FTOS(conf)#interface gigabitethernet 2/23
FTOS(conf-if-gi-2/23)#dot1x server-timeout 40
```

Figure 7-10 shows configuration information for a port for which the authenticator terminates the authentication process for an unresponsive supplicant or server after 15 seconds.

Figure 7-10. Configuring a Timeout

```
FTOS(conf-if-gi-2/1)#dot1x port-control force-authorized
FTOS(conf-if-gi-2/1)#do show dot1x interface gigabitethernet 2/1

802.1x information on Gi 2/1:
-----------------------------
Dot1x Status:              Enable
Port Control:              FORCE_AUTHORIZED
Port Auth Status:          UNAUTHORIZED
Re-Authentication:         Disable
Untagged VLAN id:          None
Guest VLAN:                Disable
Guest VLAN id:             NONE
Auth-Fail VLAN:            Disable
Auth-Fail VLAN id:         NONE
Auth-Fail Max-Attempts:    NONE
Tx Period:                 90 seconds
Quiet Period:              120 seconds
ReAuth Max:                10
Supplicant Timeout:        15 seconds
Server Timeout:            15 seconds
Re-Auth Interval:          7200 seconds
Max-EAP-Req:               10
Auth Type:                 SINGLE_HOST

Auth PAE State:            Initialize
Backend State:             Initialize
```

New Supplicant and Server Timeouts

Dynamic VLAN Assignment with Port Authentication

Dynamic VLAN Assignment with Port Authentication is supported on platforms: C S E(T)

FTOS supports dynamic VLAN assignment when using 802.1X. During 802.1x authentication, the existing VLAN configuration of a port assigned to a non-default VLAN is overwritten and the port is assigned to a specified VLAN.

- If 802.1x authentication is disabled on the port, the port is re-assigned to the previously-configured VLAN.
- If 802.1x authentication fails and if the authentication-fail VLAN is enabled for the port (see Configuring an Authentication-Fail VLAN on page 122), the port is assigned to the authentication-fail VLAN.

The dynamic VLAN assignment is based on RADIUS attribute 81, Tunnel-Private-Group-ID, and uses the following standard dot1x procedure:

1. The host sends a dot1x packet to the Dell Force10 system.
2. The system forwards a RADIUS REQUEST packet containing the host MAC address and ingress port number.
3. The RADIUS server authenticates the request and returns a RADIUS ACCEPT message with the VLAN assignment using Tunnel-Private-Group-ID.

   The dynamic VLAN assignment from the RADIUS server always overrides the configuration on the switch for the given port. This applies to ports already configured with a non-default VLAN.

Note: For the C-Series, S-Series, and E-Series TeraScale platforms, the dynamic VLAN assignment fails if a port is assigned to a non-default VLAN and if the non-default VLAN assignment was configured on an FTOS version earlier than 8.4.2.3.

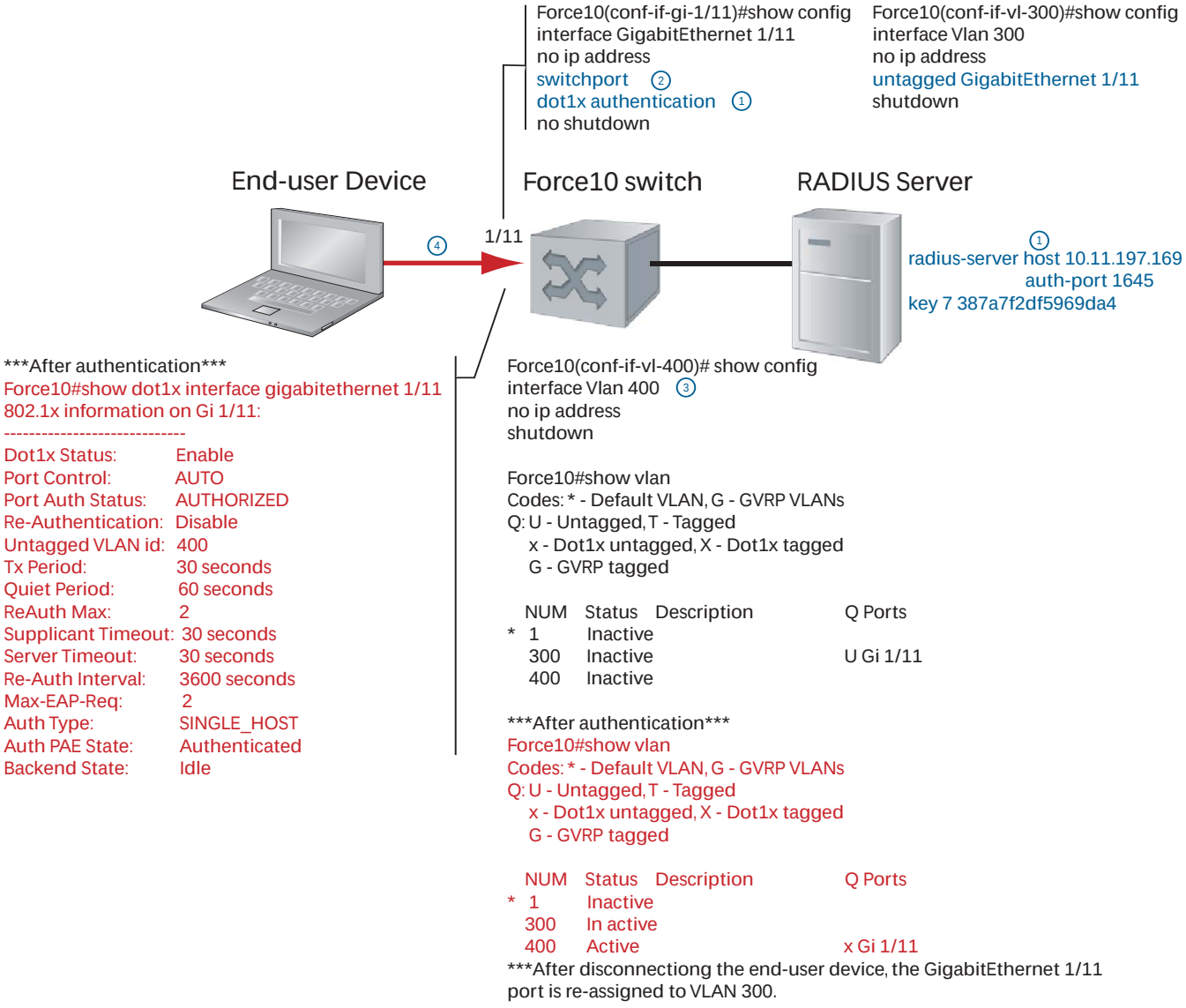
To configure dynamic VLAN assignment with 802.1x port authentication:

Step	Task
1	Configure 802.1x globally and at interface level (see Enabling 802.1X on page 112) along with relevant RADIUS server configurations.
2	Make the interface a switchport so that it can be assigned to a VLAN.
3	Create the VLAN to which the interface will be assigned.
4	Connect the supplicant to the port configured for 802.1X.
5	Verify that the port has been authorized and placed in the desired VLAN by entering the **show dot1x interface** and **show vlan** commands (red text in Figure 7-11).

Figure 7-11 shows the configuration on a Dell Force10 switch that uses dynamic VLAN assignment with 802.1X before you connect the end-user device (black and blue text), and after you connect the device (red text).

The blue text corresponds to the numbered steps on page 119. Note that the GigabitEthernet 1/11 port, on which dynamic VLAN assignment with 802.1X is configured, is initially an untagged member of VLAN 300. After a successful 802.1x authentication with dynamic VLAN configuration, the port becomes an untagged member of VLAN 400 (assigned by the RADIUS server during authentication).

Figure 7-11. Dynamic VLAN Assignment with 802.1X



Note: In the **show vlan** command output, if the statically-configured VLAN and the 802.1X dynamically-assigned VLAN are the same, the 802.1x-authorized port is displayed with `U` for `Untagged`. If the two VLANs are not the same, the 802.1x-authorized port is displayed with `x` for `Dot1X untagged`.

Guest and Authentication-Fail VLANs

Typically, the authenticator (Dell Force10 system) denies the supplicant access to the network until the supplicant is authenticated. If the supplicant is authenticated, the authenticator enables the port and places it in either the VLAN for which the port is configured, or the VLAN that the authentication server indicates in the authentication data.

Note: Ports cannot be dynamically assigned to the default VLAN.

If the supplicant fails authentication, the authenticator typically does not enable the port. In some cases this behavior is not appropriate. External users of an enterprise network, for example, might not be able to be authenticated, but still need access to the network. Also, some dumb-terminals such as network printers do not have 802.1X capability and therefore cannot authenticate themselves. To be able to connect such devices, they must be allowed access the network without compromising network security.

The Guest VLAN 802.1X extension addresses this limitation with regard to non-802.1X capable devices, and the Authentication-fail VLAN 802.1X extension addresses this limitation with regard to external users.

- If the supplicant fails authentication a specified number of times, the authenticator places the port in the Authentication-fail VLAN.
- If a port is already forwarding on the Guest VLAN when 802.1X is enabled, then the port is moved out of the Guest VLAN, and the authentication process begins.

Configuring a Guest VLAN

If the supplicant does not respond to a Request Identity frame within a determined amount of time ([*reauth-max* + 1] * *tx-period*, see Configuring Request Identity Re-transmissions on page 114) the system assumes that the host does not have 802.1X capability, and the port is placed in the Guest VLAN.

Configure a port to be placed in the Guest VLAN after failing to respond within the timeout period using the command **dot1x guest-vlan** from INTERFACE mode, as shown in Figure 7-12.

Figure 7-12. Configuring a Guest VLAN

```
FTOS(conf-if-gi-1/2)#dot1x guest-vlan 200
FTOS(conf-if-gi-1/2)#show config
!
interface GigabitEthernet 1/2
 switchport
 dot1x guest-vlan 200
 no shutdown
FTOS(conf-if-gi-1/2)#
```

View your configuration using the command **show config** from INTERFACE mode, as shown in Figure 7-12, or using the command **show dot1x interface** command from EXEC Privilege mode as shown in Figure 7-14.

Configuring an Authentication-Fail VLAN

If the supplicant fails authentication, the authenticator re-attempts to authenticate after a specified amount of time (30 seconds by default, see Configuring a Quiet Period after a Failed Authentication on page 114). You can configure the maximum number of times the authenticator re-attempts authentication after a failure (3 by default), after which the port is placed in the Authentication-fail VLAN.

Configure a port to be placed in the VLAN after failing the authentication process as specified number of times using the command **dot1x auth-fail-vlan** from INTERFACE mode, as shown in Figure 7-13. Configure the maximum number of authentication attempts by the authenticator using the keyword **max-attempts** with this command.

Figure 7-13. Configuring an Authentication-fail VLAN

```
FTOS(conf-if-gi-1/2)#dot1x auth-fail-vlan 100 max-attempts 5
FTOS(conf-if-gi-1/2)#show config
!
interface GigabitEthernet 1/2
 switchport
 dot1x guest-vlan 200
 dot1x auth-fail-vlan 100 max-attempts 5
 no shutdown
```

View your configuration using the command **show config** from INTERFACE mode, as shown in Figure 7-12, or using the command **show dot1x interface** command from EXEC Privilege mode as shown in Figure 7-14.

Figure 7-14. Viewing Guest and Authentication-fail VLAN Configurations

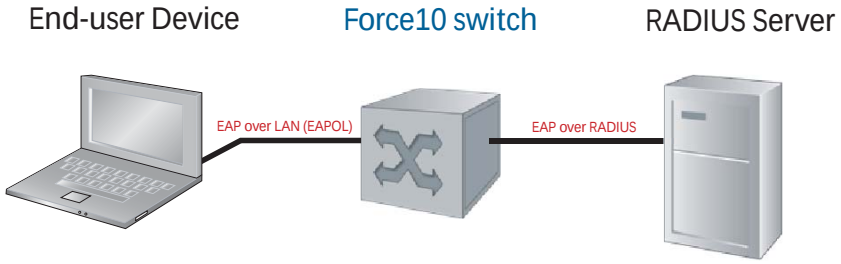
```
FTOS(conf-if-gi-2/1)#dot1x port-control force-authorized
FTOS(conf-if-gi-2/1)#do show dot1x interface gigabitethernet 2/1

802.1x information on Gi 2/1:
-----------------------------
Dot1x Status:              Enable
Port Control:              FORCE_AUTHORIZED
Port Auth Status:          UNAUTHORIZED
Re-Authentication:         Disable
Untagged VLAN id:          None
Guest VLAN:                Enable
Guest VLAN id:             200
Auth-Fail VLAN:            Enable
Auth-Fail VLAN id:         100
Auth-Fail Max-Attempts:    5
Tx Period:                 90 seconds
Quiet Period:              120 seconds
ReAuth Max:                10
Supplicant Timeout:        15 seconds
Server Timeout:            15 seconds
Re-Auth Interval:          7200 seconds
Max-EAP-Req:               10
Auth Type:                 SINGLE_HOST

Auth PAE State:            Initialize
Backend State:             Initialize
```

Multi-Host Authentication

Multi-Host Authentication is available on platforms: C E T S

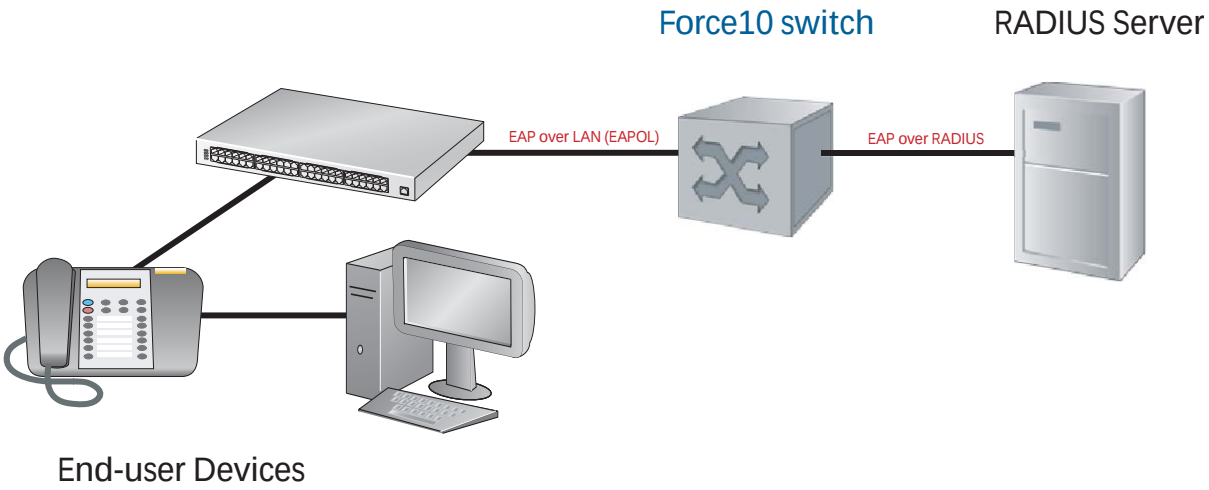
802.1x assumes that a single end-user is connected to a single authenticator port, as shown in Figure 7-15; this one-to-one mode of authentication is called Single-host mode. If multiple end-users are connected to the same port, a many-to-one configuration, only the first end-user to respond to the identity request is authenticated. Subsequent responses are ignored, and a system log is generated to indicate reception of unexpected 802.1x frames. When a port is authorized, the authenticated supplicant MAC address is associated with the port, and traffic from any other source MACs is dropped.

Figure 7-15. Single-host Authentication Mode



When multiple end-users are connected to a single authenticator port, Single-host mode authentication does not authenticate all end-users, and all but one are denied access to the network. For these cases (Figure 7-16), FTOS offers Multi-host mode authentication.

Figure 7-16. Multi-host Authentication Mode



When Multi-host mode authentication is configured, the first client to respond to an identity request is authenticated, and subsequent responses are still ignored, but since the authenticator expects the possibility of multiple responses, no system log is generated. After the first supplicant is authenticated, all end-users attached to the authorized port are allowed to access the network.

If the authorized port becomes unauthorized due to re-authentication failure or the supplicant sends an EAPOL logoff frame, all attached end-users are denied access to the network.

When the host mode is changed on a port that is already authenticated:

- **Single-host to Multi-host**: all devices attached to the port that were previously blocked may access the network; the supplicant does not re-authenticate.
- **Multi-host to Single-host**: the port restarts the authentication process, and the first end-user to respond is authenticated and allowed access.

Task	Command Syntax	Command Mode
Configure Multi-host Authentication mode on a port. Enter **no dot1x host-mode** to return to Single-host mode.	**dot1x host-mode multi-host** Default: Single-host mode	INTERFACE

```
FTOS(conf-if-gi-2/1)#dot1x port-control force-authorized
FTOS(conf-if-gi-2/1)#do show dot1x interface gigabitethernet 2/1

802.1x information on Gi 2/1:
-----------------------------
Dot1x Status:              Enable
Port Control:              FORCE_AUTHORIZED
Port Auth Status:          UNAUTHORIZED
Re-Authentication:         Disable
Untagged VLAN id:          None
Guest VLAN:                Enable
Guest VLAN id:             200
Auth-Fail VLAN:            Enable
Auth-Fail VLAN id:         100
Auth-Fail Max-Attempts:    5
Tx Period:                 90 seconds
Quiet Period:              120 seconds
ReAuth Max:                10
Supplicant Timeout:        15 seconds
Server Timeout:            15 seconds
Re-Auth Interval:          7200 seconds
Max-EAP-Req:               10
Host Mode:                 MULTI_HOST

Auth PAE State:            Initialize
Backend State:             Initialize
```

Task	Command Syntax	Command Mode
Configure Single-host Authentication mode on a port.	**dot1x host-mode single-host**	INTERFACE

```
FTOS(conf-if-gi-2/1)#dot1x port-control force-authorized
FTOS(conf-if-gi-2/1)#do show dot1x interface gigabitethernet 2/1

802.1x information on Gi 2/1:
-----------------------------
Dot1x Status:              Enable
Port Control:              FORCE_AUTHORIZED
Port Auth Status:          UNAUTHORIZED
Re-Authentication:         Disable
Untagged VLAN id:          None
Guest VLAN:                Enable
Guest VLAN id:             200
Auth-Fail VLAN:            Enable
Auth-Fail VLAN id:         100
Auth-Fail Max-Attempts:    5
Tx Period:                 90 seconds
Quiet Period:              120 seconds
ReAuth Max:                10
Supplicant Timeout:        15 seconds
Server Timeout:            15 seconds
Re-Auth Interval:          7200 seconds
Max-EAP-Req:               10
Host Mode:                 SINGLE_HOST

Auth PAE State:            Initialize
Backend State:             Initialize
```

Multi-Supplicant Authentication

Multi-Supplicant Authentication is available on platforms: C S

The 802.1X Multi-supplicant Authentication enables multiple devices on a single authenticator port to access the network by authenticating each device. In addition, Multi-supplicant Authentication uses dynamic MAC-based VLAN assignment to place devices on different VLANs. This feature is different from Multi-host Authentication in which multiple devices connected to a single authenticator port can access the network after only the one device is authenticated, and all hosts are placed in the same VLAN as the authenticated device.

Multi-supplicant authentication is needed, for example, in the case of a workstation at which a VOIP phone and PC are connected to a single authenticator port. Multi-host authentication could authenticate the first device to respond, and then both devices could access the network. However, if you wanted to place them in different VLANs—a VOIP VLAN and a data VLAN— you would need to authenticate the devices separately so that the RADIUS server can send each device's VLAN assignment during that devices authentication process.

During the authentication process, the Dell Force10 system is able to learn the MAC address of the device though the EAPoL frames, and the VLAN assignment from the RADIUS server. With this information it creates an authorized-MAC to VLAN mapping table per port. Then, the system can tag all incoming untagged frames with the appropriate VLAN-ID based on the table entries.

Task	Command Syntax	Command Mode
Enable Multi-Supplicant Authentication mode on a port.	**dot1x host-mode multi-auth** Default: Single-host mode	INTERFACE

```
FTOS#show dot1x interface gigabitethernet 1/3 details

802.1x information on Gi 1/3:
------------------------------
Dot1x Status:              Enable
Port Control:              AUTO
Port Auth Status:          MULTI-AUTH
Re-Authentication:         Disable
Untagged VLAN id:          None
Guest VLAN:                Disable
Guest VLAN id:             NONE
Auth-Fail VLAN:            Disable
Auth-Fail VLAN id:         NONE
Auth-Fail Max-Attempts:    NONE
Tx Period:                 30 seconds
Quiet Period:              60 seconds
ReAuth Max:                2
Supplicant Timeout:        30 seconds
Server Timeout:            30 seconds
Re-Auth Interval:          3600 seconds
Max-EAP-Req:               2
Host Mode:                 MULTI-AUTH
Auth PAE State:            Initialize
Backend State:             Initialize


Supplicants on Gi 1/3:
----------------------
00:01:e9:45:00:03 AUTHENTICATED
00:01:e9:55:00:10 AUTHENTICATING
00:01:e9:B5:00:03 UNAUTHENTICATED
```

Restrict the number of supplicants that can be authenticated on the port in multi-auth mode.	**dot1x max-supplicants** *number* Default: 128	INTERFACE

Note: On the C-Series, during multi-supplicant authentication, devices that fail authentication may still be counted towards the maximum number of supplicants supported by 802.1X authentication to access the port, thus preventing the full number of supplicants to be authenticated.

MAC Authentication Bypass

MAC Authentication Bypass is supported on platforms: C S

MAC Authentication Bypass (MAB) enables you to provide MAC-based security by allowing only known MAC addresses within the network using a RADIUS server.

802.1X-enabled clients can authenticate themselves using the 802.1X protocol. Other devices that do not use 802.1X—like IP phones, printers, and IP fax machines—still need connectivity to the network. The guest VLAN provides one way to access the network. However, placing trusted devices on the quarantined VLAN is not the best practice. MAB allows devices that have known static MAC addresses to be authenticated using their MAC address, and places them into a VLAN different from the VLAN in which unknown devices are placed.

For an 802.1X-incapable device, 802.1X time will out if the device does not respond to the Request Identity frame. If MAB is enabled, the port is then put into learning state and waits indefinitely until the device sends a packet. Once its MAC is learned, it is sent for authentication to the RADIUS server (as both the username and password, in hexadecimal format without any colons). If the server authenticates successfully, the port is dynamically assigned to a MAB VLAN using a RADIUS attribute 81, or is assigned to the untagged VLAN of the port. Afterwards, packets from any other MAC address are dropped. If authentication fails, the authenticator waits the quiet-period and then restarts the authentication process.

MAC authentication bypass works in conjunction and in competition with the guest VLAN and authentication-fail VLAN. When both features are enabled:

1. If authentication fails, the port it is placed into the authentication-fail VLAN.
2. If the host does not respond to the Request Identity frame, the port transitions to MAB initiation state.
3. If MAB times out or MAC authentication fails, the port is placed into the guest VLAN.

If both MAB and re-authentication are enabled, when the re-auth period finishes and whether the previous authentication was through MAB or 802.1X, 802.1X authentication is tried first. If 802.1X times out, MAB authentication is tried. The port remains authorized throughout the reauthentication process. Once a port is enabled/disabled through 802.1X authentication, changes to MAB do not take effect until the MAC is asked to re-authenticate or the port status is toggled.

Note: On the C-Series and S-Series, a MAB-authenticated port becomes unauthorized after an RPM failover.

MAB in Single-host and Multi-Host Mode

In single-host and multi-host mode, the switch attempts to authenticate a supplicant using 802.1X. If 802.1X times out because the supplicant does not respond to the Request Identity frame and MAB is enabled, the switch attempts to authenticate the *first* MAC it learns on the port. Subsequently, for single-host mode, traffic from all other MACs is dropped; for multi-host mode, all traffic from all other MACs is accepted.

After a port is authenticated by MAB, if the switch detects an 802.1X EAPoL start message from the authenticated MAC, the switch re-authenticates using 802.1X first, while keeping the port authorized.

Note: On the C-Series and S-Series, if the switch is in multi-host mode, a MAC address that was MAB-authenticated but later was disabled from MAB authentication, is not denied access but moved to the guest VLAN. If the switch is in single-host mode, the MAC address is disallowed access.

MAB in Multi-Supplicant Authentication Mode

Multi-supplicant authentication (multi-auth) mode is like the other modes in that the switch first attempts to authenticate the supplicant using 802.1X. If 802.1X times out because the supplicant does not respond to the Request Identity frame and MAB authentication is enabled, the switch attempts to authenticate every MAC it learns on the port, up to 128 MACs, which is the maximum number of supplicants 802.1X can authenticate on a single port in multi-authentication mode.

If any supplicant that has been authenticated using MAB starts to speak EAPoL, the switch re-authenticates that supplicant using 802.1X first, while keeping the MAC authorized through the re-authentication process.

Step	Task	Command Syntax	Command Mode
1	Configure the following attributes on the RADIUS Server: • Attribute 1—User-name: Use the supplicant MAC address in hex format without any colons. For example, enter 10:34:AA:33:44:F8 as 1034AA3344F8. • Attribute 2—Password: Use the supplicant MAC address, but encrypted in MD5. • Attribute 4—NAS-IP-Address: IPv4 address of the switch that is used to communicate with the RADIUS server. • Attribute 5—NAS -Port: The port number of the interface being authorized entered as an integer. • Attribute 30—Called-Station-Id: MAC address of the ingress interfaces of the authenticator. • Attribute 31—Calling-Station-Id: MAC address of the 802.1X supplicant. • Attribute 87—NAS-Port-Id: The name of the interface being authorized entered as a string. **Note:** Only attributes 1 and 2 are used for MAB; Attributes 30 and 31 are not mandatory in the MAB method.		
2	Enable MAB.	**dot1x mac-auth-bypass**	INTERFACE

Step	Task	Command Syntax	Command Mode
3	(Optional) Use MAB authentication only—do not use 802.1X authentication first. If MAB fails the port or the MAC address is blocked, the port is placed in the guest VLAN (if configured). 802.1x authentication is not even attempted. Re-authentication is performed using 802.1X timers.	**dot1x auth-type mab-only**	INTERFACE
4	Display the 802.1X and MAB configuration.	**show dot1x interface**	EXEC Privilege

```
FTOS#show dot1x int Gi 2/32

802.1X information on Gi 2/32:
-----------------------------
Dot1x Status:Enable
Port Control:AUTO
Port Auth Status:UNAUTHORIZED
Re-Authentication:Disable
Untagged VLAN id:None
Guest VLAN:Enable
Guest VLAN id:10
Auth-Fail VLAN:Enable
Auth-Fail VLAN id:11
Auth-Fail Max-Attempts:3
Mac-Auth-Bypass:Enable
Tx Period:30 seconds
Quiet Period:60 seconds
ReAuth Max:2
Supplicant Timeout:30 seconds
Server Timeout:30 seconds
Re-Auth Interval:3600 seconds
Max-EAP-Req:2
Auth Type:SINGLE_HOST

Auth PAE State:Initialize
Backend State:Initialize
```

Dynamic CoS with 802.1X

Dynamic CoS with 802.1X is supported on platforms: C S

Class of Service (CoS) is a method of traffic management that groups similar types of traffic so that they are serviced differently. One way of classifying traffic is 802.1p, which uses the 3-bit Priority field in the VLAN tag to mark frames (other classification methods include ToS, ACL, and DSCP). Once traffic is classified, you can use Quality of Service (QoS) traffic management to control the level of service for a class in terms of bandwidth and delivery time.

For incoming traffic, FTOS allows you to set a static priority value on a per-port basis or dynamically set a priority on a per-port basis by leveraging 802.1X.

Note: When priority is statically configured using **dynamic dot1p** and dynamically configured using Dynamic CoS with 802.1X, the dynamic configuration takes precedence.

One use for Dynamic CoS with 802.1X is when the traffic from a server should be classified based on the application that it is running. Static dot1p priority configuration done from the switch is not sufficient in this case, as the server application might change. You would instead need to push the CoS configuration to the switches based on the application the server is running.

Dynamic CoS uses RADIUS attribute 59, called User-Priority-Table, to specify the priority value for incoming frames. Attribute 59 has an 8-octet field that maps the incoming dot1p values to new values; it is essentially a dot1p re-mapping table. The position of each octet corresponds to a priority value: the first octet maps to incoming priority 0, the second octet maps to incoming priority 1, etc. The value in each octet represents the corresponding new priority.

To use the Dynamic CoS with 802.1X authentication, no configuration command is required. You must only configure the supplicant records on the RADIUS server, including VLAN assignment and CoS priority re-mapping table. VLAN and priority values are automatically applied to incoming packets. The RADIUS server finds the appropriate record based on the supplicant's credentials and sends the priority re-mapping table to the Dell Force10 system by including Attribute 59 in the AUTH-ACCEPT packet.

FTOS Behavior: The following conditions are applied to the use of dynamic CoS with 802.1X authentication on C-Series and S-Series platforms:

- In accordance with port-based QoS, incoming dot1p values can be mapped to only four priority values: 0, 2, 4, and 6. If the RADIUS server returns any other dot1p value (1, 3, 5, or 7), the value is not used and frames are forwarded on egress queue 0 without changing the incoming dot1p value. The example shows how dynamic CoS remaps (or does not remap) the dot1p priority in 802.1X-authenticated traffic and how the frames are forwarded:

```
Incoming Frame   RADIUS-based     Outgoing Frame  Egress Queue
Tagged dot1p     CoS Remap Table  Tagged dot1p
--------------   ---------------  --------------  ------------
0                7                0               0
1                5                1               0
2                4                4               2
3                6                6               3
4                3                4               0
5                1                5               0
6                2                2               0
7                4                4               2
```

- The priority of untagged packets is assigned according to the remapped value of priority 0 traffic in the RADIUS-based table. For example, in the following remapping table, untagged packets are tagged with priority 2:

```
FTOS#show dot1x cos-mapping interface Gigabitethernet 2/32

802.1Xp CoS remap table on Gi 2/32:
----------------------------
Dot1p  Remapped Dot1p
0      2
1      6
2      5
3      4
4      3
5      2
6      1
7      0
```

- After being re-tagged by dynamic CoS for 802.1X, packets are forwarded in the switch according to their new CoS priority.
- When a supplicant logs off from an 802.1X authentication session, the dynamic CoS table is deleted or reset.

  When an 802.1x session is re-authenticated, the previously assigned CoS table is retained through the re-authentication process. If the re-authentication fails, the CoS table is deleted. If the re-authentication is successful and the authentication server does not include a CoS table in the AUTH-ACCEPT packet, the previously assigned CoS table MUST be deleted. If the re-authentication is successful and the server sends a CoS table, the old CoS table is overwritten with the new one.
- If multi-supplicant authentication mode is enabled on a port, you can configure a CoS mapping table for specified MAC addresses in the RADIUS server. FTOS will then maintain a per-MAC CoS table for each port, and mark the priority of all traffic originating from a configured MAC address with the corresponding table value.
- To display the CoS priority-mapping table provided by the RADIUS server and applied to authenticated supplicants on an 802.1X-enabled port, enter the **show dot1x cos-mapping interface**
- command.

8

IP Access Control Lists (ACL), Prefix Lists, and Route-maps

IP Access Control Lists, Prefix Lists, and Route-maps are supported on platforms: C E S

Ingress IP ACLs are supported on platforms: C E S

Egress IP ACLs are supported on platform: E

Overview

At their simplest, Access Control Lists (ACLs), Prefix lists, and Route-maps permit or deny traffic based on MAC and/or IP addresses. This chapter discusses implementing IP ACLs, IP Prefix lists and Route-maps. For MAC ACLS, refer to the Access Control Lists (ACLs) chapter in the *FTOS Command Line Reference Guide*.

An ACL is essentially a filter containing some criteria to match (examine IP, TCP, or UDP packets) and an action to take (permit or deny). ACLs are processed in sequence so that if a packet does not match the criterion in the first filter, the second filter (if configured) is applied. When a packet matches a filter, the switch drops or forwards the packet based on the filter's specified action. If the packet does not match any of the filters in the ACL, the packet is dropped (implicit deny).

The number of ACLs supported on a system depends on your CAM size. See CAM Profiling, CAM Allocation, and CAM Optimization in this chapter for more information. Refer to Chapter 11, Content Addressable Memory, on page 281 for complete CAM profiling information.

This chapter covers the following topics:

IP Access Control Lists (ACLs)

In the Dell Force10 switch/routers, you can create two different types of IP ACLs: standard or extended. A standard ACL filters packets based on the source IP packet. An extended ACL filters traffic based on the following criteria (for more information on ACL supported options see the *FTOS Command Reference*):

- IP protocol number
- Source IP address
- Destination IP address
- Source TCP port number
- Destination TCP port number
- Source UDP port number
- Destination UDP port number

For extended ACL TCP and UDP filters, you can match criteria on specific or ranges of TCP or UDP ports. For extended ACL TCP filters, you can also match criteria on established TCP sessions.

When creating an access list, the sequence of the filters is important. You have a choice of assigning sequence numbers to the filters as you enter them, or FTOS will assign numbers in the order the filters are created. The sequence numbers, whether configured or assigned by FTOS, are listed in the **show config** and **show ip accounting access-list** command display output.

Ingress and egress Hot Lock ACLs allow you to append or delete new rules into an existing ACL (already written into CAM) without disrupting traffic flow. Existing entries in CAM are shuffled to accommodate the new entries. Hot Lock ACLs are enabled by default and support both standard and extended ACLs on all platforms.

Note: Hot Lock ACLs are supported on Ingress ACLs only.

CAM Profiling, CAM Allocation, and CAM Optimization

CAM Profiling is supported on platform E

User Configurable CAM Allocations are supported on platform C

CAM optimization is supported on platforms C S

CAM Profiling

CAM optimization is supported on platforms E T

CAM profiling for ACLs is supported on E-Series TeraScale only. For complete information regarding E-Series TeraScale CAM profiles and configuration, refer to Chapter 11, Content Addressable Memory.

The default CAM profile has 1K Layer 2 ingress ACL entries. If you need more memory for Layer 2 ingress ACLs, select the profile *l2-ipv4-inacl*.

When budgeting your CAM allocations for ACLs and QoS configurations, remember that ACL and QoS rules might consume more than one CAM entry depending on complexity. For example, TCP and UDP rules with port range options might require more than one CAM entry.

The Layer 2 ACL CAM partition has sub-partitions for several types of information. Table 8-1 lists the sub-partition and the percentage of the Layer 2 ACL CAM partition that FTOS allocates to each by default.

Table 8-1. Layer 2 ACL CAM Sub-partition Sizes

Partition	% Allocated
Sysflow	6
L2ACL	14
*PVST	50
QoS	12
L2PT	13
FRRP	5

You can re-configure the amount of space, in percentage, allocated to each sub-partition. As with the IPv4Flow partition, you can configure the Layer 2 ACL partition from EXEC Privilege mode or CONFIGURATION mode.

The amount of space that you can distribute to the sub-partitions is equal to the amount of CAM space that the selected CAM profile allocates to the Layer 2 ACL partition. FTOS requires that you specify the amount of CAM space for all sub-partitions and that the sum of all sub-partitions is 100%. FTOS displays the following message if the total allocated space is not correct:

```
% Error: Sum of all regions does not total to 100%.
```

User Configurable CAM Allocation

User Configurable CAM Allocations are supported on platform C

Allocate space for IPV6 ACLs on the C-Series by using the **cam-acl** command in CONFIGURATION mode.

The CAM space is allotted in FP blocks. The total space allocated must equal 13 FP blocks. Note that there are 16 FP blocks, but the System Flow requires 3 blocks that cannot be reallocated. The default CAM Allocation settings on a C-Series matching are:

- L3 ACL (ipv4acl): 6
- L2 ACL(l2acl) : 5
- IPv6 L3 ACL (ipv6acl): 0
- L3 QoS (ipv4qos): 1
- L2 QoS (l2qos): 1

The **ipv6acl** allocation must be entered as a factor of 2 (2, 4, 6, 8, 10). All other profile allocations can use either even or odd numbered ranges.

You must save the new CAM settings to the startup-config (**write-mem or copy run start**) then reload the system for the new settings to take effect.

CAM optimization

CAM optimization is supported on platforms C S

When this command is enabled, if a Policy Map containing classification rules (ACL and/or dscp/ ip-precedence rules) is applied to more than one physical interface on the same port-pipe, only a single copy of the policy is written (only 1 FP entry will be used). When the command is disabled, the system behaves as described in this chapter.

Test CAM Usage

The test cam-usage command is supported on platforms C E S

This command applies to both IPv4 and IPv6 CAM profiles, but is best used when verifying QoS optimization for IPv6 ACLs.

Use this command to determine whether sufficient ACL CAM space is available to enable a service-policy. Create a Class Map with all required ACL rules, then execute the **test cam-usage** command in Privilege mode to verify the actual CAM space required. Figure 8-1 gives a sample of the output shown when executing the command. The status column indicates whether or not the policy can be enabled.

Figure 8-1. Command Example: test cam-usage (C-Series)

```
FTOS#test cam-usage service-policy input TestPolicy linecard all

Linecard | Portpipe | CAM Partition | Available CAM | Estimated CAM per Port | Status
--------------------------------------------------------------------------------------
       2 |        1 | IPv4Flow      |           232 |                      0 | Allowed
       2 |        1 | IPv6Flow      |             0 |                      0 | Allowed
       4 |        0 | IPv4Flow      |           232 |                      0 | Allowed
       4 |        0 | IPv6Flow      |             0 |                      0 | Allowed
FTOS#
```

Implementing ACLs on FTOS

One IP ACL can be assigned per interface with FTOS. If an IP ACL is not assigned to an interface, it is not used by the software in any other capacity.

The number of entries allowed per ACL is hardware-dependent. Refer to your line card documentation for detailed specification on entries allowed per ACL.

If counters are enabled on IP ACL rules that are already configured, those counters are reset when a new rule is inserted or prepended. If a rule is appended, the existing counters are not affected. This is applicable to the following features:

- L2 Ingress Access list
- L2 Egress Access list
- L3 Egress Access list

Note: IP ACLs are supported over VLANs in Version 6.2.1.1 and higher.

ACLs and VLANs

There are some differences when assigning ACLs to a VLAN rather than a physical port. For example, when using a single port-pipe, if you apply an ACL to a VLAN, one copy of the ACL entries would get installed in the ACL CAM on the port-pipe. The entry would look for the incoming VLAN in the packet. Whereas if you apply an ACL on individual ports of a VLAN, separate copies of the ACL entries would be installed for each port belonging to a port-pipe.

When you use the **log** keyword, CP processor will have to log details about the packets that match. Depending on how many packets match the log entry and at what rate, CP might become busy as it has to log these packets' details. However the other processors (RP1 and RP2) should be unaffected. This option is typically useful when debugging some problem related to control traffic. We have used this option numerous times in the field and have not encountered any problems in such usage so far.

ACL Optimization

If an access list contains duplicate entries, FTOS deletes one entry to conserve CAM space.

Standard and Extended ACLs take up the same amount of CAM space. A single ACL rule uses 2 CAM entries whether it is identified as a Standard or Extended ACL.

Determine the order in which ACLs are used to classify traffic

When you link class-maps to queues using the command **service-queue**, FTOS matches the class-maps according to queue priority (queue numbers closer to 0 have lower priorities). For example, in Figure 8-2, class-map *cmap2* is matched against ingress packets before *cmap1*.

ACLs *acl1* and *acl2* have overlapping rules because the address range 20.1.1.0/24 is within 20.0.0.0/8. Therefore, (without the keyword **order**) packets within the range 20.1.1.0/24 match positive against *cmap1* and are buffered in queue 7, though you intended for these packets to match positive against *cmap2* and be buffered in queue 4.

In cases such as these, where class-maps with overlapping ACL rules are applied to different queues, use the **order** keyword to specify the order in which you want to apply ACL rules, as shown in Figure 8-2. The order can range from 0 to 254. FTOS writes to the CAM ACL rules with lower order numbers (order numbers closer to 0) before rules with higher order numbers so that packets are matched as you intended. By default, all ACL rules have an order of 254.

Figure 8-2. Using the Order Keyword in ACLs

```
FTOS(conf)#ip access-list standard acl1
FTOS(config-std-nacl)#permit 20.0.0.0/8
FTOS(config-std-nacl)#exit
FTOS(conf)#ip access-list standard acl2
FTOS(config-std-nacl)#permit 20.1.1.0/24 order 0
FTOS(config-std-nacl)#exit
FTOS(conf)#class-map match-all cmap1
FTOS(conf-class-map)#match ip access-group acl1
FTOS(conf-class-map)#exit
FTOS(conf)#class-map match-all cmap2
FTOS(conf-class-map)#match ip access-group acl2
FTOS(conf-class-map)#exit
FTOS(conf)#policy-map-input pmap
FTOS(conf-policy-map-in)#service-queue 7 class-map cmap1
FTOS(conf-policy-map-in)#service-queue 4 class-map cmap2
FTOS(conf-policy-map-in)#exit
FTOS(conf)#interface gig 1/0
FTOS(conf-if-gi-1/0)#service-policy input pmap
```

IP Fragment Handling

FTOS supports a configurable option to explicitly deny IP fragmented packets, particularly second and subsequent packets. It extends the existing ACL command syntax with the **fragments** keyword for all Layer 3 rules applicable to all Layer protocols (permit/deny ip/tcp/udp/icmp).

- Both standard and extended ACLs support IP fragments.

- Second and subsequent fragments are allowed because a Layer 4 rule cannot be applied to these fragments. If the packet is to be denied eventually, the first fragment would be denied and hence the packet as a whole cannot be reassembled.
- Implementing the required rules will use a significant number of CAM entries per TCP/UDP entry.
- For IP ACL, FTOS always applies implicit deny. You do not have to configure it.
- For IP ACL, FTOS applies implicit permit for second and subsequent fragment just prior to the implicit deny.
- If an *explicit* deny is configured, the second and subsequent fragments will not hit the implicit permit rule for fragments.
- Loopback interfaces do not support ACLs using the IP fragment option. If you configure an ACL with the fragments option and apply it to a loopback interface, the command is accepted, but the ACL entries are not actually installed the offending rule in CAM.

IP fragments ACL examples

The following configuration permits all packets (both fragmented & non-fragmented) with destination IP 10.1.1.1. The second rule does not get hit at all.

```
FTOS(conf)#ip access-list extended ABC
FTOS(conf-ext-nacl)#permit ip any 10.1.1.1/32
FTOS(conf-ext-nacl)#deny ip any 10.1.1.1./32 fragments
FTOS(conf-ext-nacl)
```

To deny second/subsequent fragments, use the same rules in a different order. These ACLs deny all second & subsequent fragments with destination IP 10.1.1.1 but permit the first fragment & non fragmented packets with destination IP 10.1.1.1 .

```
FTOS(conf)#ip access-list extended ABC
FTOS(conf-ext-nacl)#deny ip any 10.1.1.1/32 fragments
FTOS(conf-ext-nacl)#permit ip any 10.1.1.1/32
FTOS(conf-ext-nacl)
```

Layer 4 ACL rules examples

In the below scenario, first fragments non-fragmented TCP packets from 10.1.1.1 with TCP destination port equal to 24 are permitted. All other fragments are denied.

```
FTOS(conf)#ip access-list extended ABC
FTOS(conf-ext-nacl)#permit tcp host 10.1.1.1 any eq 24
FTOS(conf-ext-nacl)#deny ip any any fragment
FTOS(conf-ext-nacl)
```

In the following, TCP packets that are first fragments or non-fragmented from host 10.1.1.1 with TCP destination port equal to 24 are permitted. Additionally, all TCP non-first fragments from host 10.1.1.1 are permitted. All other IP packets that are non-first fragments are denied.

```
FTOS(conf)#ip access-list extended ABC
FTOS(conf-ext-nacl)#permit tcp host 10.1.1.1 any eq 24
FTOS(conf-ext-nacl)#permit tcp host 10.1.1.1 any fragment
FTOS(conf-ext-nacl)#deny ip any any fragment
FTOS(conf-ext-nacl)
```

To log all the packets denied and to override the implicit deny rule and the implicit permit rule for TCP/UDP fragments, use a configuration similar to the following.

```
FTOS(conf)#ip access-list extended ABC
FTOS(conf-ext-nacl)#permit tcp any any fragment
FTOS(conf-ext-nacl)#permit udp any any fragment
FTOS(conf-ext-nacl)#deny ip any any log
FTOS(conf-ext-nacl)
```

Note the following when configuring ACLs with the **fragments** keyword.

When an ACL filters packets it looks at the Fragment Offset (FO) to determine whether or not it is a fragment.

FO = 0 means it is either the first fragment or the packet is a non-fragment.

FO > 0 means it is dealing with the fragments of the original packet.

Permit ACL line with L3 information only, and the fragments keyword is present:
If a packet's L3 information matches the L3 information in the ACL line, the packet's fragment offset (FO) is checked.

- If a packet's FO > 0, the packet is permitted.
- If a packet's FO = 0 , the next ACL entry is processed.

Deny ACL line with L3 information only, and the fragments keyword is present:
If a packet's L3 information does match the L3 information in the ACL line, the packet's fragment offset (FO) is checked.

- If a packet's FO > 0, the packet is denied.
- If a packet's FO = 0, the next ACL line is processed.

Configure a standard IP ACL

To configure an ACL, use commands in the IP ACCESS LIST mode and the INTERFACE mode. The following list includes the configuration tasks for IP ACLs:

For a complete listing of all commands related to IP ACLs, refer to the *FTOS Command Line Interface Reference* document.

Refer to Configure an extended IP ACL on page 143 to set up extended ACLs.

A standard IP ACL uses the source IP address as its match criterion.

Note: On E-Series ExaScale systems, TCP ACL flags are not supported in standard or extended ACLs with IPv6 microcode. An error message is shown if IPv6 microcode is configured and an ACL is entered with a TCP filter included.

```
FTOS(conf-ipv6-acl)#seq 8 permit tcp any any urg
May 5 08:32:34: %E90MJ:0 %ACL_AGENT-2-ACL_AGENT_ENTRY_ERROR: Unable to write seq 8 of
list test as individual TCP flags are not supported on linecard 0
```

To configure a standard IP ACL, use these commands in the following sequence:

Step	Command Syntax	Command Mode	Purpose
1	**ip access-list standard** *access-listname*	CONFIGURATION	Enter IP ACCESS LIST mode by naming a standard IP access list.
2	**seq** *sequence-number* {**deny** | **permit**} {*source* [*mask*] | **any | host** *ip-address*} [**count** [**byte**] | **log**] [**order**] [**monitor**] [**fragments**]	CONFIG-STD-NACL	Configure a drop or forward filter. The parameters are: • **log** and **monitor** options are supported on E-Series only.

Note: When assigning sequence numbers to filters, keep in mind that you might need to insert a new filter. To prevent reconfiguring multiple filters, assign sequence numbers in multiples of five or another number.

When you use the **log** keyword, CP processor logs details about the packets that match. Depending on how many packets match the log entry and at what rate, the CP may become busy as it has to log these packets' details.

To view the rules of a particular ACL configured on a particular interface, use the **show ip accounting access-list** *ACL-name* **interface** *interface* command (Figure 226) in EXEC Privilege mode.

Figure 8-3. Command Example: show ip accounting access-list

```
FTOS#show ip accounting access ToOspf interface gig 1/6
Standard IP access list ToOspf
 seq 5 deny any
 seq 10 deny 10.2.0.0 /16
 seq 15 deny 10.3.0.0 /16
 seq 20 deny 10.4.0.0 /16
 seq 25 deny 10.5.0.0 /16
 seq 30 deny 10.6.0.0 /16
 seq 35 deny 10.7.0.0 /16
 seq 40 deny 10.8.0.0 /16
 seq 45 deny 10.9.0.0 /16
 seq 50 deny 10.10.0.0 /16
FTOS#
```

Figure 8-4 illustrates how the **seq** command orders the filters according to the sequence number assigned. In the example, filter 25 was configured before filter 15, but the **show config** command displays the filters in the correct order.

Figure 8-4. Command example: seq

```
FTOS(config-std-nacl)#seq 25 deny ip host 10.5.0.0 any log
FTOS(config-std-nacl)#seq 15 permit tcp 10.3.0.0 /16 any
FTOS(config-std-nacl)#show config
!
ip access-list standard dilling
 seq 15 permit tcp 10.3.0.0/16 any
 seq 25 deny ip host 10.5.0.0 any log
FTOS(config-std-nacl)#
```

To delete a filter, use the **no seq** *sequence-number* command in the IP ACCESS LIST mode.

If you are creating a standard ACL with only one or two filters, you can let FTOS assign a sequence number based on the order in which the filters are configured. The software assigns filters in multiples of 5.

To configure a filter without a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip access-list standard** *access-list-name*	CONFIGURATION	Create a standard IP ACL and assign it a unique name.
2	{**deny** | **permit**} {*source* [*mask*] | **any** | **host** *ip-address*} [**count** [**byte**] | **log**] [**order**] [**monitor**] [**fragments**]	CONFIG-STD-NACL	Configure a drop or forward IP ACL filter. • **log** and **monitor** options are supported on E-Series only.

When you use the **log** keyword, CP processor logs details about the packets that match. Depending on how many packets match the log entry and at what rate, the CP may become busy as it has to log these packets' details.

Figure 8-5 illustrates a standard IP ACL in which the sequence numbers were assigned by the FTOS. The filters were assigned sequence numbers based on the order in which they were configured (for example, the first filter was given the lowest sequence number). The **show config** command in the IP ACCESS LIST mode displays the two filters with the sequence numbers 5 and 10.

Figure 8-5. Standard IP ACL

```
FTOS(config-route-map)#ip access standard kigali
FTOS(config-std-nacl)#permit 10.1.0.0/16
FTOS(config-std-nacl)#show config
!
ip access-list standard kigali
 seq 5 permit 10.1.0.0/16
FTOS(config-std-nacl)#
```

To view all configured IP ACLs, use the **show ip accounting access-list** command (Figure 229) in the EXEC Privilege mode.

Figure 8-6. Command Example: show ip accounting access-list

```
FTOS#show ip accounting access example interface gig 4/12
Extended IP access list example
seq 10 deny tcp any any eq 111
 seq 15 deny udp any any eq 111
 seq 20 deny udp any any eq 2049
 seq 25 deny udp any any eq 31337
 seq 30 deny tcp any any range 12345 12346
 seq 35 permit udp host 10.21.126.225 10.4.5.0 /28
 seq 40 permit udp host 10.21.126.226 10.4.5.0 /28
 seq 45 permit udp 10.8.0.0 /16 10.50.188.118 /31 range 1812 1813
 seq 50 permit tcp 10.8.0.0 /16 10.50.188.118 /31 eq 49
 seq 55 permit udp 10.15.1.0 /24 10.50.188.118 /31 range 1812 1813
```

To delete a filter, enter the **show config** command in the IP ACCESS LIST mode and locate the sequence number of the filter you want to delete. Then use the **no seq** *sequence-number* command in the IP ACCESS LIST mode.

Configure an extended IP ACL

Extended IP ACLs filter on source and destination IP addresses, IP host addresses, TCP addresses, TCP host addresses, UDP addresses, and UDP host addresses.

Since traffic passes through the filter in the order of the filter's sequence, you can configure the extended IP ACL by first entering the IP ACCESS LIST mode and then assigning a sequence number to the filter.

Note: On E-Series ExaScale systems, TCP ACL flags are not supported in standard or extended ACLs with IPv6 microcode. An error message is shown if IPv6 microcode is configured and an ACL is entered with a TCP filter included.

```
FTOS(conf-ipv6-acl)#seq 8 permit tcp any any urg
May 5 08:32:34: %E90MJ-0-%ACL_AGENT-2-ACL_AGENT_ENTRY_ERROR: Unable to write seq 8 of
list test as individual TCP flags are not supported on linecard 0
```

Configure filters with sequence number

To create a filter for packets with a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip access-list extended** *access-list-name*	CONFIGURATION	Enter the IP ACCESS LIST mode by creating an extended IP ACL.

Step	Command Syntax	Command Mode	Purpose
2	**seq** *sequence-number* {**deny** | **permit**} {*ip-protocol-number* | **icmp | ip | tcp | udp**} {*source mask* | **any** | **host** *ip-address*} {*destination mask* | **any** | **host** *ip-address*} [*operator port* [*port*]] [**count** [**byte**] | **log**] [**order**] [**monitor**] [**fragments**]	CONFIG-EXT-NACL	Configure a drop or forward filter. • **log** and **monitor** options are supported on E-Series only.

When you use the **log** keyword, CP processor logs details about the packets that match. Depending on how many packets match the log entry and at what rate, the CP may become busy as it has to log these packets' details.

TCP packets: To create a filter for TCP packets with a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip access-list extended** *access-list-name*	CONFIGURATION	Create an extended IP ACL and assign it a unique name.
2	**seq** *sequence-number* {**deny** | **permit**} **tcp** {*source mask* | **any** | **host** *ip-address*}} [**count** [**byte**] | **log**] [**order**] [**monitor]** [**fragments**]	CONFIG-EXT-NACL	Configure an extended IP ACL filter for TCP packets. • **log** and **monitor** options are supported on E-Series only.

When you use the **log** keyword, CP processor logs details about the packets that match. Depending on how many packets match the log entry and at what rate, the CP may become busy as it has to log these packets' details.

UDP packets: To create a filter for UDP packets with a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip access-list extended** *access-list-name*	CONFIGURATION	Create a extended IP ACL and assign it a unique name.
2	**seq** *sequence-number* {**deny** | **permit**} {*ip-protocol-number* **udp**} {*source mask* | **any** | **host** *ip-address*} {*destination mask* | **any** | **host** *ip-address*} [*operator port* [*port*]] [**count** [**byte**] | **log**] [**order**] [**monitor**] [**fragments**]	CONFIG-EXT-NACL	Configure an extended IP ACL filter for UDP packets. • **log** and **monitor** options are supported on E-Series only.

When you create the filters with a specific sequence number, you can create the filters in any order and the filters are placed in the correct order.

Note: When assigning sequence numbers to filters, keep in mind that you might need to insert a new filter. To prevent reconfiguring multiple filters, assign sequence numbers in multiples of five or another number.

Figure 8-7 illustrates how the **seq** command orders the filters according to the sequence number assigned. In the example, filter 15 was configured before filter 5, but the **show config** command displays the filters in the correct order.

Figure 8-7. Command Example: seq

```
FTOS(config-ext-nacl)#seq 15 deny ip host 112.45.0.0 any log
FTOS(config-ext-nacl)#seq 5 permit tcp 12.1.3.45 0.0.255.255 any
FTOS(config-ext-nacl)#show confi
!
ip access-list extended dilling
 seq 5 permit tcp 12.1.0.0 0.0.255.255 any
 seq 15 deny ip host 112.45.0.0 any log
FTOS(config-ext-nacl)#
```

Configure filters without sequence number

If you are creating an extended ACL with only one or two filters, you can let FTOS assign a sequence number based on the order in which the filters are configured. FTOS assigns filters in multiples of 5.

To configure a filter for an extended IP ACL without a specified sequence number, use any or all of the following commands in the IP ACCESS LIST mode:

Command Syntax	Command Mode	Purpose
{**deny** | **permit**} {*source mask* | **any** | **host** *ip-address*} [**count** [**byte**] | **log**] [**order**] [**monitor**] [**fragments**]	CONFIG-EXT-NACL	Configure a deny or permit filter to examine IP packets. • **log** and **monitor** options are supported on E-Series only.
{**deny** | **permit**} **tcp** {*source mask*] | **any** | **host** *ip-address*}} [**count** [**byte**] | **log**] [**order**] [**monitor**] [**fragments**]	CONFIG-EXT-NACL	Configure a deny or permit filter to examine TCP packets. • **log** and **monitor** options are supported on E-Series only.
{**deny** | **permit**} **udp** {*source mask* | **any** | **host** *ip-address*}} [**count** [**byte**] | **log**] [**order**] [**monitor**] [**fragments**]	CONFIG-EXT-NACL	Configure a deny or permit filter to examine UDP packets. • **log** and **monitor** options are supported on E-Series only.

When you use the **log** keyword, CP processor logs details about the packets that match. Depending on how many packets match the log entry and at what rate, the CP may become busy as it has to log these packets' details.

Figure 8-8 illustrates an extended IP ACL in which the sequence numbers were assigned by the software. The filters were assigned sequence numbers based on the order in which they were configured (for example, the first filter was given the lowest sequence number). The **show config** command in the IP ACCESS LIST mode displays the two filters with the sequence numbers 5 and 10.

Figure 8-8. Extended IP ACL

```
FTOS(config-ext-nacl)#deny tcp host 123.55.34.0 any
FTOS(config-ext-nacl)#permit udp 154.44.123.34 0.0.255.255 host 34.6.0.0
FTOS(config-ext-nacl)#show config
!
ip access-list extended nimule
 seq 5 deny tcp host 123.55.34.0 any
 seq 10 permit udp 154.44.0.0 0.0.255.255 host 34.6.0.0
FTOS(config-ext-nacl)#
```

To view all configured IP ACLs and the number of packets processed through the ACL, use the **show ip accounting access-list** command (Figure 232) in the EXEC Privilege mode.

Established Flag

The **est** (established) flag is deprecated for Terascale series line cards.The flag is only available on legacy Etherscale linecards. Employ the **ack** and **rst** flags in their stead to achieve the same functionality.

To obtain the functionality of **est,** use the following ACLs:

- permit tcp any any rst
- permit tcp any any ack

Configuring Layer 2 and Layer 3 ACLs on an Interface

Both Layer 2 and Layer 3 ACLs may be configured on an interface in Layer 2 mode. If both L2 and L3 ACLs are applied to an interface, the following rules apply:

- The packets routed by FTOS are governed by the L3 ACL only, since they are not filtered against an L2 ACL.
- The packets switched by FTOS are first filtered by the L3 ACL, then by the L2 ACL.
- When packets are switched by FTOS, the egress L3 ACL does not filter the packet.

For the following features, if counters are enabled on rules that have already been configured and a new rule is either inserted or prepended, all the existing counters will be reset:

- L2 Ingress Access list
- L3 Egress Access list
- L2 Egress Access list

If a rule is simply appended, existing counters are not affected.

Table 8-2. L2 and L3 ACL Filtering on Switched Packets

L2 ACL Behavior	L3 ACL Behavior	Decision on Targeted Traffic
Deny	Deny	Denied by L3 ACL
Deny	Permit	Permitted by L3 ACL
Permit	Deny	Denied by L2 ACL
Permit	Permit	Permitted by L2 ACL

Note: If an interface is configured as a "**vlan-stack access**" port, the packets are filtered by an L2 ACL only. The L3 ACL applied to such a port does not affect traffic. That is, existing rules for other features (such as trace-list, PBR, and QoS) are applied accordingly to the permitted traffic.

For information on MAC ACLs, refer to the Access Control Lists (ACLs) chapter in the *FTOS Command Line Reference Guide*.

Assign an IP ACL to an Interface

Ingress IP ACLs are supported on platforms: C and S

Ingress and Egress IP ACL are supported on platform: E

To pass traffic through a configured IP ACL, you must assign that ACL to a physical interface, a port channel interface, or a VLAN. The IP ACL is applied to all traffic entering a physical or port channel interface and the traffic is either forwarded or dropped depending on the criteria and actions specified in the ACL.

The same ACL may be applied to different interfaces and that changes its functionality. For example, you can take ACL "ABCD", and apply it using the **in** keyword and it becomes an ingress access list. If you apply the same ACL using the **out** keyword, it becomes an egress access list. If you apply the same ACL to the loopback interface, it becomes a loopback access list.

This chapter covers the following topics:

For more information on Layer-3 interfaces, refer to Chapter 13, Interfaces, on page 47.

To apply an IP ACL (standard or extended) to a physical or port channel interface, use these commands in the following sequence in the INTERFACE mode:

Step	Command Syntax	Command Mode	Purpose
1	**interface interface slot/port**	CONFIGURATION	Enter the interface number.
2	**ip address** *ip-address*	INTERFACE	Configure an IP address for the interface, placing it in Layer-3 mode.
3	**ip access-group** *access-list-name* {**in | out**} [**implicit-permit**] [**vlan** *vlan-range*]	INTERFACE	Apply an IP ACL to traffic entering or exiting an interface. • **out:** configure the ACL to filter outgoing traffic. This keyword is supported only on E-Series. **Note:** The number of entries allowed per ACL is hardware-dependent. Refer to your line card documentation for detailed specification on entries allowed per ACL.
4	**ip access-list [standard | extended]** *name*	INTERFACE	Apply rules to the new ACL.

To view which IP ACL is applied to an interface, use the **show config** command (Figure 232) in the INTERFACE mode or the **show running-config** command in the EXEC mode.

Figure 8-9. Command example: show config in the INTERFACE Mode

```
FTOS(conf-if)#show conf
!
interface GigabitEthernet 0/0
 ip address 10.2.1.100 255.255.255.0
 ip access-group nimule in
 no shutdown
FTOS(conf-if)#
```

Use only Standard ACLs in the **access-class** command to filter traffic on Telnet sessions.

Counting ACL Hits

You can view the number of packets matching the ACL by using the **count** option when creating ACL entries. E-Series supports packet and byte counts simultaneously. C-Series and S-Series support only one at any given time.

To view the number of packets matching an ACL that is applied to an interface:

Step	Task
1	Create an ACL that uses rules with the count option. See Configure a standard IP ACL on page 140
2	Apply the ACL as an inbound or outbound ACL on an interface. See Assign an IP ACL to an Interface on page 147

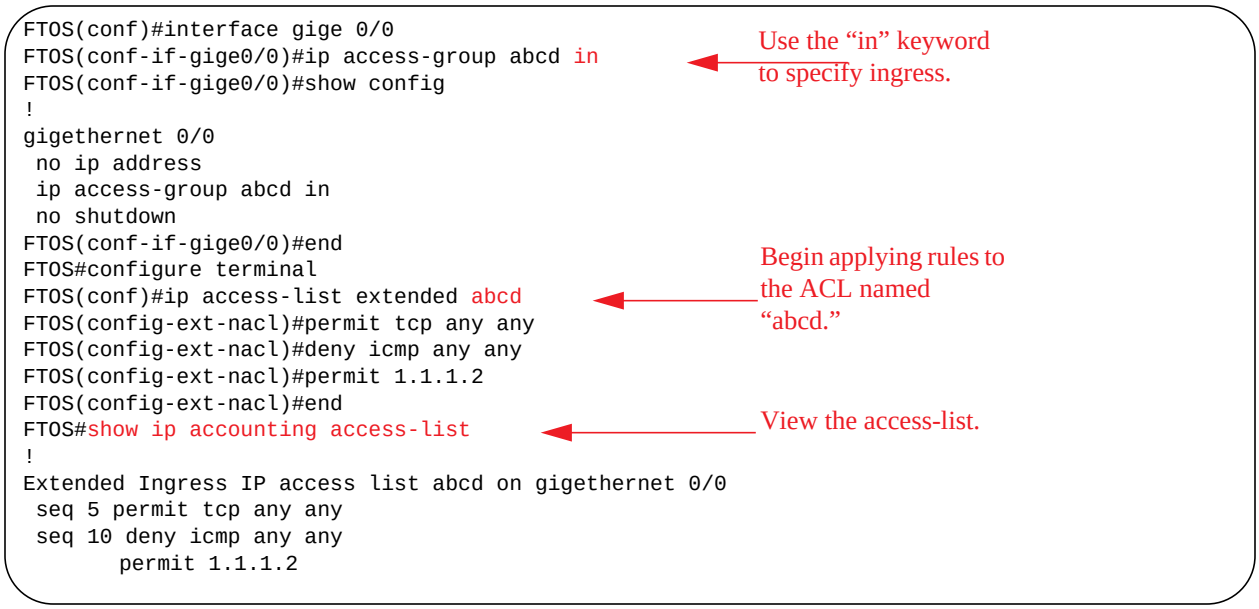
Step	Task
3	View the number of packets matching the ACL using the **show ip accounting access-list** from EXEC Privilege mode.

Configuring Ingress ACLs

Ingress ACLs are applied to interfaces and to traffic entering the system.These system-wide ACLs eliminate the need to apply ACLs onto each interface and achieves the same results. By localizing target traffic, it is a simpler implementation.

To create an ingress ACLs, use the **ip access-group** command (Figure 233) in the EXEC Privilege mode. This example also shows applying the ACL, applying rules to the newly created access group, and viewing the access list:

Figure 8-10. Creating an Ingress ACL



Configuring Egress ACLs

Layer 2 and Layer 3 ACLs are supported on platform E

Egress ACLs are applied to line cards and affect the traffic leaving the system. Configuring egress ACLs onto physical interfaces protects the system infrastructure from attack—malicious and incidental—by explicitly allowing only authorized traffic.These system-wide ACLs eliminate the need to apply ACLs onto each interface and achieves the same results. By localizing target traffic, it is a simpler implementation.

An egress ACL is used when users would like to restrict egress traffic. For example, when a DOS attack traffic is isolated to one particular interface, you can apply an egress ACL to block that particular flow from exiting the box, thereby protecting downstream devices.

To create an egress ACLs, use the **ip access-group** command (Figure 234) in the EXEC Privilege mode. This example also shows viewing the configuration, applying rules to the newly created access group, and viewing the access list:

Figure 8-11. Creating an Egress ACL

```
FTOS(conf)#interface gige 0/0
FTOS(conf-if-gige0/0)#ip access-group abcd out
FTOS(conf-if-gige0/0)#show config
!
gigethernet 0/0
 no ip address
 ip access-group abcd out
 no shutdown
FTOS(conf-if-gige0/0)#end
FTOS#configure terminal
FTOS(conf)#ip access-list extended abcd
FTOS(config-ext-nacl)#permit tcp any any
FTOS(config-ext-nacl)#deny icmp any any
FTOS(config-ext-nacl)#permit 1.1.1.2
FTOS(config-ext-nacl)#end
FTOS#show ip accounting access-list
!
Extended Ingress IP access list abcd on gigethernet 0/0
 seq 5 permit tcp any any
 seq 10 deny icmp any any
        permit 1.1.1.2
```

Use the "out" keyword to specify egress.

Begin applying rules to the ACL named "abcd."

View the access-list.

Egress Layer 3 ACL Lookup for Control-plane IP Traffic

By default, packets originated from the system are not filtered by egress ACLs. If you initiate a ping session from the system, for example, and apply an egress ACL to block this type of traffic on the interface, the ACL does not affect that ping traffic. The Control Plane Egress Layer 3 ACL feature enhances IP reachability debugging by implementing control-plane ACLs for CPU-generated and CPU-forwarded traffic. Using **permit** rules with the **count** option, you can track on a per-flow basis whether CPU-generated and CPU-forwarded packets were transmitted successfully..

Task	Command Syntax	Command Mode
Apply Egress ACLs to IPv4 system traffic.	**ip control-plane** [**egress filter**]	CONFIGURATION
Apply Egress ACLs to IPv6 system traffic.	**ipv6 control-plane** [**egress filter**]	CONFIGURATION
Create a Layer 3 ACL using **permit** rules with the **count** option to describe the desired CPU traffic	**permit ip** {*source mask* | **any** | **host** *ip-address}* {*destination mask* | **any** | **host** *ip-address*} **count**	CONFIG-NACL

FTOS Behavior: VRRP hellos and IGMP packets are not affected when egress ACL filtering for CPU traffic is enabled. Packets sent by the CPU with the source address as the VRRP virtual IP address have the interface MAC address instead of VRRP virtual MAC address.

Configuring ACLs to Loopback

ACLs can be supplied on Loopback interfaces supported on platform E

Configuring ACLs onto the CPU in a loopback interface protects the system infrastructure from attack—malicious and incidental—by explicate allowing only authorized traffic.

The ACLs on loopback interfaces are applied only to the CPU on the RPM—this eliminates the need to apply specific ACLs onto all ingress interfaces and achieves the same results. By localizing target traffic, it is a simpler implementation.

The ACLs target and handle Layer 3 traffic destined to terminate on the system including routing protocols, remote access, SNMP, ICMP, and etc. Effective filtering of Layer 3 traffic from Layer 3 routers reduces the risk of attack.

Note: Loopback ACLs are supported only on ingress traffic.

Loopback interfaces do not support ACLs using the IP fragment option. If you configure an ACL with the fragments option and apply it to a loopback interface, the command is accepted, but the ACL entries are not actually installed the offending rule in CAM.

See also Loopback Interfaces in the Interfaces chapter.

Applying an ACL on Loopback Interfaces

ACLs can be applied on Loopback interfaces supported on platform E

To apply an ACL (standard or extended) for loopback, use these commands in the following sequence:

Step	Command Syntax	Command Mode	Purpose
1	**interface loopback 0**	CONFIGURATION	Only loopback 0 is supported for the loopback ACL.

Step	Command Syntax	Command Mode	Purpose
2	[**seq** *number*] **permit loopback-logging any any**	CONFIGURATION	If you are applying an extended ACL, and it has a *deny ip any any* entry, this entry denies internally generated packets as well as packets received from external devices. To prevent internally generated packets from being dropped, make sure that the ACL you intend to apply has the following entry: [**seq** *number*] **permit loopback-logging any any**. This line may be anywhere in the ACL.
3	**ip access-list** [**standard** | **extended**] *name*	CONFIGURATION	Apply rules to the new ACL.
4	**ip access-group** *name* **in**	INTERFACE	Apply an ACL to traffic entering loopback. • **in:** configure the ACL to filter incoming traffic **Note:** ACLs for loopback can only be applied to incoming traffic.

To apply ACLs on loopback, use the **ip access-group** command (Figure 235) in the INTERFACE mode. This example also shows the interface configuration status, adding rules to the access group, and displaying the list of rules in the ACL:

Figure 8-12. Applying an ACL to the Loopback Interface

```
FTOS(conf)#interface loopback 0
FTOS(conf-if-lo-0)#ip access-group abcd in
FTOS(conf-if-lo-0)#show config
!
interface Loopback 0
 no ip address
 ip access-group abcd in
 no shutdown
FTOS(conf-if-lo-0)#end
FTOS#configure terminal
FTOS(conf)#ip access-list extended abcd
FTOS(config-ext-nacl)#permit tcp any any
FTOS(config-ext-nacl)#deny icmp any any
FTOS(config-ext-nacl)#permit 1.1.1.2
FTOS(config-ext-nacl)#end
FTOS#show ip accounting access-list
!
Extended Ingress IP access list abcd on Loopback 0
        seq 5 permit tcp any any
        seq 10 deny icmp any any
        seq 10 deny icmp any any
```

Use the **in** keyword.

Add rules to the ACL named “abcd.”

Display the ACL.

Note: See also the section VTY Line Local Authentication and Authorization on page 948.

IP Prefix Lists

Prefix Lists are supported on platforms: C E S

IP prefix lists control routing policy. An IP prefix list is a series of sequential filters that contain a matching criterion (examine IP route prefix) and an action (permit or deny) to process routes. The filters are processed in sequence so that if a route prefix does not match the criterion in the first filter, the second filter (if configured) is applied. When the route prefix matches a filter, FTOS drops or forwards the packet based on the filter's designated action. If the route prefix does not match any of the filters in the prefix list, the route is dropped (that is, implicit deny).

A route prefix is an IP address pattern that matches on bits within the IP address. The format of a route prefix is A.B.C.D/X where A.B.C.D is a dotted-decimal address and /X is the number of bits that should be matched of the dotted decimal address. For example, in 112.24.0.0/16, the first 16 bits of the address 112.24.0.0 match all addresses between 112.24.0.0 to 112.24.255.255.

Below are some examples that permit or deny filters for specific routes using the **le** and **ge** parameters, where x.x.x.x/x represents a route prefix:

- To deny only /8 prefixes, enter `deny x.x.x.x/x ge 8 le 8`
- To permit routes with the mask greater than /8 but less than /12, enter `permit x.x.x.x/x ge 8 le 12`
- To deny routes with a mask less than /24, enter `deny x.x.x.x/x le 24`
- To permit routes with a mask greater than /20, enter `permit x.x.x.x/x ge 20`

The following rules apply to prefix lists:

- A prefix list without any permit or deny filters allows all routes.
- An "implicit deny" is assumed (that is, the route is dropped) for all route prefixes that do not match a permit or deny filter in a configured prefix list.
- Once a route matches a filter, the filter's action is applied. No additional filters are applied to the route.

Implementation Information

In FTOS, prefix lists are used in processing routes for routing protocols (for example, RIP, OSPF, and BGP).

Note: The S-Series platform does not support all protocols. It is important to know which protocol you are supporting prior to implementing Prefix-Lists.

Configuration Task List for Prefix Lists

To configure a prefix list, you must use commands in the PREFIX LIST, the ROUTER RIP, ROUTER OSPF, and ROUTER BGP modes. Basically, you create the prefix list in the PREFIX LIST mode, and assign that list to commands in the ROUTER RIP, ROUTER OSPF and ROUTER BGP modes.

The following list includes the configuration tasks for prefix lists:

- Configure a prefix list on page 154
- Use a prefix list for route redistribution on page 156

For a complete listing of all commands related to prefix lists, refer to the *FTOS Command Line Interface Reference* document.

Configure a prefix list

To configure a prefix list, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip prefix-list** *prefix-name*	CONFIGURATION	Create a prefix list and assign it a unique name. You are in the PREFIX LIST mode.
2	**seq** *sequence-number* {**deny** | **permit**} *ip-prefix* [**ge** *min-prefix-length*] [**le** *max-prefix-length*]	CONFIG-NPREFIXL	Create a prefix list with a sequence number and a deny or permit action. The optional parameters are: • **ge** *min-prefix-length:* is the minimum prefix length to be matched (0 to 32). • **le** *max-prefix-length:* is the maximum prefix length to be matched (0 to 32).

If you want to forward all routes that do not match the prefix list criteria, you must configure a prefix list filter to permit all routes (**permit 0.0.0.0/0 le 32**). The "permit all" filter should be the last filter in your prefix list. To permit the default route only, enter **permit 0.0.0.0/0**.

Figure 8-13 illustrates how the **seq** command orders the filters according to the sequence number assigned. In the example, filter 20 was configured before filter 15 and 12, but the **show config** command displays the filters in the correct order.

Figure 8-13. Command Example: seq

```
FTOS(conf-nprefixl)#seq 20 permit 0.0.0.0/0 le 32
FTOS(conf-nprefixl)#seq 12 deny 134.23.0.0 /16
FTOS(conf-nprefixl)#seq 15 deny 120.23.14.0 /8 le 16
FTOS(conf-nprefixl)#show config
!
ip prefix-list juba
 seq 12 deny 134.23.0.0/16
 seq 15 deny 120.0.0.0/8 le 16
 seq 20 permit 0.0.0.0/0 le 32
FTOS(conf-nprefixl)#
```

Note the last line in the prefix list Juba contains a "permit all" statement. By including this line in a prefix list, you specify that all routes not matching any criteria in the prefix list are forwarded.

To delete a filter, use the **no seq** *sequence-number* command in the PREFIX LIST mode.

If you are creating a standard prefix list with only one or two filters, you can let FTOS assign a sequence number based on the order in which the filters are configured. The FTOS assigns filters in multiples of five.

To configure a filter without a specified sequence number, use these commands in the following sequence starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip prefix-list** *prefix-name*	CONFIGURATION	Create a prefix list and assign it a unique name.
2	{**deny** | **permit**} *ip-prefix* [**ge** *min-prefix-length*] [**le** *max-prefix-length*]	CONFIG-NPREFIXL	Create a prefix list filter with a deny or permit action. The optional parameters are: • **ge** *min-prefix-length:* is the minimum prefix length to be matched (0 to 32). • **le** *max-prefix-length:* is the maximum prefix length to be matched (0 to 32).

Figure 8-14 illustrates a prefix list in which the sequence numbers were assigned by the software. The filters were assigned sequence numbers based on the order in which they were configured (for example, the first filter was given the lowest sequence number). The **show config** command in the PREFIX LIST mode displays the two filters with the sequence numbers 5 and 10.

Figure 8-14. Prefix List

```
FTOS(conf-nprefixl)#permit 123.23.0.0 /16
FTOS(conf-nprefixl)#deny 133.24.56.0 /8
FTOS(conf-nprefixl)#show conf
!
ip prefix-list awe
 seq 5 permit 123.23.0.0/16
 seq 10 deny 133.0.0.0/8
FTOS(conf-nprefixl)#
```

To delete a filter, enter the **show config** command in the PREFIX LIST mode and locate the sequence number of the filter you want to delete; then use the **no seq** *sequence-number* command in the PREFIX LIST mode.

To view all configured prefix lists, use either of the following commands in the EXEC mode:

Command Syntax	Command Mode	Purpose
show ip prefix-list detail [*prefix-name*]	EXEC Privilege	Show detailed information about configured Prefix lists.
show ip prefix-list summary [*prefix-name*]	EXEC Privilege	Show a table of summarized information about configured Prefix lists.

Figure 8-15. Command example: show ip prefix-list detail

```
FTOS>show ip prefix detail
Prefix-list with the last deletion/insertion: filter_ospf
ip prefix-list filter_in:
count: 3, range entries: 3, sequences: 5 - 10
   seq 5 deny 1.102.0.0/16 le 32 (hit count: 0)
   seq 6 deny 2.1.0.0/16 ge 23 (hit count: 0)
   seq 10 permit 0.0.0.0/0 le 32 (hit count: 0)
ip prefix-list filter_ospf:
count: 4, range entries: 1, sequences: 5 - 10
   seq 5 deny 100.100.1.0/24 (hit count: 0)
   seq 6 deny 200.200.1.0/24 (hit count: 0)
   seq 7 deny 200.200.2.0/24 (hit count: 0)
   seq 10 permit 0.0.0.0/0 le 32 (hit count: 0)
FTOS>
```

Figure 8-16. Command Example: show ip prefix-list summary

```
FTOS>show ip prefix summary
Prefix-list with the last deletion/insertion: filter_ospf
ip prefix-list filter_in:
count: 3, range entries: 3, sequences: 5 - 10
ip prefix-list filter_ospf:
count: 4, range entries: 1, sequences: 5 - 10
FTOS>
```

Use a prefix list for route redistribution

To pass traffic through a configured prefix list, you must use the prefix list in a route redistribution command. The prefix list is applied to all traffic redistributed into the routing process and the traffic is either forwarded or dropped depending on the criteria and actions specified in the prefix list.

To apply a filter to routes in RIP (RIP is supported on C and E-Series.), use either of the following commands in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
router rip	CONFIGURATION	Enter RIP mode
distribute-list *prefix-list-name* **in** [*interface*]	CONFIG-ROUTER-RIP	Apply a prefix list to filter the network prefixes in incoming route updates. You can specify an interface. If you enter the name of a nonexistent prefix list, all routes are forwarded.
distribute-list *prefix-list-name* **out** [*interface* | **connected** | **static** | **ospf**]	CONFIG-ROUTER-RIP	Apply a prefix list to filter network prefixes advertised in outgoing route updates. You can specify an interface or type of route. If you enter the name of a non-existent prefix list, all routes are forwarded.

To view the configuration, use the **show config** command in the ROUTER RIP mode (Figure 240) or the **show running-config rip** command in the EXEC mode.

Figure 8-17. Command Example: show config in the ROUTER RIP Mode

```
FTOS(conf-router_rip)#show config
!
router rip
 distribute-list prefix juba out
 network 10.0.0.0
FTOS(conf-router_rip)#router ospf 34
```

To apply a filter to routes in OSPF, use either of the following commands in the ROUTER OSPF mode:

Command Syntax	Command Mode	Purpose
router ospf	CONFIGURATION	Enter OSPF mode
distribute-list *prefix-list-name* **in** [*interface*]	CONFIG-ROUTER-OSPF	Apply a configured prefix list to incoming routes. You can specify an interface. If you enter the name of a non-existent prefix list, all routes are forwarded.
distribute-list *prefix-list-name* **out** [**connected** | **rip** | **static**]	CONFIG-ROUTER-OSPF	Apply a configured prefix list to incoming routes. You can specify which type of routes are affected. If you enter the name of a non-existent prefix list, all routes are forwarded.

To view the configuration, use the **show config** command in the ROUTER OSPF mode (Figure 241) or the **show running-config ospf** command in the EXEC mode.

Figure 8-18. Command Example: show config in ROUTER OSPF Mode

```
FTOS(conf-router_ospf)#show config
!
router ospf 34
 network 10.2.1.1 255.255.255.255 area 0.0.0.1
 distribute-list prefix awe in
FTOS(conf-router_ospf)#
```

ACL Resequencing

Resequencing an ACL or Prefix List is supported on platform E

ACL Resequencing allows you to re-number the rules and remarks in an access or prefix list. The placement of rules within the list is critical because packets are matched against rules in sequential order. Use Resequencing whenever there is no longer an opportunity to order new rules as desired using current numbering scheme.

For example, Table 8-3 contains some rules that are numbered in increments of 1. No new rules can be placed between these, so apply resequencing to create numbering space, as shown in Table 8-4. In the same example, apply resequencing if more than two rules must be placed between rules 7 and 10.

IPv4 and IPv6 ACLs and prefixes and MAC ACLs can be resequenced. No CAM writes happen as a result of resequencing, so there is no packet loss; the behavior is like Hot-lock ACLs.

Note: ACL Resequencing does not affect the rules or remarks or the order in which they are applied. It merely renumbers them so that new rules can be placed within the list as desired.

Table 8-3. ACL Resequencing Example (Insert New Rules)

seq 5 permit any host 1.1.1.1
seq 6 permit any host 1.1.1.2
seq 7 permit any host 1.1.1.3
seq 10 permit any host 1.1.1.4

Table 8-4. ACL Resequencing Example (Resequenced)

seq 5 permit any host 1.1.1.1
seq 10 permit any host 1.1.1.2
seq 15 permit any host 1.1.1.3
seq 20 permit any host 1.1.1.4

Resequencing an ACL or Prefix List

Resequencing is available for IPv4 and IPv6 ACLs and prefix lists and MAC ACLs. To resequence an ACL or prefix list use the appropriate command in Table 8-5. You must specify the list name, starting number, and increment when using these commands.

Table 8-5. Resequencing ACLs and Prefix Lists

List	Command	Command Mode
IPv4, IPv6, or MAC ACL	**resequence access-list** {**ipv4** | **ipv6** | **mac**} {*access-list-name StartingSeqNum Step-to-Increment*}	Exec
IPv4 or IPv6 prefix-list	**resequence prefix-list** {**ipv4** | **ipv6**} {*prefix-list-name StartingSeqNum Step-to-Increment*}	Exec

Figure 8-19 shows the resequencing of an IPv4 access-list beginning with the number 2 and incrementing by 2.

Figure 8-19. Resequencing ACLs

```
FTOS(config-ext-nacl)# show config
!
ip access-list extended test
remark 4 XYZ
remark 5 this remark corresponds to permit any host 1.1.1.1
seq 5 permit ip any host 1.1.1.1
remark 9 ABC
remark 10 this remark corresponds to permit ip any host 1.1.1.2
seq 10 permit ip any host 1.1.1.2
seq 15 permit ip any host 1.1.1.3
seq 20 permit ip any host 1.1.1.4
FTOS# end
FTOS# resequence access-list ipv4 test 2 2
FTOS# show running-config acl
!
ip access-list extended test
remark 2 XYZ
remark 4 this remark corresponds to permit any host 1.1.1.1
seq 4 permit ip any host 1.1.1.1
remark 6 this remark has no corresponding rule
remark 8 this remark corresponds to permit ip any host 1.1.1.2
seq 8 permit ip any host 1.1.1.2
seq 10 permit ip any host 1.1.1.3
seq 12 permit ip any host 1.1.1.4
```

Remarks and rules that originally have the same sequence number have the same sequence number after the **resequence** command is applied. Remarks that do not have a corresponding rule will be incremented as as a rule. These two mechanisms allow remarks to retain their original position in the list.

For example, in Figure 8-20, remark 10 corresponds to rule 10 and as such they have the same number before and after the command is entered. Remark 4 is incremented as a rule, and all rules have retained their original positions.

Figure 8-20. Resequencing Remarks

```
FTOS(config-ext-nacl)# show config
!
ip access-list extended test
remark 4 XYZ
remark 5 this remark corresponds to permit any host 1.1.1.1
seq 5 permit ip any host 1.1.1.1
remark 9 ABC
remark 10 this remark corresponds to permit ip any host 1.1.1.2
seq 10 permit ip any host 1.1.1.2
seq 15 permit ip any host 1.1.1.3
seq 20 permit ip any host 1.1.1.4
FTOS# end
FTOS# resequence access-list ipv4 test 2 2
FTOS# show running-config acl
!
ip access-list extended test
remark 2 XYZ
remark 4 this remark corresponds to permit any host 1.1.1.1
seq 4 permit ip any host 1.1.1.1
remark 6 this remark has no corresponding rule
remark 8 this remark corresponds to permit ip any host 1.1.1.2
seq 8 permit ip any host 1.1.1.2
seq 10 permit ip any host 1.1.1.3
seq 12 permit ip any host 1.1.1.4
```

Route Maps

Route-maps are supported on platforms: C E S

Like ACLs and prefix lists, route maps are composed of a series of commands that contain a matching criterion and an action, yet route maps can change the packets meeting the criterion. ACLs and prefix lists can only drop or forward the packet or traffic. Route maps process routes for route redistribution. For example, a route map can be called to filter only specific routes and to add a metric.

Route maps also have an “implicit deny.” Unlike ACLs and prefix lists, however, where the packet or traffic is dropped, in route maps, if a route does not match any of the route map conditions, the route is not redistributed.

Implementation Information

The FTOS implementation of route maps allows route maps with no match command or no set command. When there is no match command, all traffic matches the route map and the set command applies.

Important Points to Remember

- For route-maps with more than one match clause:
  - Two or more match clauses within the same route-map sequence have the *same* match commands (though the values are different), matching a packet against these clauses is a logical OR operation.
  - Two or more match clauses within the same route-map sequence have *different* match commands, matching a packet against these clauses is a logical AND operation.
- If no match is found in a route-map sequence, the process moves to the next route-map sequence until a match is found, or there are no more sequences.
- When a match is found, the packet is forwarded; no more route-map sequences are processed.
  - If a continue clause is included in the route-map sequence, the next or a specified route-map sequence is processed after a match is found.

Configuration Task List for Route Maps

You configure route maps in the ROUTE-MAP mode and apply them in various commands in the ROUTER RIP and ROUTER OSPF modes.

The following list includes the configuration tasks for route maps:

- Create a route map on page 161 (mandatory)
- Configure route map filters on page 163 (optional)
- Configure a route map for route redistribution on page 166 (optional)
- Configure a route map for route tagging on page 167 (optional)

Create a route map

Route maps, ACLs, and prefix lists are similar in composition because all three contain filters, but route map filters are do not contain the permit and deny actions found in ACLs and prefix lists. Route map filters match certain routes and set or specify values.

To create a route map and enter the ROUTE-MAP mode, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
route-map *map-name* [**permit** | **deny**] [*sequence-number*]	CONFIGURATION	Create a route map and assign it a unique name. The optional **permit** and **deny** keywords are the action of the route map. The default is **permit**. The optional parameter **seq** allows you to assign a sequence number to the route map instance.

The default action is permit and the default sequence number starts at 10. When the keyword **deny** is used in configuring a route map, routes that meet the match filters are not redistributed.

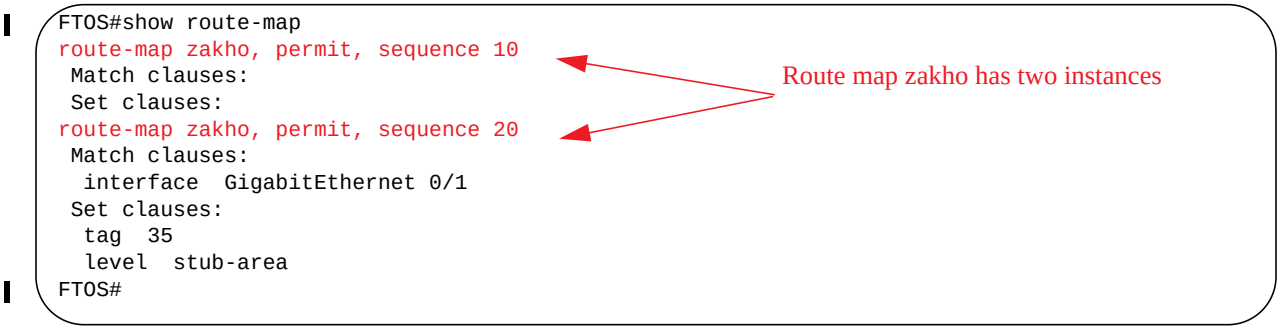
To view the configuration, use the **show config** command in the ROUTE-MAP mode (Figure 8-21).

Figure 8-21. Command Example: show config in the ROUTE-MAP Mode

```
FTOS(config-route-map)#show config
!
route-map dilling permit 10
FTOS(config-route-map)#
```

You can create multiple instances of this route map by using the sequence number option to place the route maps in the correct order. FTOS processes the route maps with the lowest sequence number first. When a configured route map is applied to a command, like **redistribute**, traffic passes through all instances of that route map until a match is found. Figure 8-22 shows an example with two instances of a route map.

Figure 8-22. Command Example: show route-map with Multiple Instances of a Route Map



To delete all instances of that route map, use the **no route-map** *map-name* command. To delete just one instance, add the sequence number to the command syntax (Figure 8-24).

Figure 8-23. Deleting One Instance of a Route Map

```
FTOS(conf)#no route-map zakho 10
FTOS(conf)#end
FTOS#show route-map
route-map zakho, permit, sequence 20
 Match clauses:
  interface  GigabitEthernet 0/1
 Set clauses:
  tag  35
  level  stub-area
FTOS#
```

Figure 8-24 shows an example of a route map with multiple instances. The **show config** command displays only the configuration of the current route map instance. To view all instances of a specific route map, use the **show route-map** command.

Figure 8-24. Command Example: show route-map

```
FTOS#show route-map dilling
route-map dilling, permit, sequence 10
 Match clauses:
 Set clauses:
route-map dilling, permit, sequence 15
 Match clauses:
  interface  Loopback 23
 Set clauses:
  tag  3444
FTOS#
```

To delete a route map, use the **no route-map** *map-name* command in the CONFIGURATION mode.

Configure route map filters

Within the ROUTE-MAP mode, there are **match** and **set** commands. Basically, **match** commands search for a certain criterion in the routes and the **set** commands change the characteristics of those routes, either adding something or specifying a level.

When there are multiple match commands of the same parameter under one instance of route-map, then FTOS does a match between either of those match commands. If there are multiple match commands of different parameter, then FTOS does a match ONLY if there is a match among ALL match commands. The following example explains better:

Example 1

```
FTOS(conf)#route-map force permit 10
FTOS(config-route-map)#match tag 1000
FTOS(config-route-map)#match tag 2000
FTOS(config-route-map)#match tag 3000
```

In the above route-map, if a route has any of the tag value specified in the match commands, then there is a match.

Example 2

```
FTOS(conf)#route-map force permit 10
FTOS(config-route-map)#match tag 1000
FTOS(config-route-map)#match metric 2000
```

In the above route-map, *only* if a route has *both* the characteristics mentioned in the route-map, it is matched. Explaining further, the route *must* have a tag value of 1000 *and* a metric value of 2000. Only then is there a match.

Also, if there are different instances of the same route-map, then it's sufficient if a permit match happens in *any* instance of that route-map. As an example:

```
FTOS(conf)#route-map force permit 10
FTOS(config-route-map)#match tag 1000

FTOS(conf)#route-map force deny 20
FTOS(config-route-map)#match tag 1000

FTOS(conf)#route-map force deny 30
FTOS(config-route-map)#match tag 1000
```

In the above route-map, instance 10 permits the route having a tag value of 1000 and instances 20 & 30 denies the route having a tag value of 1000. In the above scenario, FTOS scans all the instances of the route-map for any permit statement. If there is a match anywhere, the route is permitted, though other instances of the route-map denies it.

To configure match criterion for a route map, use any or all of the following commands in the ROUTE-MAP mode:

Command Syntax	Command Mode	Purpose
match as-path *as-path-name*	CONFIG-ROUTE-MAP	Match routes with the same AS-PATH numbers.
match community *community-list-name* [**exact**]	CONFIG-ROUTE-MAP	Match routes with COMMUNITY list attributes in their path.
match interface *interface*	CONFIG-ROUTE-MAP	Match routes whose next hop is a specific interface. The parameters are: • For a Fast Ethernet interface, enter the keyword **FastEthernet** followed by the slot/port information. • For a 1-Gigabit Ethernet interface, enter the keyword **gigabitEthernet** followed by the slot/port information. • For a loopback interface, enter the keyword **loopback** followed by a number between zero (0) and 16383. • For a port channel interface, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale. • For a SONET interface, enter the keyword **sonet** followed by the slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **tengigabitEthernet** followed by the slot/port information. • For a VLAN, enter the keyword vlan followed by a number from 1 to 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

Command Syntax	Command Mode	Purpose
match ip address *prefix-list-name*	CONFIG-ROUTE-MAP	Match destination routes specified in a prefix list (IPv4).
match ipv6 address *prefix-list-name*	CONFIG-ROUTE-MAP	Match destination routes specified in a prefix list (IPv6).
match ip next-hop {*access-list-name* | **prefix-list** *prefix-list-name*}	CONFIG-ROUTE-MAP	Match next-hop routes specified in a prefix list (IPv4).
match ipv6 next-hop {*access-list-name* | **prefix-list** *prefix-list-name*}	CONFIG-ROUTE-MAP	Match next-hop routes specified in a prefix list (IPv6).
match ip route-source {*access-list-name* | **prefix-list** *prefix-list-name*}	CONFIG-ROUTE-MAP	Match source routes specified in a prefix list (IPv4).
match ipv6 route-source {*access-list-name* | **prefix-list** *prefix-list-name*}	CONFIG-ROUTE-MAP	Match source routes specified in a prefix list (IPv6).
match metric *metric-value*	CONFIG-ROUTE-MAP	Match routes with a specific value.
match origin {**egp** | **igp** | **incomplete**}	CONFIG-ROUTE-MAP	Match BGP routes based on the ORIGIN attribute.
match route-type {**external** [**type-1** | **type-2**] **| internal | level-1 | level-2** | **local** }	CONFIG-ROUTE-MAP	Match routes specified as internal or external to OSPF, ISIS level-1, ISIS level-2, or locally generated.
match tag *tag-value*	CONFIG-ROUTE-MAP	Match routes with a specific tag.

To configure a set condition, use any or all of the following commands in the ROUTE-MAP mode:

Command Syntax	Command Mode	Purpose
set as-path prepend *as-number* [... *as-number*]	CONFIG-ROUTE-MAP	Add an AS-PATH number to the beginning of the AS-PATH
set automatic-tag	CONFIG-ROUTE-MAP	Generate a tag to be added to redistributed routes.
set level {**backbone** | **level-1** | **level-1-2** | **level-2** | **stub-area** }	CONFIG-ROUTE-MAP	Specify an OSPF area or ISIS level for redistributed routes.
set local-preference *value*	CONFIG-ROUTE-MAP	Specify a value for the BGP route's LOCAL_PREF attribute.
set metric {+ | - | *metric-value*}	CONFIG-ROUTE-MAP	Specify a value for redistributed routes.
set metric-type {**external** | **internal** | **type-1** | **type-2**}	CONFIG-ROUTE-MAP	Specify an OSPF or ISIS type for redistributed routes.
set next-hop *ip-address*	CONFIG-ROUTE-MAP	Assign an IP address as the route's next hop.

Command Syntax	Command Mode	Purpose
set ipv6 next-hop *ip-address*	CONFIG-ROUTE-MAP	Assign an IPv6 address as the route's next hop.
set origin {**egp** | **igp** | **incomplete**}	CONFIG-ROUTE-MAP	Assign an ORIGIN attribute.
set tag *tag-value*	CONFIG-ROUTE-MAP	Specify a tag for the redistributed routes.
set weight *value*	CONFIG-ROUTE-MAP	Specify a value as the route's weight.

Use these commands to create route map instances. There is no limit to the number of set and match commands per route map, but the convention is to keep the number of match and set filters in a route map low. **Set** commands do not require a corresponding **match** command.

Configure a route map for route redistribution

Route maps on their own cannot affect traffic and must be included in different commands to affect routing traffic. To apply a route map to traffic on the E-Series, you must call or include that route map in a command such as the **redistribute** or **default-information originate** commands in OSPF, ISIS, and BGP.

Route redistribution occurs when FTOS learns the advertising routes from static or directly connected routes or another routing protocol. Different protocols assign different values to redistributed routes to identify either the routes and their origins. The metric value is the most common attribute that is changed to properly redistribute other routes into a routing protocol. Other attributes that can be changed include the metric type (for example, external and internal route types in OSPF) and route tag. Use the **redistribute** command in OSPF, RIP, ISIS, and BGP to set some of these attributes for routes that are redistributed into those protocols.

Route maps add to that redistribution capability by allowing you to match specific routes and set or change more attributes when redistributing those routes.

In Figure 8-25, the **redistribute** command calls the route map `static ospf` to redistribute only certain static routes into OSPF. According to the route map `static ospf`, only routes that have a next hop of Gigabitethernet interface 0/0 and that have a metric of 255 will be redistributed into the OSPF backbone area.

Note: When re-distributing routes using route-maps, the user must take care to create the route-map defined in the **redistribute** command under the routing protocol. If no route-map is created, then NO routes are redistributed.

Figure 8-25. Route Redistribution into OSPF

```
router ospf 34
 default-information originate metric-type 1
 redistribute static metric 20 metric-type 2 tag 0 route-map staticospf
!
route-map staticospf permit 10
 match interface  GigabitEthernet 0/0
 match metric  255
 set level  backbone
```

Configure a route map for route tagging

One method for identifying routes from different routing protocols is to assign a tag to routes from that protocol. As the route enters a different routing domain, it is tagged and that tag is passed along with the route as it passes through different routing protocols. This tag can then be used when the route leaves a routing domain to redistribute those routes again.

In Figure 8-26, the **redistribute ospf** command with a route map is used in the ROUTER RIP mode to apply a tag of 34 to all internal OSPF routes that are redistributed into RIP.

Figure 8-26. Tagging OSPF Routes Entering a RIP Routing Domain

```
!
router rip
 redistribute ospf 34 metric 1 route-map torip
!
route-map torip permit 10
 match route-type  internal
 set tag  34
!
```

Continue clause

Normally, when a match is found, set clauses are executed, and the packet is then forwarded; no more route-map modules are processed. If the **continue** command is configured at the end of a module, the next module (or a specified module) is processed even after a match is found. Figure 8-27 shows a continue clause at the end of a route-map module. In this example, if a match is found in the route-map “test” module 10, module 30 will be processed.

Note: If the continue clause is configured without specifying a module, the next sequential module is processed.

Figure 8-27. Command Example: continue

```
!
route-map test permit 10
match commu comm-list1
set community 1:1 1:2 1:3
set as-path prepend 1 2 3 4 5
continue 30!
```

9

Bidirectional Forwarding Detection

Bidirectional Forwarding Detection is supported only on platforms: C E

BFD is supported on E-Series ExaScale E X with FTOS 8.2.1.0 and later.

Protocol Overview

Bidirectional Forwarding Detection (BFD) is a protocol that is used to rapidly detect communication failures between two adjacent systems. It is a simple and lightweight replacement for existing routing protocol link state detection mechanisms. It also provides a failure detection solution for links on which no routing protocol is used.

BFD is a simple hello mechanism. Two neighboring systems running BFD establish a session using a three-way handshake. After the session has been established, the systems exchange periodic control packets at sub-second intervals. If a system does not receive a hello packet within a specified amount of time, routing protocols are notified that the forwarding path is down.

BFD provides forwarding path failure detection times on the order of milliseconds rather than seconds as with conventional routing protocol hellos. It is independent of routing protocols, and as such provides a consistent method of failure detection when used across a network. Networks converge faster because BFD triggers link state changes in the routing protocol sooner and more consistently, because BFD can eliminate the use of multiple protocol-dependent timers and methods.

BFD also carries less overhead than routing protocol hello mechanisms. Control packets can be encapsulated in any form that is convenient, and, on Dell Force10 routers, sessions are maintained by BFD Agents that reside on the line card, which frees resources on the RPM. Only session state changes are reported to the BFD Manager (on the RPM), which in turn notifies the routing protocols that are registered with it.

BFD is an independent and generic protocol, which all media, topologies, and routing protocols can support using any encapsulation. Dell Force10 has implemented BFD at Layer 3 and with UDP encapsulation. BFD functionality will be implemented in phases. OSPF, IS-IS (not on C-Series), VRRP, VLANs, LAGs, static routes, and physical ports support BFD, based on the IETF internet draft *draft-ietf-bfd-base-03*.

How BFD Works

Two neighboring systems running BFD establish a session using a three-way handshake. After the session has been established, the systems exchange control packets at agreed upon intervals. In addition, systems send a control packet anytime there is a state change or change in a session parameter; these control packets are sent without regard to transmit and receive intervals.

Note: FTOS does not support multi-hop BFD sessions.

If a system does not receive a control packet within an agreed-upon amount of time, the BFD Agent changes the session state to Down. It then notifies the BFD Manager of the change, and sends a control packet to the neighbor that indicates the state change (though it might not be received if the link or receiving interface is faulty). The BFD Manager notifies the routing protocols that are registered with it (clients) that the forwarding path is down, and a link state change is triggered in all protocols.

Note: A session state change from Up to Down is the only state change that triggers a link state change in the routing protocol client.

BFD packet format

Control packets are encapsulated in UDP packets. Figure 9-1 shows the complete encapsulation of a BFD control packet inside an IPv4 packet.

Figure 9-1. BFD in IPv4 Packet Format

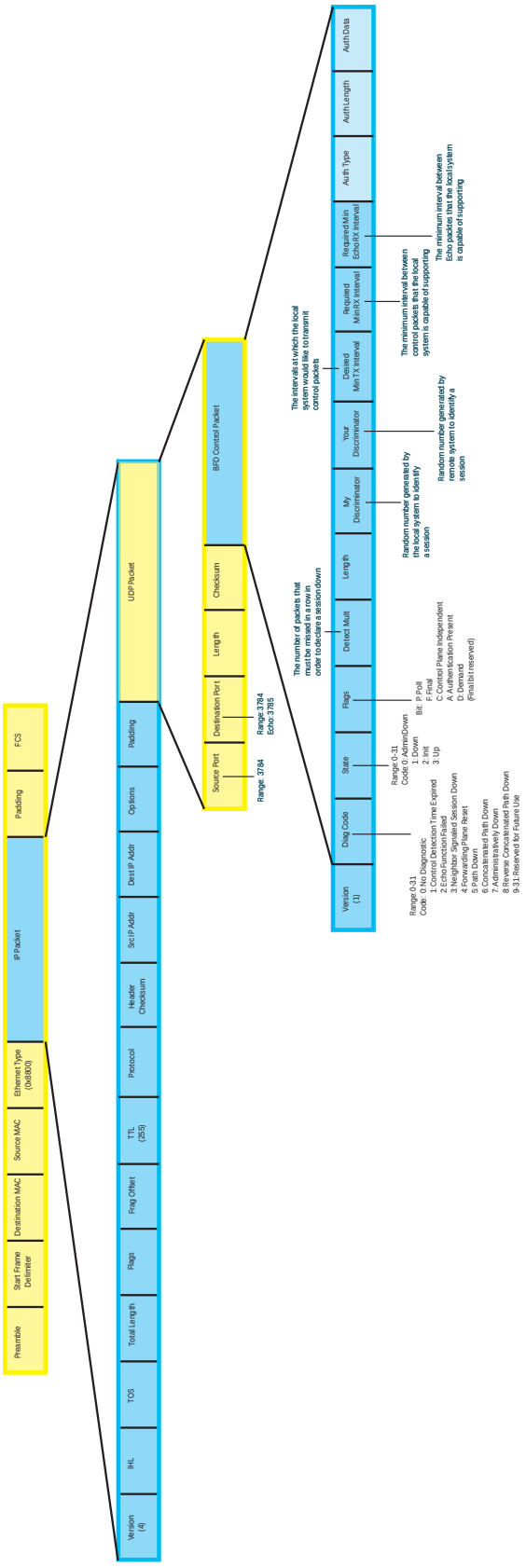


Table 9-1. BFD Packet Fields

Field	Description
Diagnostic Code	The reason that the last session failed.
State	The current local session state. See BFD sessions.
Flag	A bit that indicates packet function. If the poll bit is set, the receiving system must respond as soon as possible, without regard to its transmit interval. The responding system clears the poll bit and sets the final bit in its response. The poll and final bits are used during the handshake and Demand mode (see BFD sessions). **Note:** FTOS does not currently support multi-point sessions, Demand mode, authentication, or control plane independence; these bits are always clear.
Detection Multiplier	The number of packets that must be missed in order to declare a session down.
Length	The entire length of the BFD packet.
My Discriminator	A random number generated by the local system to identify the session.
Your Discriminator	A random number generated by the remote system to identify the session. Discriminator values are necessary to identify the session to which a control packet belongs since there can be many sessions running on a single interface.
Desired Min TX Interval	The minimum rate at which the local system would like to send control packets to the remote system.
Required Min RX Interval	The minimum rate at which the local system would like to receive control packets from the remote system.
Required Min Echo RX	The minimum rate at which the local system would like to receive echo packets. **Note:** FTOS does not currently support the echo function.
Authentication Type Authentication Length Authentication Data	An optional method for authenticating control packets. **Note:** FTOS does not currently support the BFD authentication function.

Two important parameters are calculated using the values contained in the control packet.

- **Transmit interval —** Transmit interval is the agreed-upon rate at which a system sends control packets. Each system has its own transmit interval, which is the greater of the last received remote Desired TX Interval and the local Required Min RX Interval.
- **Detection time —** Detection time is the amount of time that a system does not receive a control packet, after which the system determines that the session has failed. Each system has its own detection time.
  - In Asynchronous mode: Detection time is the remote Detection Multiplier multiplied by greater of the remote Desired TX Interval and the local Required Min RX Interval.
  - In Demand mode: Detection time is the local Detection Multiplier multiplied by the greater of the local Desired Min TX and the remote Required Min RX Interval.

BFD sessions

BFD must be enabled on both sides of a link in order to establish a session. The two participating systems can assume either of two roles:

- **Active**—The active system initiates the BFD session. Both systems can be active for the same session.
- **Passive**—The passive system does not initiate a session. It only responds to a request for session initialization from the active system.

A BFD session has two modes:

- **Asynchronous mode**—In Asynchronous mode, both systems send periodic control messages at an agreed upon interval to indicate that their session status is Up.
- **Demand mode**—If one system requests Demand mode, the other system stops sending periodic control packets; it only sends a response to status inquiries from the Demand mode initiator. Either system (but not both) can request Demand mode at any time.

Note: FTOS supports asynchronous mode only.

A session can have four states: Administratively Down, Down, Init, and Up.

- **Administratively Down**—The local system will not participate in a particular session.
- **Down**—The remote system is not sending any control packets or at least not within the detection time for a particular session.
- **Init**—The local system is communicating.
- **Up**—The both systems are exchanging control packets.

The session is declared down if:

- A control packet is not received within the detection time.
- Sufficient echo packets are lost.
- Demand mode is active and a control packet is not received in response to a poll packet.

BFD three-way handshake

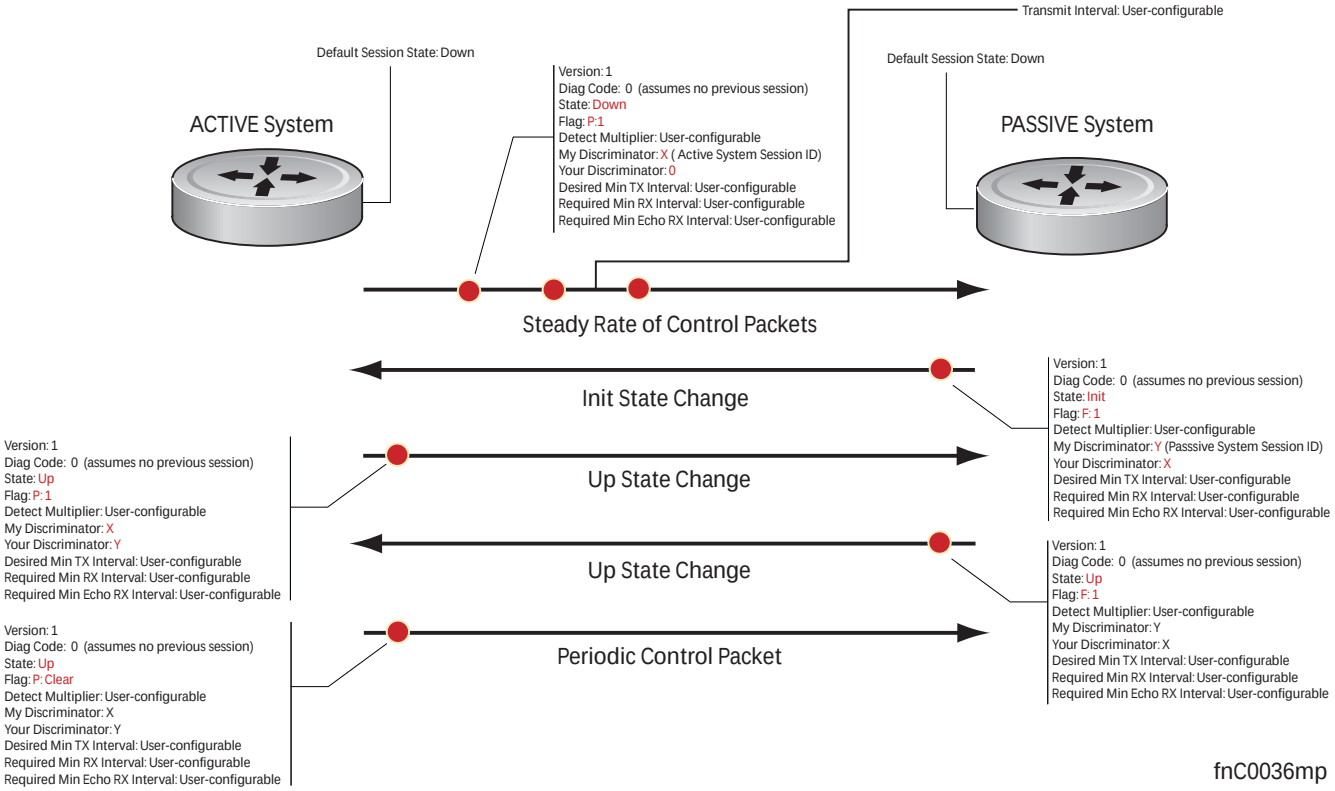
A three-way handshake must take place between the systems that will participate in the BFD session. The handshake shown in Figure 9-2 assumes that there is one active and one passive system, and that this is the first session established on this link. The default session state on both ports is Down.

1. The active system sends a steady stream of control packets that indicates that its session state is Down, until the passive system responds. These packets are sent at the desired transmit interval of the Active system, and the Your Discriminator field is set to zero.
2. When the passive system receives any of these control packets, it changes its session state to Init, and sends a response that indicates its state change. The response includes its session ID in the My Discriminator field, and the session ID of the remote system in the Your Discriminator field.
3. The active system receives the response from the passive system, and changes its session state to Up. It then sends a control packet indicating this state change. This is the third and final part of of the

handshake. At this point, the discriminator values have been exchanged, and the transmit intervals have been negotiated.

4. The passive system receives the control packet, changes its state to Up. Both systems agree that a session has been established. However, since both members must send a control packet—that requires a response—anytime there is a state change or change in a session parameter, the passive system sends a final response indicating the state change. After this, periodic control packets are exchanged.

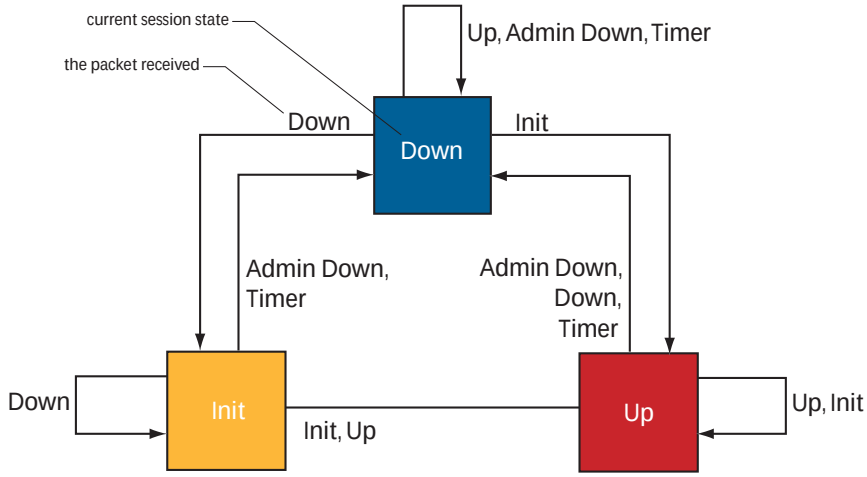
Figure 9-2. BFD Three-way Handshake



Session state changes

Figure 9-3 shows how the session state on a system changes based on the status notification it receives from the remote system. For example, if a session on a system is down, and it receives a Down status notification from the remote system, the session state on the local system changes to Init.

Figure 9-3. BFD State Machine



Important Points to Remember

- BFD for line card ports is hitless, but is not hitless for VLANs since they are instantiated on the RPM.
- BFD is supported on C-Series and E-Series only.
- FTOS supports a maximum of 100 sessions per BFD agent. Each linecard processor has a BFD Agent, so the limit translates to 100 BFD sessions per linecard (plus, on the E-Series, 100 BFD sessions on RP2, which handles LAG and VLANs).
- BFD must be enabled on both ends of a link.
- Demand mode, authentication, and the Echo function are not supported.
- BFD is not supported on multi-hop and virtual links.
- Protocol Liveness is supported for routing protocols only.
- FTOS supports only OSPF, ISIS (E-Series only), and VRRP protocols as BFD clients.

Configuring Bidirectional Forwarding Detection

The remainder of this chapter is divided into the following sections:

Configuring BFD for Physical Ports

BFD on physical ports is useful when no routing protocol is enabled. Without BFD, if the remote system fails, the local system does not remove the connected route until the first failed attempt to send a packet. When BFD is enabled, the local system removes the route as soon as it stops receiving periodic control packets from the remote system.

Configuring BFD for a physical port is a two-step process:

1. Enable BFD globally. See page 176.
2. Establish a session with a next-hop neighbor. See page 176.

Related configuration tasks

- Change session parameters. See page 178.
- Disable or re-enable BFD on an interface. See page 179.

Enabling BFD globally

BFD must be enabled globally on both routers, as shown in Figure 9-5.

To enable BFD globally:

Step	Task	Command Syntax	Command Mode
1	Enable BFD globally.	**bfd enable**	CONFIGURATION

Verify that BFD is enabled globally using the command **show running bfd**, as shown in Figure 9-4.

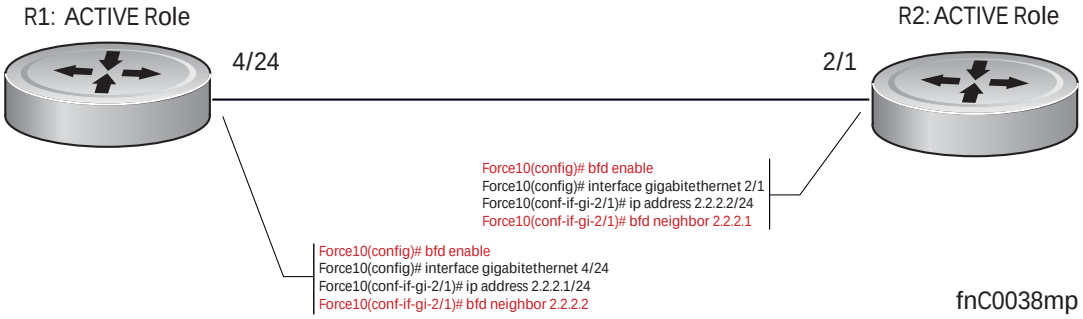
Figure 9-4. Enabling BFD Globally

```
R1(conf)#bfd ?
enable                  Enable BFD protocol
protocol-liveness       Enable BFD protocol-liveness
R1(conf)#bfd enable

R1(conf)#do show running-config bfd
!
bfd enable                                  <-- BFD Enabled Globally
R1(conf)#
```

Establishing a session on physical ports

To establish a session, BFD must be enabled at interface level on both ends of the link, as shown in Figure 9-5. The configuration parameters do not need to match.

Figure 9-5. Establishing a BFD Session for Physical Ports



To establish a session:

Step	Task	Command Syntax	Command Mode
1	Enter interface mode	**interface**	CONFIGURATION
2	Assign an IP address to the interface if one is not already assigned.	**ip address** *ip-address*	INTERFACE
3	Identify the neighbor with which the interface will participate in the BFD session.	**bfd neighbor** *ip-address*	INTERFACE

Verify that the session is established using the command **show bfd neighbors**, as shown in Figure 9-6.

Figure 9-6. Viewing Established Sessions for Physical Ports

```
R1(conf-if-gi-4/24)#do show bfd neighbors
*       - Active session role
Ad Dn   - Admin Down
C       - CLI
I       - ISIS
O       - OSPF
R       - Static Route (RTM)

  LocalAddr       RemoteAddr      Interface State Rx-int Tx-int Mult Clients
* 2.2.2.1         2.2.2.2         Gi 4/24   Up    100    100    3    C
                                                                     BFD Session Enabled
```

The command **show bfd neighbors detail** shows more specific information about BFD sessions (Figure 9-7).

Figure 9-7. Viewing Session Details

```
R1(conf-if-gi-4/24)#do show bfd neighbors detail

Session Discriminator: 1
Neighbor Discriminator: 1
Local Addr: 2.2.2.1
Local MAC Addr: 00:01:e8:09:c3:e5
Remote Addr: 2.2.2.2
Remote MAC Addr: 00:01:e8:06:95:a2
Int: GigabitEthernet 4/24
State: Up
Configured parameters:
 TX:  100ms, RX:  100ms, Multiplier: 3
Neighbor parameters:
 TX:  100ms, RX:  100ms, Multiplier: 3
Actual parameters:
 TX:  100ms, RX:  100ms, Multiplier: 3
Role: Active
Delete session on Down: False
Client Registered: CLI
Uptime: 00:03:57
Statistics:
 Number of packets received from neighbor: 1775
 Number of packets sent to neighbor: 1775
 Number of state changes: 1
 Number of messages from IFA about port state change: 0
 Number of messages communicated b/w Manager and Agent: 4
```

When both interfaces are configured for BFD, log messages are displayed indicating state changes, as shown in Message 1.

Message 1 BFD Session State Changes

```
R1(conf-if-gi-4/24)#00:36:01: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE: Changed session state to Down for
neighbor 2.2.2.2 on interface Gi 4/24 (diag: 0)
00:36:02: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE: Changed session state to Up for neighbor 2.2.2.2 on
interface Gi 4/24 (diag: 0)
```

Changing physical port session parameters

BFD sessions are configured with default intervals and a default role (active). The parameters that can be configured are: Desired TX Interval, Required Min RX Interval, Detection Multiplier, and system role. These parameters are configured per interface; if you change a parameter, the change affects all physical port sessions on that interface. Dell Force10 recommends maintaining the default values.

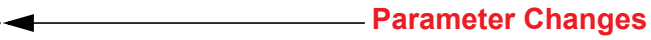
To change session parameters on an interface:

Step	Task	Command Syntax	Command Mode
1	Change session parameters for all sessions on an interface.	**bfd interval** *milliseconds* **min_rx** *milliseconds* **multiplier** *value* **role** [**active** | **passive**]	INTERFACE

View session parameters using the **show bfd neighbors detail** command.

Figure 9-8. Changing Session Parameters for Physical Ports

```
R1(conf-if-gi-4/24)#bfd interval 100 min_rx 100 multiplier 4 role passive
R1(conf-if-gi-4/24)#do show bfd neighbors detail

Session Discriminator: 1
Neighbor Discriminator: 1
Local Addr: 2.2.2.1
Local MAC Addr: 00:01:e8:09:c3:e5
Remote Addr: 2.2.2.2
Remote MAC Addr: 00:01:e8:06:95:a2
Int: GigabitEthernet 4/24
State: Up
Configured parameters:
 TX:  100ms, RX:  100ms, Multiplier: 4
Neighbor parameters:
 TX:  100ms, RX:  100ms, Multiplier: 3
Actual parameters:
 TX:  100ms, RX:  100ms, Multiplier: 4
Role: Passive
Delete session on Down: False
Client Registered: CLI
Uptime: 00:09:06
Statistics:
 Number of packets received from neighbor: 4092
 Number of packets sent to neighbor: 4093
 Number of state changes: 1
 Number of messages from IFA about port state change: 0
 Number of messages communicated b/w Manager and Agent: 7
```



Disabling and re-enabling BFD

BFD is enabled on all interfaces by default, though sessions are not created unless explicitly configured. If BFD is disabled, all of the sessions on that interface are placed in an Administratively Down state (Message 2), and the remote systems are notified of the session state change (Message 3).

To disable BFD on an interface:

Step	Task	Command Syntax	Command Mode
1	Disable BFD on an interface.	**no bfd enable**	INTERFACE

Message 2 Disabling BFD on a Local Interface

```
R1(conf-if-gi-4/24)#01:00:52: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE: Changed session state to Ad Dn for
neighbor 2.2.2.2 on interface Gi 4/24 (diag: 0)
```

Message 3 Remote System State Change due to Local State Admin Down

```
R2>01:32:53: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE: Changed session state to Down for neighbor 2.2.2.1
on interface Gi 2/1 (diag: 7)
```

To re-enable BFD on an interface:

Step	Task	Command Syntax	Command Mode
1	Enable BFD on an interface.	**bfd enable**	INTERFACE

Configuring BFD for Static Routes

BFD gives systems a link state detection mechanism for static routes. With BFD, systems are notified to remove static routes from the routing table as soon as the link state change occurs, rather than having to wait until packets fail to reach their next hop.

Configuring BFD for static routes is a three-step process:

1. Enable BFD globally. See Enabling BFD globally on page 176.
2. On the local system, establish a session with the next hop of a static route. See page 180.
3. On the remote system, establish a session with the physical port that is the origin of the static route. See Establishing a session on physical ports on page 176.

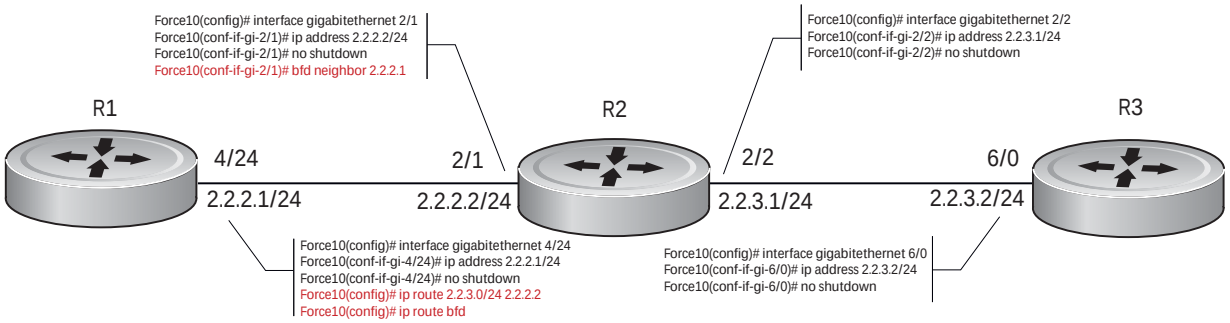
Related configuration tasks

- Change session parameters. See page 181.
- Disable BFD for all static routes. See page 181.

Establishing sessions for static routes

Sessions are established for all neighbors that are the next hop of a static route.

Figure 9-9. Enabling BFD for Static Routes



fnC0039mp

To establish a BFD session:

Step	Task	Command Syntax	Command Mode
1	Establish BFD sessions for all neighbors that are the next hop of a static route.	**ip route bfd**	CONFIGURATION

Verify that sessions have been created for static routes using the command **show bfd neighbors**, as shown in Figure 9-10. View detailed session information using the command **show bfd neighbors detail**, as shown in Figure 9-8.

Figure 9-10. Viewing Established Sessions for Static Routes

```
R1(conf)#ip route 2.2.3.0/24 2.2.2.2
R1(conf)#ip route bfd
R1(conf)#do show bfd neighbors

*       - Active session role
Ad Dn   - Admin Down
C       - CLI
I       - ISIS
O       - OSPF
R       - Static Route (RTM)

  LocalAddr       RemoteAddr      Interface State Rx-int Tx-int Mult Clients
  2.2.2.1         2.2.2.2         Gi 4/24   Up    100    100    4       R
```

BFD for Static Routes Enabled

Changing static route session parameters

BFD sessions are configured with default intervals and a default role. The parameters that can be configured are: Desired TX Interval, Required Min RX Interval, Detection Multiplier, and system role. These parameters are configured for all static routes; if you change a parameter, the change affects all sessions for static routes.

To change parameters for static route sessions:

Step	Task	Command Syntax	Command Mode
1	Change parameters for all static route sessions.	**ip route bfd interval** *milliseconds* **min_rx** *milliseconds* **multiplier** *value* **role** [**active** | **passive**]	CONFIGURATION

View session parameters using the command **show bfd neighbors detail**, as shown in Figure 9-8 on page 179.

Disabling BFD for static routes

If BFD is disabled, all static route BFD sessions are torn down. A final Admin Down packet is sent to all neighbors on the remote systems, and those neighbors change to the Down state (Message 3 on page 179).

To disable BFD for static routes:

Step	Task	Command Syntax	Command Mode
1	Disable BFD for static routes.	**no ip route bfd**	CONFIGURATION

Configuring BFD for OSPF

When using BFD with OSPF, the OSPF protocol registers with the BFD manager on the RPM. BFD sessions are established with all neighboring interfaces participating in OSPF. If a neighboring interface fails, the BFD agent on the line card notifies the BFD manager, which in turn notifies the OSPF protocol that a link state change occurred.

Configuring BFD for OSPF is a two-step process:

1. Enable BFD globally. See Enabling BFD globally on page 176.
2. Establish sessions for all or particular OSPF neighbors. See page 182.

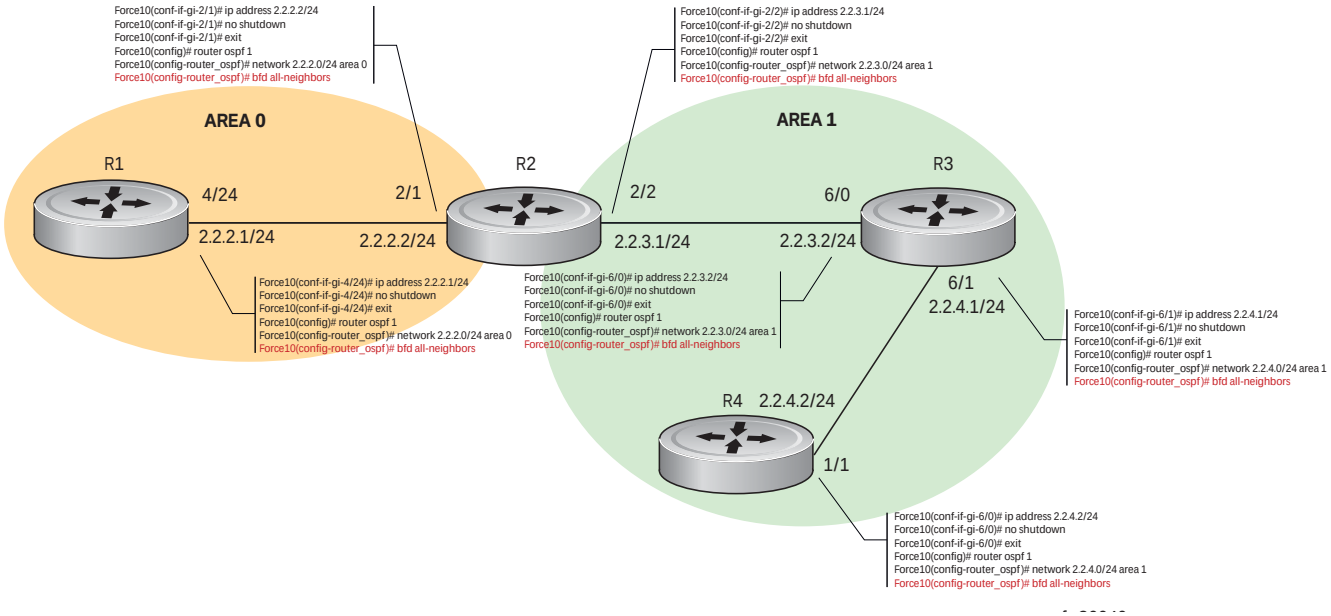
Related configuration tasks

- Change session parameters. See page 184.
- Disable BFD sessions for OSPF. See page 184.

Establishing sessions with OSPF neighbors

BFD sessions can be established with all OSPF neighbors at once, or sessions can be established with all neighbors out of a specific interface. Sessions are only established when the OSPF adjacency is in the full state.

Figure 9-11. Establishing Sessions with OSPF Neighbors



To establish BFD with all OSPF neighbors:

Step	Task	Command Syntax	Command Mode
1	Establish sessions with all OSPF neighbors.	**bfd all-neighbors**	ROUTER-OSPF

To establish BFD for all OSPF neighbors on a single interface:

Step	Task	Command Syntax	Command Mode
1	Establish sessions with all OSPF neighbors on a single interface.	**ip ospf bfd all-neighbors**	INTERFACE

View the established sessions using the command **show bfd neighbors**, as shown in Figure 9-12.

Figure 9-12. Viewing Established Sessions for OSPF Neighbors

```
R2(conf-router_ospf)#bfd all-neighbors
R2(conf-router_ospf)#do show bfd neighbors

*       - Active session role
Ad Dn   - Admin Down
C       - CLI
I       - ISIS
O       - OSPF
R       - Static Route (RTM)

  LocalAddr       RemoteAddr      Interface State Rx-int Tx-int Mult Clients
* 2.2.2.2         2.2.2.1         Gi 2/1    Up    100    100    3    0
```

OSPF BFD Sessions Enabled

Changing OSPF session parameters

BFD sessions are configured with default intervals and a default role. The parameters that can be configured are: Desired TX Interval, Required Min RX Interval, Detection Multiplier, and system role. These parameters are configured for all OSPF sessions or all OSPF sessions on a particular interface; if you change a parameter globally, the change affects all OSPF neighbors sessions. If you change a parameter at interface level, the change affects all OSPF sessions on that interface.

To change parameters for all OSPF sessions:

Step	Task	Command Syntax	Command Mode
1	Change parameters for OSPF sessions.	**bfd all-neighbors interval** *milliseconds* **min_rx** *milliseconds* **multiplier** *value* **role** [**active** | **passive**]	ROUTER-OSPF

To change parameters for OSPF sessions on an interface:

Step	Task	Command Syntax	Command Mode
1	Change parameters for all OSPF sessions on an interface.	**ip ospf bfd all-neighbors interval** *milliseconds* **min_rx** *milliseconds* **multiplier** *value* **role** [**active** | **passive**]	INTERFACE

View session parameters using the command **show bfd neighbors detail**, as shown in Figure 9-8 on page 179.

Disabling BFD for OSPF

If BFD is disabled globally, all sessions are torn down, and sessions on the remote system are placed in a Down state. If BFD is disabled on an interface, sessions on the interface are torn down, and sessions on the remote system are placed in a Down state (Message 3 on page 179). Disabling BFD does not trigger a change in BFD clients; a final Admin Down packet is sent before the session is terminated.

To disable BFD sessions with all OSPF neighbors:

Step	Task	Command Syntax	Command Mode
1	Disable BFD sessions with all OSPF neighbors.	**no bfd all-neighbors**	ROUTER-OSPF

To disable BFD sessions with all OSPF neighbors out of an interface:

Step	Task	Command Syntax	Command Mode
1	Disable BFD sessions with all OSPF neighbors out of an interface	**ip ospf bfd all-neighbors disable**	INTERFACE

Configuring BFD for BGP

BFD for BGP is only supported on platforms: E C S4810

In a BGP core network, BFD provides rapid detection of communication failures in BGP fast-forwarding paths between internal BGP (iBGP) and external BGP (eBGP) peers for faster network reconvergence. BFD for BGP is supported on 1GE, 10GE, 40GE, port-channel, and VLAN interfaces. BFD for BGP does not support IPv6 and the BGP multihop feature.

Prerequisites

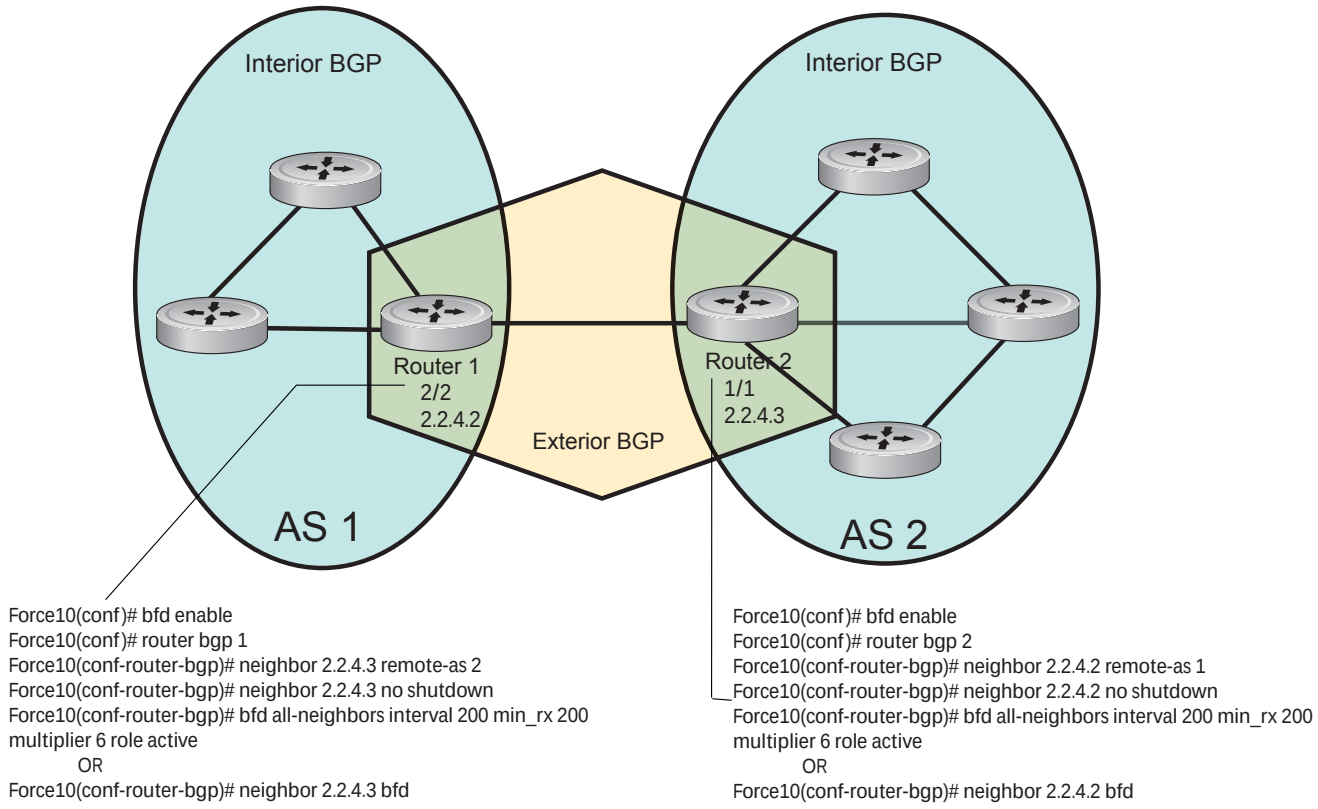
Before configuring BFD for BGP, you must first perform the following tasks:

1. Configure BGP on the routers that you want to interconnect as described in BGP Configuration on page 225.
2. Enable fast fall-over for BGP neighbors to reduce convergence time (**neighbor fall-over** command) as described in BGP fast fall-over on page 235.

Establishing sessions with BGP neighbors

Figure 9-13 shows a sample BFD configuration on Router 1 and Router 2 that use eBGP in a transit network to interconnect AS1 and AS2. The eBGP routers exchange information with each other as well as with iBGP routers to maintain connectivity and accessibility within each autonomous system.

Figure 9-13. BFD Session Between BGP Neighbors



Note that the sample configuration shows alternative ways to establish a BFD session with a BGP neighbor:

- By establishing BFD sessions with all neighbors discovered by BGP (**bfd all-neighbors** command)
- By establishing a BFD session with a specified BGP neighbor (**neighbor** {*ip-address* | *peer-group-name*} **bfd** command)

BFD packets originating from a router are assigned to the highest priority egress queue to minimize transmission delays. Incoming BFD control packets received from the BGP neighbor are assigned to the highest priority queue within the Control Plane Policing (COPP) framework to avoid BFD packets drops due to queue congestion.

BFD notifies BGP of any failure conditions that it detects on the link. Recovery actions are initiated by BGP.

BFD for BGP is supported only on directly-connected BGP neighbors and only in BGP IPv4 networks.

- On an E-Series TeraScale or C-Series router, up to 100 simultaneous BFD sessions are supported per line card.
- On an S4810 router, up to 64 simultaneous BFD sessions are supported.

As long as each BFD for BGP neighbor receives a BFD control packet within the configured BFD interval for failure detection, the BFD session remains up and BGP maintains its adjacencies. If a BFD for BGP neighbor does not receive a control packet within the detection interval, the router informs any clients of the BFD session (other routing protocols) about the failure. It then depends on the individual routing protocols that uses the BGP link to determine the appropriate response to the failure condition. The typical response is usually to terminate the peering session for the routing protocol and reconverge by bypassing the failed neighboring router. A log message is generated whenever BFD detects a failure condition.

You can configure BFD for BGP on the following types of interfaces: physical port (10GE or 40GE), port channel, and VLAN.

To establish a BFD session with one or all BGP neighbors, follow these steps:

Step	Task	Command Syntax	Command Mode
1	Enable BFD globally.	**bfd enable**	CONFIGURATION
2	Specify the AS number and enter ROUTER BGP configuration mode.	**router bgp** *as-number*	CONFIGURATION
3	Add a BGP neighbor or peer group in a remote AS.	**neighbor** {*ip-address* | *peer-group name*} **remote-as** *as-number*	CONFIG-ROUTER-BGP
4	Enable the BGP neighbor.	**neighbor** {*ip-address* | *peer-group-name*} **no shutdown**	CONFIG-ROUTER-BGP

Step	Task	Command Syntax	Command Mode
5	Configure parameters for a BFD session established with all neighbors discovered by BGP. OR Establish a BFD session with a specified BGP neighbor or peer group using the default BFD session parameters.	**bfd all-neighbors** [**interval** *millisecs* **min_rx** *millisecs* **multiplier** *value* **role** {**active | passive**}] OR **neighbor** {*ip-address* | *peer-group-name*} **bfd**	CONFIG-ROUTER-BGP
	Notes: - When you establish a BFD session with a specified BGP neighbor or peer group using the **neighbor bfd** command, the default BFD session parameters are used (interval: 100 milliseconds, min_rx: 100 milliseconds, multiplier: 3 packets, and role: active). - When you explicitly enable or disable a BGP neighbor for a BFD session with the **neighbor bfd** or **neighbor bfd disable** commands: - The neighbor does not inherit the BFD enable/disable values configured with the **bfd all-neighbors** command or configured for the peer group to which the neighbor belongs. - The neighbor only inherits the global timer values configured with the **bfd all-neighbors** command (**interval**, **min_rx**, and **multiplier**).		
6	Repeat Steps 1 to 5 on each BGP peer participating in a BFD session.		

Disabling BFD for BGP

To disable a BFD for BGP session with a specified neighbor, enter the **neighbor** {*ip-address* | *peer-group-name*} **bfd disable** command in ROUTER BGP configuration mode.

To remove the disabled state of a BFD for BGP session with a specified neighbor, enter the **no neighbor** {*ip-address* | *peer-group-name*} **bfd disable** command in ROUTER BGP configuration mode. The BGP link with the neighbor returns to normal operation and uses the BFD session parameters globally configured with the **bfd all-neighbors** command or configured for the peer group to which the neighbor belongs.

Using BFD in a BGP Peer Group

If you establish a BFD session for the members of a peer group (**neighbor** *peer-group-name* **bfd** command in ROUTER BGP configuration mode), members of the peer group may have BFD:

- Explicitly enabled (**neighbor** *ip-address* **bfd** command)
- Explicitly disabled (**neighbor** *ip-address* **bfd disable** command)
- Inherited (neither explicitly enabled or disabled) according to the current BFD configuration of the peer group. For information on BGP peer groups, see Configure Peer Groups on page 232.

If you explicitly enable (or disable) a BGP neighbor for BFD that belongs to a peer group:

- The neighbor does not inherit the BFD enable/disable values configured with the **bfd all-neighbors** command or configured for the peer group to which the neighbor belongs.

- The neighbor inherits only the global timer values that are configured with the **bfd all-neighbors** command (interval, min_rx, and multiplier).

If you explicitly enable (or disable) a peer group for BFD that has no BFD parameters configured (e.g. advertisement interval) using the **neighbor** *peer-group-name* **bfd** command, the peer group inherits any BFD settings configured with the **bfd all-neighbors** command.

Displaying BFD for BGP Information

To display information about BFD for BGP sessions on a router, enter one of the following **show** commands:

Task	Command	Command Mode
Verify a BFD for BGP configuration.	**show running-config bgp** Figure 9-14	EXEC Privilege
Verify that a BFD for BGP session has been successfully established with a BGP neighbor. A line-by-line listing of established BFD adjacencies is displayed.	**show bfd neighbors** [*interface*] [**detail**] Figure 9-15 and Figure 9-16	EXEC Privilege
Display BFD packet counters for sessions with BGP neighbors.	**show bfd counters bgp** [*interface*] Figure 9-17	EXEC Privilege
Check to see if BFD is enabled for BGP connections.	**show ip bgp summary** Figure 9-18	EXEC Privilege
Displays routing information exchanged with BGP neighbors, including BFD for BGP sessions.	**show ip bgp neighbors** [*ip-address*] Figure 9-19	EXEC Privilege

The following examples show the BFD for BGP output displayed for these **show** commands.

Figure 9-14. Verifying a BFD for BGP Configuration: show running-config bgp Command

```
R2# show running-config bgp
!
router bgp 2
 neighbor 1.1.1.2 remote-as 1
 neighbor 1.1.1.2 no shutdown
 neighbor 2.2.2.2 remote-as 1
 neighbor 2.2.2.2 no shutdown
 neighbor 3.3.3.2 remote-as 1
 neighbor 3.3.3.2 no shutdown
 bfd all-neighbors
```

Figure 9-15. Verifying BFD Sessions with BGP Neighbors: show bfd neighbors Command

```
R2# show bfd neighbors

*       - Active session role
Ad Dn   - Admin Down
B       - BGP
C       - CLI
I       - ISIS
O       - OSPF
R       - Static Route (RTM)
M       - MPLS
V       - VRRP

  LocalAddr       RemoteAddr      Interface State Rx-int Tx-int Mult Clients
* 1.1.1.3         1.1.1.2         Te 6/0    Up    100    100    3    B
* 2.2.2.3         2.2.2.2         Te 6/1    Up    100    100    3    B
* 3.3.3.3         3.3.3.2         Te 6/2    Up    100    100    3    B
```

Figure 9-16. Verifying BFD Sessions with BGP Neighbors: show bfd neighbors detail Command

```
R2# show bfd neighbors detail

Session Discriminator: 9
Neighbor Discriminator: 10
Local Addr: 1.1.1.3
Local MAC Addr: 00:01:e8:66:da:33
Remote Addr: 1.1.1.2
Remote MAC Addr: 00:01:e8:8a:da:7b
Int: TenGigabitEthernet 6/0
State: Up
Configured parameters:
TX:  100ms, RX:  100ms, Multiplier: 3
Neighbor parameters:
TX:  100ms, RX:  100ms, Multiplier: 3
Actual parameters:
TX:  100ms, RX:  100ms, Multiplier: 3
Role: Active
Delete session on Down: True
Client Registered: BGP
Uptime: 00:07:55
Statistics:
Number of packets received from neighbor: 4762
Number of packets sent to neighbor: 4490
Number of state changes: 2
Number of messages from IFA about port state change: 0
Number of messages communicated b/w Manager and Agent: 5

Session Discriminator: 10
Neighbor Discriminator: 11
Local Addr: 2.2.2.3
Local MAC Addr: 00:01:e8:66:da:34
Remote Addr: 2.2.2.2
Remote MAC Addr: 00:01:e8:8a:da:7b
Int: TenGigabitEthernet 6/1
State: Up
Configured parameters:
TX:  100ms, RX:  100ms, Multiplier: 3
Neighbor parameters:
TX:  100ms, RX:  100ms, Multiplier: 3
Actual parameters:
TX:  100ms, RX:  100ms, Multiplier: 3
Role: Active
Delete session on Down: True
Client Registered: BGP
Uptime: 00:02:22
Statistics:
Number of packets received from neighbor: 1428
Number of packets sent to neighbor: 1428
Number of state changes: 1
Number of messages from IFA about port state change: 0
Number of messages communicated b/w Manager and Agent: 4
```

BFD session parameters: TX (packet transmission), RX (packet reception), and multiplier (maximum number of missed packets)

Figure 9-17. Displaying BFD Packet Counters: show bfd counters bgp Command

```
R2# show bfd counters bgp

Interface TenGigabitEthernet 6/0
Protocol BGP
Messages:
Registration     : 5
De-registration  : 4
Init             : 0
Up               : 6
Down             : 0
Admin Down       : 2

Interface TenGigabitEthernet 6/1
Protocol BGP
Messages:
Registration     : 5
De-registration  : 4
Init             : 0
Up               : 6
Down             : 0
Admin Down       : 2

Interface TenGigabitEthernet 6/2
Protocol BGP
Messages:
Registration     : 1
De-registration  : 0
Init             : 0
Up               : 1
Down             : 0
Admin Down       : 2
```

Figure 9-18. Displaying BFD for BGP Status: show ip bgp summary Command

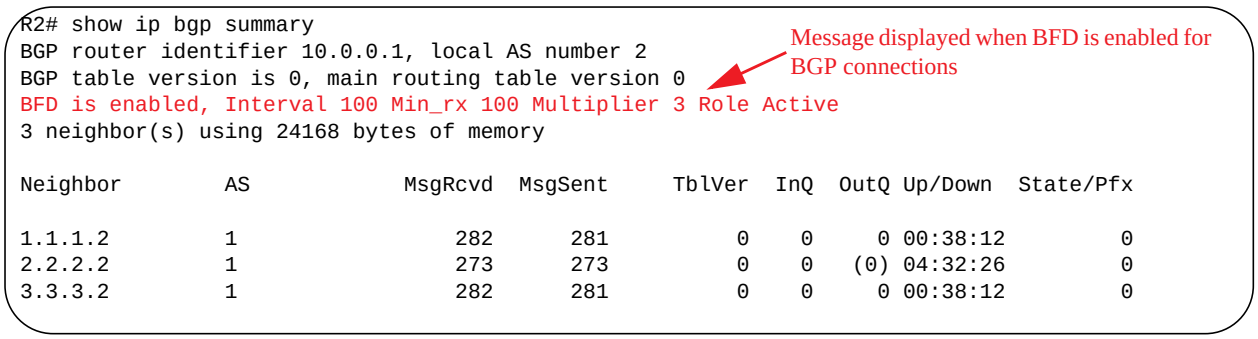


Figure 9-19. Displaying Routing Sessions with BGP Neighbors: show ip bgp neighbors Command

```
R2# show ip bgp neighbors 2.2.2.2

BGP neighbor is 2.2.2.2, remote AS 1, external link
  BGP version 4, remote router ID 12.0.0.4
  BGP state ESTABLISHED, in this state for 00:05:33
  Last read 00:00:30, last write 00:00:30
  Hold time is 180, keepalive interval is 60 seconds
  Received 8 messages, 0 in queue
    1 opens, 0 notifications, 0 updates
    7 keepalives, 0 route refresh requests
  Sent 9 messages, 0 in queue
    2 opens, 0 notifications, 0 updates
    7 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 30 seconds
  Minimum time before advertisements start is 0 seconds

  Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Capabilities advertised to neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Neighbor is using BGP global mode BFD configuration

  For address family: IPv4 Unicast
  BGP table version 0, neighbor version 0
  Prefixes accepted 0 (consume 0 bytes), withdrawn 0 by peer, martian prefixes ignored 0
  Prefixes advertised 0, denied 0, withdrawn 0 from peer

  Connections established 1; dropped 0
  Last reset never
Local host: 2.2.2.3, Local port: 63805
Foreign host: 2.2.2.2, Foreign port: 179
E1200i_ExaScale#


R2# show ip bgp neighbors 2.2.2.3

BGP neighbor is 2.2.2.3, remote AS 1, external link
  Member of peer-group pg1 for session parameters
  BGP version 4, remote router ID 12.0.0.4
  BGP state ESTABLISHED, in this state for 00:05:33
  ...
  Neighbor is using BGP neighbor mode BFD configuration
  Peer active in peer-group outbound optimization
...


R2# show ip bgp neighbors 2.2.2.4

BGP neighbor is 2.2.2.4, remote AS 1, external link
  Member of peer-group pg1 for session parameters
  BGP version 4, remote router ID 12.0.0.4
  BGP state ESTABLISHED, in this state for 00:05:33
  ...
  Neighbor is using BGP peer-group mode BFD configuration
  Peer active in peer-group outbound optimization
  ...
```

Message displayed when a BFD session is enabled with a BGP neighbor that inherits the global BFD session settings configured with the global **bfd all-neighbors** command

Message displayed when a BFD session with a BGP neighbor has been explicitly enabled using the **neighbor** *ip-address* **bfd** command

Message displayed when a BGP neighbor is in a peer group for which a BFD session has been explicitly enabled using the **neighbor** *peer-group-name* **bfd** command

Configuring BFD for IS-IS

BFD for IS-IS is supported on platform: E

When using BFD with IS-IS, the IS-IS protocol registers with the BFD manager on the RPM. BFD sessions are then established with all neighboring interfaces participating in IS-IS. If a neighboring interface fails, the BFD agent on the line card notifies the BFD manager, which in turn notifies the IS-IS protocol that a link state change occurred.

Configuring BFD for IS-IS is a two-step process:

1. Enable BFD globally. See Enabling BFD globally on page 176.
2. Establish sessions for all or particular IS-IS neighbors. See page 193.

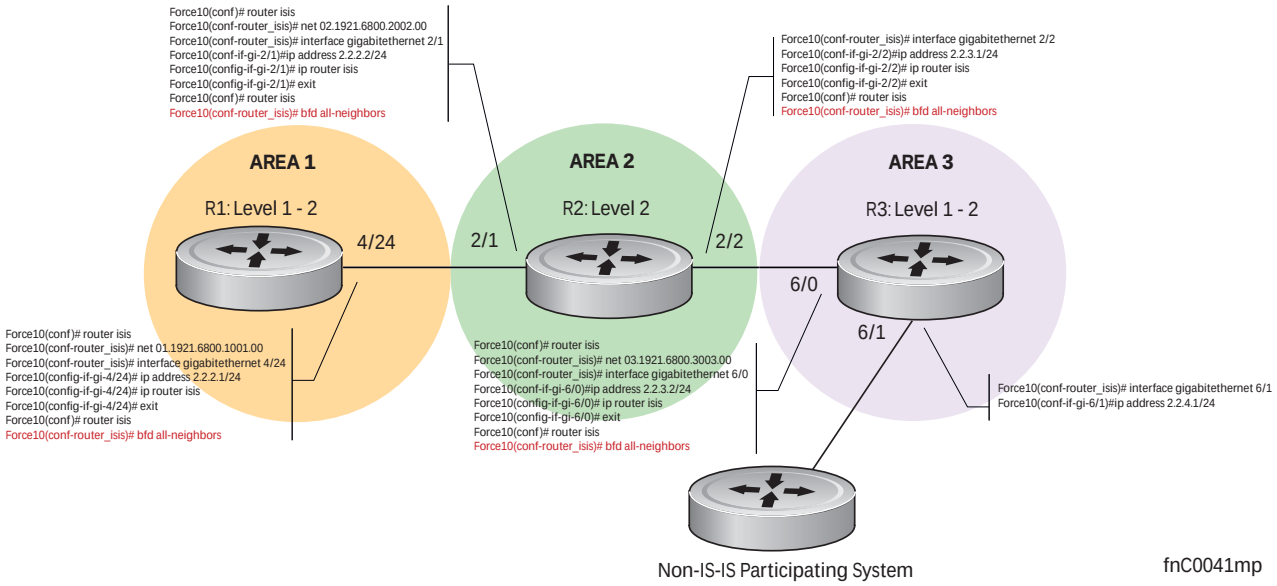
Related configuration tasks

- Change session parameters. See page 194.
- Disable BFD sessions for IS-IS. See page 195.

Establishing sessions with IS-IS neighbors

BFD sessions can be established for all IS-IS neighbors at once or sessions can be established for all neighbors out of a specific interface.

Figure 9-20. Establishing Sessions with IS-IS Neighbors



To establish BFD with all IS-IS neighbors:

Step	Task	Command Syntax	Command Mode
1	Establish sessions with all IS-IS neighbors.	**bfd all-neighbors**	ROUTER-ISIS

To establish BFD with all IS-IS neighbors out of a single interface:

Step	Task	Command Syntax	Command Mode
1	Establish sessions with all IS-IS neighbors out of an interface.	**isis bfd all-neighbors**	INTERFACE

View the established sessions using the command **show bfd neighbors**, as shown in Figure 9-21.

Figure 9-21. Viewing Established Sessions for IS-IS Neighbors

```
R2(conf-router_isis)#bfd all-neighbors
R2(conf-router_isis)#do show bfd neighbors

*       - Active session role
Ad Dn   - Admin Down
C       - CLI
I       - ISIS
O       - OSPF
R       - Static Route (RTM)

  LocalAddr       RemoteAddr      Interface State Rx-int Tx-int Mult Clients
* 2.2.2.2         2.2.2.1         Gi 2/1    Up    100    100    3    I
* 2.2.3.1         2.2.3.2         Gi 2/2    Up    100    100    3    I
```

IS-IS BFD Sessions Enabled

Changing IS-IS session parameters

BFD sessions are configured with default intervals and a default role. The parameters that can be configured are: Desired TX Interval, Required Min RX Interval, Detection Multiplier, and system role. These parameters are configured for all IS-IS sessions or all IS-IS sessions out of an interface; if you change a parameter globally, the change affects all IS-IS neighbors sessions. If you change a parameter at interface level, the change affects all IS-IS sessions on that interface.

To change parameters for all IS-IS sessions:

Step	Task	Command Syntax	Command Mode
1	Change parameters for all IS-IS sessions.	**bfd all-neighbors interval** *milliseconds* **min_rx** *milliseconds* **multiplier** *value* **role** [**active** | **passive**]	ROUTER-ISIS

To change parameters for IS-IS sessions on an interface:

Step	Task	Command Syntax	Command Mode
1	Change parameters for all IS-IS sessions out of an interface.	**isis bfd all-neighbors interval** *milliseconds* **min_rx** *milliseconds* **multiplier** *value* **role** [**active** | **passive**]	INTERFACE

View session parameters using the command **show bfd neighbors detail**, as shown in Figure 9-8 on page 179.

Disabling BFD for IS-IS

If BFD is disabled globally, all sessions are torn down, and sessions on the remote system are placed in a Down state. If BFD is disabled on an interface, sessions on the interface are torn down, and sessions on the remote system are placed in a Down state (Message 3 on page 179). Disabling BFD does not trigger a change in BFD clients; a final Admin Down packet is sent before the session is terminated.

To disable BFD sessions with all IS-IS neighbors:

Step	Task	Command Syntax	Command Mode
1	Disable BFD sessions with all IS-IS neighbors.	**no bfd all-neighbors**	ROUTER-ISIS

To disable BFD sessions with all IS-IS neighbors out of an interface:

Step	Task	Command Syntax	Command Mode
1	Disable BFD sessions with all IS-IS neighbors out of an interface	**isis bfd all-neighbors disable**	INTERFACE

Configuring BFD for VRRP

When using BFD with VRRP, the VRRP protocol registers with the BFD manager on the RPM. BFD sessions are established with all neighboring interfaces participating in VRRP. If a neighboring interface fails, the BFD agent on the line card notifies the BFD manager, which in turn notifies the VRRP protocol that a link state change occurred.

Configuring BFD for VRRP is a three-step process:

1. Enable BFD globally. See Enabling BFD globally on page 176.
2. Establish VRRP BFD sessions with all VRRP-participating neighbors.
3. On the master router, establish a VRRP BFD sessions with the backup routers. See page 195.

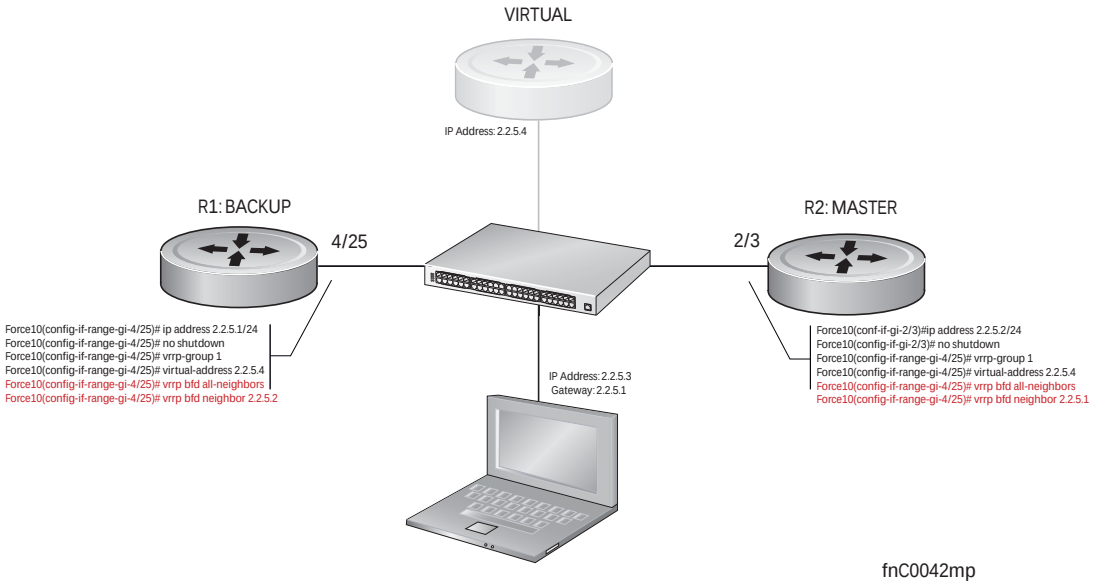
Related configuration tasks

- Change session parameters. See page 197.
- Disable or re-enable BFD on an interface. See page 182.

Establishing sessions with all VRRP neighbors

BFD sessions can be established for all VRRP neighbors at once, or a session can be established with a particular neighbor.

Figure 9-22. Establishing Sessions with VRRP Neighbors



To establish sessions with all VRRP neighbors:

Step	Task	Command Syntax	Command Mode
1	Establish sessions with all VRRP neighbors.	**vrrp bfd all-neighbors**	INTERFACE

Establishing VRRP sessions on VRRP neighbors

The master router does not care about the state of the backup router, so it does not participate in any VRRP BFD sessions. Therefore, VRRP BFD sessions on the backup router cannot change to the UP state. The master router must be configured to establish an individual VRRP session the backup router.

To establish a session with a particular VRRP neighbor:

Step	Task	Command Syntax	Command Mode
1	Establish a session with a particular VRRP neighbor.	**vrrp bfd neighbor** *ip-address*	INTERFACE

View the established sessions using the command **show bfd neighbors**, as shown in Figure 9-23.

Figure 9-23. Viewing Established Sessions for VRRP Neighbors

```
R1(conf-if-gi-4/25)#vrrp bfd all-neighbors
R1(conf-if-gi-4/25)#do show bfd neighbor

*       - Active session role
Ad Dn   - Admin Down
C       - CLI
I       - ISIS
O       - OSPF
R       - Static Route (RTM)
V       - VRRP

  LocalAddr       RemoteAddr      Interface State Rx-int Tx-int Mult Clients
* 2.2.5.1         2.2.5.2         Gi 4/25    Down  1000   1000   3    V
```

VRRP BFD Sessions Enabled

Session state information is also shown in the **show vrrp** command output, as shown in Figure 9-24.

Figure 9-24. Viewing Established Sessions for VRRP Neighbors

```
R1(conf-if-gi-4/25)#do show vrrp
------------------
GigabitEthernet 4/1, VRID: 1, Net: 2.2.5.1
State: Backup, Priority: 1, Master: 2.2.5.2
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 95, Bad pkts rcvd: 0, Adv sent: 933, Gratuitous ARP sent: 3
Virtual MAC address:
 00:00:5e:00:01:01
Virtual IP address:
 2.2.5.4
Authentication: (none)
BFD Neighbors:
RemoteAddr      State
2.2.5.2         Up
```

VRRP BFD Session State

Changing VRRP session parameters

BFD sessions are configured with default intervals and a default role. The parameters that can be configured are: Desired TX Interval, Required Min RX Interval, Detection Multiplier, and system role. You can change parameters for all VRRP sessions for a particular neighbor.

To change parameters for all VRRP sessions:

Step	Task	Command Syntax	Command Mode
1	Change parameters for all VRRP sessions.	**vrrp bfd all-neighbors interval** *milliseconds* **min_rx** *milliseconds* **multiplier** *value* **role** [**active** | **passive**]	INTERFACE

To change parameters for a particular VRRP session:

Step	Task	Command Syntax	Command Mode
1	Change parameters for a particular VRRP session.	**vrrp bfd neighbor** *ip-address* **interval** *milliseconds* **min_rx** *milliseconds* **multiplier** *value* **role** [**active** | **passive**]	INTERFACE

View session parameters using the command **show bfd neighbors detail**, as shown in Figure 9-8 on page 179.

Disabling BFD for VRRP

If any or all VRRP sessions are disabled, the sessions are torn down. A final Admin Down control packet is sent to all neighbors and sessions on the remote system change to the Down state (Message 3 on page 179).

To disable all VRRP sessions on an interface:

Step	Task	Command Syntax	Command Mode
1	Disable all VRRP sessions on an interface.	**no vrrp bfd all-neighbors**	INTERFACE

To disable all VRRP sessions in a particular VRRP group:

Step	Task	Command Syntax	Command Mode
1	Disable all VRRP sessions in a VRRP group.	**bfd disable**	VRRP

To disable a particular VRRP session:

Step	Task	Command Syntax	Command Mode
1	Disable a particular VRRP session on an interface.	**no vrrp bfd neighbor** *ip-address*	INTERFACE

Configuring BFD for VLANs

BFD on Dell Force10 systems is a Layer 3 protocol. Therefore, BFD is used with routed VLANs. BFD on VLANs is analogous to BFD on physical ports. If no routing protocol is enabled, and a remote system fails, the local system does not remove the connected route until the first failed attempt to send a packet. If BFD is enabled, the local system removes the route when it stops receiving periodic control packets from the remote system.

There is one BFD Agent for VLANs and port-channels, which resides on RP2 as opposed to the other agents which are on the line card. Therefore, the 100 total possible sessions that this agent can maintain is shared for VLANs and port-channels.

Configuring BFD for VLANs is a two-step process:

1. Enable BFD globally on all participating routers. See Enabling BFD globally on page 176.
2. Establish sessions with VLAN neighbors. See page 199.

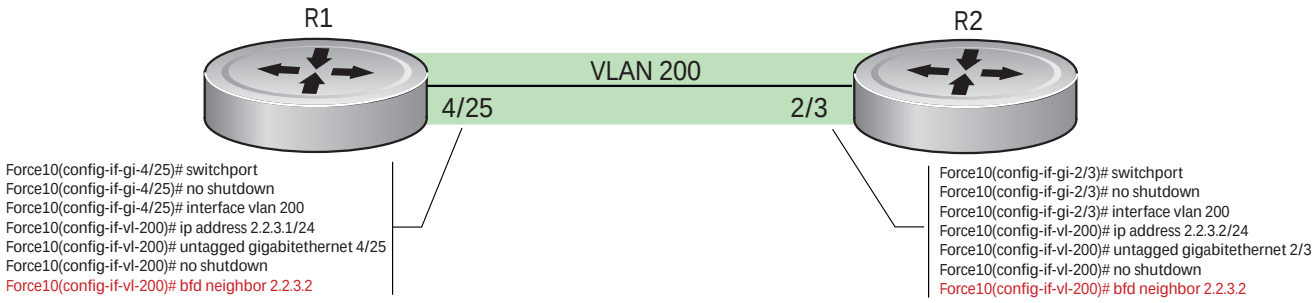
Related configuration tasks

- Change session parameters. See page 200.
- Disable BFD for VLANs. See page 182.

Establishing sessions with VLAN neighbors

To establish a session, BFD must be enabled at interface level on both ends of the link, as shown in Figure 9-25. The session parameters do not need to match.

Figure 9-25. Establishing Sessions with VLAN Neighbors



fnC0043mp

To establish a BFD session with a VLAN neighbor:

Step	Task	Command Syntax	Command Mode
1	Establish sessions with a VLAN neighbor.	**bfd neighbor** *ip-address*	INTERFACE VLAN

View the established sessions using the command **show bfd neighbors**, as shown in Figure 9-26.

Figure 9-26. Viewing Established Sessions for VLAN Neighbors

```
R2(conf-if-vl-200)#bfd neighbor 2.2.3.2
R2(conf-if-vl-200)#do show bfd neighbors

*       - Active session role
Ad Dn   - Admin Down
C       - CLI
I       - ISIS
O       - OSPF
R       - Static Route (RTM)
V       - VRRP

  LocalAddr       RemoteAddr      Interface State Rx-int Tx-int Mult Clients
* 2.2.3.2         2.2.3.1         Vl 200    Up    100    100    3    C
```

VLAN BFD Sessions Enabled

Changing session parameters

BFD sessions are configured with default intervals and a default role. The parameters that can be configured are: Desired TX Interval, Required Min RX Interval, Detection Multiplier, and system role. These parameters are configured per interface; if a configuration change is made, the change affects all sessions on that interface.

△ **Caution:** When configuring BFD on VLAN or LAG interfaces on the C-Series, Dell Force10 recommends a minimum value of 500 milliseconds for both the transmit and minimum receive time, which yields a final detection time of (500ms *3) 1500 milliseconds.

To change session parameters on an interface:

Step	Task	Command Syntax	Command Mode
1	Change session parameters for all sessions on an interface.	**bfd interval** *milliseconds* **min_rx** *milliseconds* **multiplier** *value* **role** [**active** | **passive**]	INTERFACE VLAN

View session parameters using the command **show bfd neighbors detail**, as shown in Figure 9-8 on page 179.

Disabling BFD for VLANs

If BFD is disabled on an interface, sessions on the interface are torn down. A final Admin Down control packet is sent to all neighbors, and sessions on the remote system change to the Down state (Message 3 on page 179).

To disable BFD on a VLAN interface:

Step	Task	Command Syntax	Command Mode
1	Disable all sessions on a VLAN interface.	**no bfd enable**	INTERFACE VLAN

Configuring BFD for Port-Channels

BFD on port-channels is analogous to BFD on physical ports. If no routing protocol is enabled, and a remote system fails, the local system does not remove the connected route until the first failed attempt to send a packet. If BFD is enabled, the local system removes the route when it stops receiving periodic control packets from the remote system.

There is one BFD Agent for VLANs and port-channels, which resides on RP2 as opposed to the other agents which are on the line card. Therefore, the 100 total possible sessions that this agent can maintain is shared for VLANs and port-channels.

Configuring BFD for port-channels is a two-step process:

1. Enable BFD globally on all participating routers. See Enabling BFD globally on page 176.
2. Enable BFD at interface level at both ends of the port-channel. See page 201.

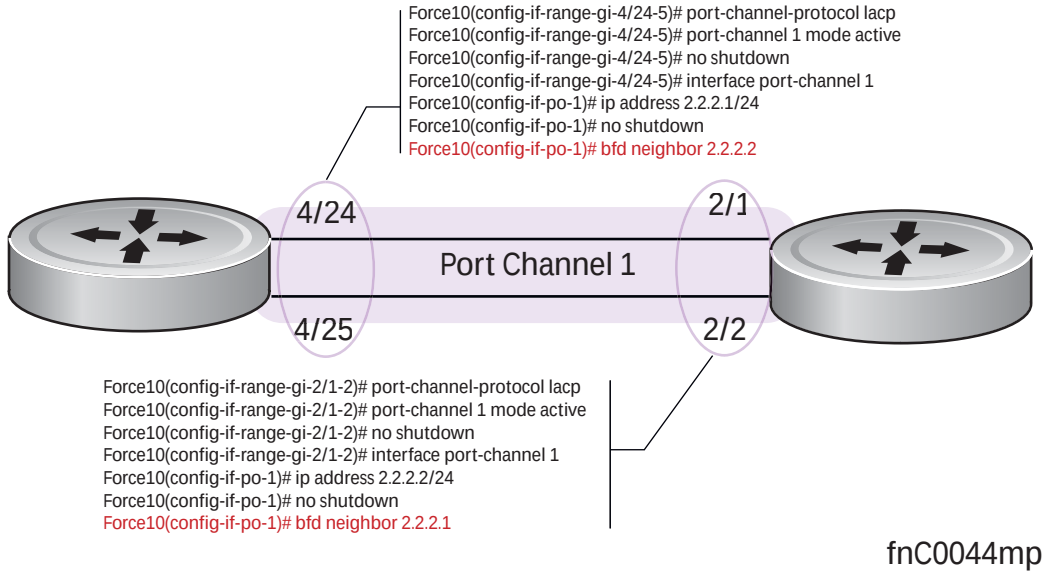
Related configuration tasks

- Change session parameters. See page 202.
- Disable BFD a port-channel. See page 202.

Establishing sessions on port-channels

To establish a session, BFD must be enabled at interface level on both ends of the link, as shown in Figure 9-5. The session parameters do not need to match.

Figure 9-27. Establishing Sessions on Port-Channels



To establish a session on a port-channel:

Step	Task	Command Syntax	Command Mode
1	Establish a session on a port-channel.	**bfd neighbor** *ip-address*	INTERFACE PORT-CHANNEL

View the established sessions using the command **show bfd neighbors**, as shown in Figure 9-21.

Figure 9-28. Viewing Established Sessions for VLAN Neighbors

```
R2(conf-if-po-1)#bfd neighbors 2.2.2.1
R2(conf-if-po-1)#do show bfd neighors

*       - Active session role
Ad Dn   - Admin Down
C       - CLI
I       - ISIS
O       - OSPF
R       - Static Route (RTM)
V       - VRRP

  LocalAddr       RemoteAddr      Interface State Rx-int Tx-int Mult Clients
* 2.2.2.2         2.2.2.1         Po 1      Up    100    100    3    C
```

Port-channel BFD Sessions Enabled

Changing port-channel session parameters

BFD sessions are configured with default intervals and a default role. The parameters that can be configured are: Desired TX Interval, Required Min RX Interval, Detection Multiplier, and system role. These parameters are configured per interface; if you change a parameter, the change affects all sessions on that interface.

△ **Caution:** When configuring BFD on VLAN or LAG interfaces on the C-Series, Dell Force10 recommends a minimum value of 500 milliseconds for both the transmit and minimum receive time, which yields a final detection time of (500ms *3) 1500 milliseconds.

To change session parameters on an interface:

Step	Task	Command Syntax	Command Mode
1	Change session parameters for all sessions on a port-channel interface.	**bfd interval** *milliseconds* **min_rx** *milliseconds* **multiplier** *value* **role** [**active** | **passive**]	INTERFACE PORT-CHANNEL

View session parameters using the command **show bfd neighbors detail**, as shown in Figure 9-8 on page 179.

Disabling BFD for port-channels

If BFD is disabled on an interface, sessions on the interface are torn down. A final Admin Down control packet is sent to all neighbors, and sessions on the remote system are placed in a Down state (Message 3 on page 179).

To disable BFD for a port-channel:

Step	Task	Command Syntax	Command Mode
1	Disable BFD for a port-channel.	**no bfd enable**	INTERFACE PORT-CHANNEL

Configuring Protocol Liveness

Protocol Liveness is a feature that notifies the BFD Manager when a client protocol is disabled. When a client is disabled, all BFD sessions for that protocol are torn down. Neighbors on the remote system receive an Admin Down control packet and are placed in the Down state (Message 3 on page 179).

To enable Protocol Liveness:

Step	Task	Command Syntax	Command Mode
1	Enable Protocol Liveness	**bfd protocol-liveness**	CONFIGURATION

Troubleshooting BFD

Examine control packet field values using the command **debug bfd detail**. Figure 9-29 shows a three-way handshake using this command.

Figure 9-29. debug bfd detail Command Output

```
R1(conf-if-gi-4/24)#00:54:38: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE: Changed session state
to Down for neighbor 2.2.2.2 on interface Gi 4/24 (diag: 0)
00:54:38 : Sent packet for session with neighbor 2.2.2.2 on Gi 4/24
TX packet dump:
    Version:1, Diag code:0, State:Down, Poll bit:0, Final bit:0, Demand bit:0
    myDiscrim:4, yourDiscrim:0, minTx:1000000, minRx:1000000, multiplier:3, minEchoRx:0
00:54:38 : Received packet for session with neighbor 2.2.2.2 on Gi 4/24
RX packet dump:
    Version:1, Diag code:0, State:Init, Poll bit:0, Final bit:0, Demand bit:0
    myDiscrim:6, yourDiscrim:4, minTx:1000000, minRx:1000000, multiplier:3, minEchoRx:0
00:54:38: %RPM0-P:RP2 %BFDMGR-1-BFD_STATE_CHANGE: Changed session state to Up for neighbor
2.2.2.2 on interface Gi 4/24 (diag: 0)
```

Examine control packets in hexadecimal format using the command **debug bfd packet**.

Figure 9-30. debug bfd packet Command Output

```
RX packet dump:
         20 c0 03 18 00 00 00 05 00 00 00 04 00 01 86 a0
         00 01 86 a0 00 00 00 00
00:34:13 : Sent packet for session with neighbor 2.2.2.2 on Gi 4/24
TX packet dump:
         20 c0 03 18 00 00 00 04 00 00 00 05 00 01 86 a0
         00 01 86 a0 00 00 00 00
00:34:14 : Received packet for session with neighbor 2.2.2.2 on Gi 4/24
RX packet dump:
         20 c0 03 18 00 00 00 05 00 00 00 04 00 01 86 a0
         00 01 86 a0 00 00 00 00
00:34:14 : Sent packet for session with neighbor 2.2.2.2 on Gi 4/24
TX packet dump:
```

The output for the command **debug bfd event** is the same as the log messages that appear on the console by default.

10

Border Gateway Protocol IPv4 (BGPv4)

Border Gateway Protocol IPv4 (BGPv4) version 4 (BGPv4) is supported on platforms: C E S

Platforms support BGP according to the following table:

FTOS version	Platform support	
8.1.1.0	E-Series ExaScale	E X
7.8.1.0	S-Series	S
7.7.1.0.	C-Series	C
pre-7.7.1.0	E-Series TeraScale	E T

This chapter is intended to provide a general description of Border Gateway Protocol version 4 (BGPv4) as it is supported in the Dell Force10 Operating System (FTOS).

This chapter includes the following topics:

BGP protocol standards are listed in the Appendix 63, Standards Compliance chapter.

Protocol Overview

Border Gateway Protocol (BGP) is an external gateway protocol that transmits interdomain routing information within and between Autonomous Systems (AS). Its primary function is to exchange network reachability information with other BGP systems. BGP generally operates with an Internal Gateway Protocol (IGP) such as OSPF or RIP, allowing you to communicate to external ASs smoothly. BGP adds reliability to network connections be having multiple paths from one router to another.

Autonomous Systems (AS)

BGP Autonomous Systems (ASs) are a collection of nodes under common administration, with common network routing policies. Each AS has a number, already assigned by an internet authority. You do not assign the BGP number.

AS Numbers (ASNs) are important because the ASN uniquely identifies each network on the Internet. The IANA has reserved AS numbers 64512 through 65534 to be used for private purposes. The ASNs 0 and 65535 are reserved by the IANA and should not be used in a live environment.

Autonomous Systems can be grouped into three categories, defined by their connections and operation.

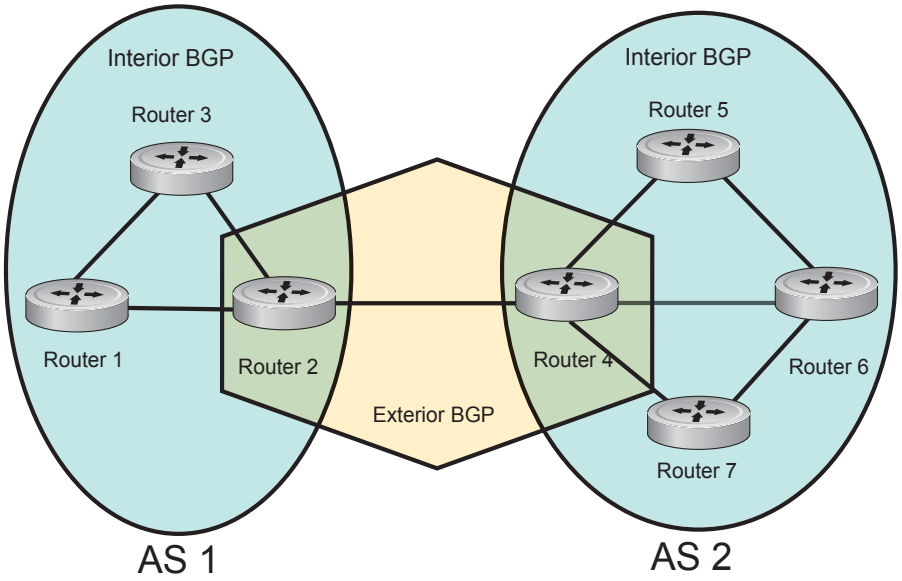
A **multihomed** AS is one that maintains connections to more than one other AS. This allows the AS to remain connected to the internet in the event of a complete failure of one of their connections. However, this type of AS does not allow traffic from one AS to pass through on its way to another AS. A simple example of this is seen in Figure 10-1.

A **stub** AS is one that is connected to only one other AS.

A **transit** AS is one that provides connections through itself to separate networks. For example as seen in Figure 10-1, Router 1 can use Router 2 (the transit AS) to connect to Router 4. ISPs are always transit ASs, because they provide connections from one network to another. The ISP is considered to be "selling transit service" to the customer network, so thus the term Transit AS.

When BGP operates inside an Autonomous System (AS1 or AS2 as seen in Figure 10-1), it is referred to as Internal BGP (IBGP *Interior Border Gateway Protocol*). When BGP operates between Autonomous Systems (AS1 and AS2), it is called External BGP (EBGP *Exterior Border Gateway Protocol*). IBGP provides routers inside the AS with the knowledge to reach routers external to the AS. EBGP routers exchange information with other EBGP routers as well as IBGP routers to maintain connectivity and accessibility.

Figure 10-1. BGP Autonomous Zones

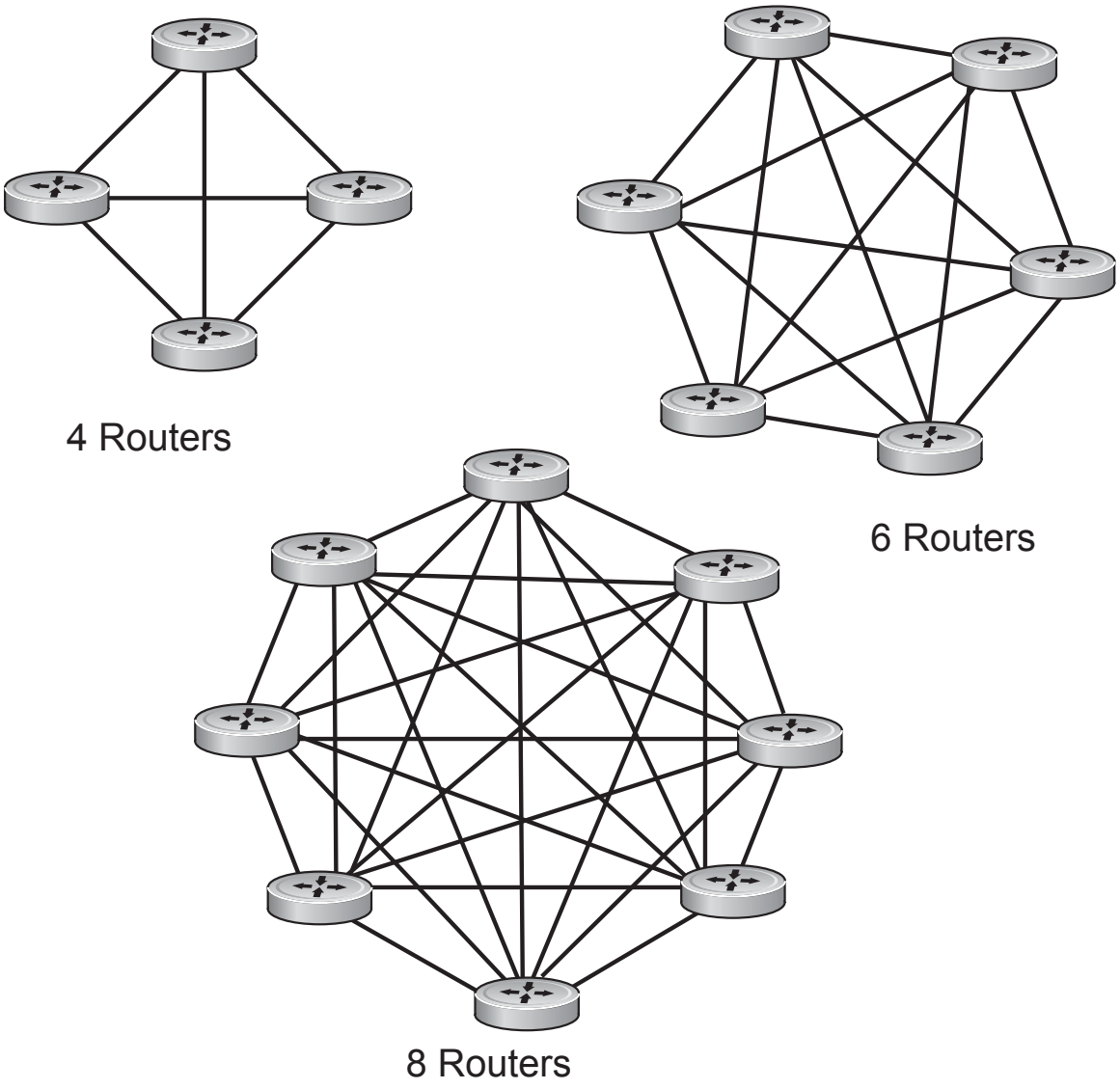


BGP version 4 (BGPv4) supports classless interdomain routing and aggregate routes and AS paths. BGP is a path vector protocol - a computer network in which BGP maintains the path that update information takes as it diffuses through the network. Updates traveling through the network and returning to the same node are easily detected and discarded.

BGP does not use traditional Interior Gateway Protocol (IGP) matrix, but makes routing decisions based on path, network policies and/or rulesets. Unlike most protocols, BGP uses TCP as its transport protocol.

Since each BGP routers talking to another router is a session, a BGP network needs to be in "full mesh". This is a topology that has every router directly connected to every other router. For example, as seen in Figure 10-2, four routers connected in a full mesh have three peers each, six routers have 5 peers each, and eight routers in full mesh will have seven peers each.

Figure 10-2. Full Mesh Examples



The number of BGP speakers each BGP peer must maintain increases exponentially. Network management quickly becomes impossible.

Sessions and Peers

When two routers communicate using the BGP protocol, a BGP session is started. The two end-points of that session are Peers. A Peer is also called a Neighbor.

Establishing a session

Information exchange between peers is driven by events and timers. The focus in BGP is on the traffic routing policies.

In order to make decisions in its operations with other BGP peers, a BGP peer uses a simple finite state machine that consists of six states: Idle, Connect, Active, OpenSent, OpenConfirm, and Established. For each peer-to-peer session, a BGP implementation tracks which of these six states the session is in. The BGP protocol defines the messages that each peer should exchange in order to change the session from one state to another.

The first state is the **Idle** mode. BGP initializes all resources, refuses all inbound BGP connection attempts, and initiates a TCP connection to the peer.

The next state is **Connect**. In this state the router waits for the TCP connection to complete, transitioning to the **OpenSent** state if successful.

If that transition is not successful, BGP resets the ConnectRetry timer and transitions to the **Active** state when the timer expires.

In the **Active** state, the router resets the ConnectRetry timer to zero, and returns to the **Connect** state.

Upon successful **OpenSent** transition, the router sends an Open message and waits for one in return.

Keepalive messages are exchanged next, and upon successful receipt, the router is placed in the **Established** state. Keepalive messages continue to be sent at regular periods (established by the Keepalive timer) to verify connections.

Once established, the router can now send/receive Keepalive, Update, and Notification messages to/from its peer.

Peer Groups

Peer Groups are neighbors grouped according to common routing policies. They enable easier system configuration and management by allowing groups of routers to share and inherit policies.

Peer groups also aid in convergence speed. When a BGP process needs to send the same information to a large number of peers, it needs to set up a long output queue to get that information to all the proper peers. If they are members of a peer group, however, the information can be sent to one place then passed onto the peers within the group.

Route Reflectors

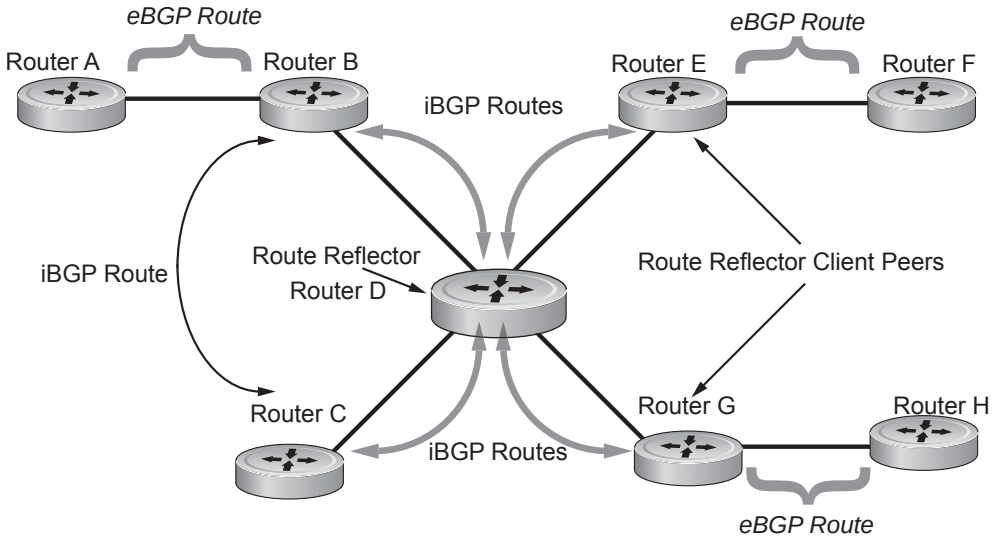
Route Reflectors reorganize the iBGP core into a hierarchy and allows some route advertisement rules.

Route reflection divides iBGP peers into two groups: client peers and nonclient peers. A route reflector and its client peers form a route reflection cluster. Since BGP speakers announce only the best route for a given prefix, route reflector rules are applied after the router makes its best path decision.

- If a route was received from a nonclient peer, reflect the route to all client peers.
- If the route was received from a client peer, reflect the route to all nonclient and all client peers.

To illustrate how these rules affect routing, see Figure 10-3 and the following steps.Routers B, C, D, E, and G are members of the same AS - AS100. These routers are also in the same Route Reflection Cluster, where Router D is the Route Reflector. Router E and H are client peers of Router D; Routers B and C and nonclient peers of Router D.

Figure 10-3. Route Reflection Example



1. Router B receives an advertisement from Router A through eBGP. Since the route is learned through eBGP, Router B advertises it to all its iBGP peers: Routers C and D.
2. Router C receives the advertisement but does not advertise it to any peer because its only other peer is Router D, an iBGP peer, and Router D has already learned it through iBGP from Router B.
3. Router D does not advertise the route to Router C because Router C is a nonclient peer and the route advertisement came from Router B who is also a non-client peer.
4. Router D does reflect the advertisement to Routers E and G because they are client peers of Router D.
5. Routers E and G then advertise this iBGP learned route to their eBGP peers Routers F and H.

Confederations

Communities

BGP communities are sets of routes with one or more common attributes. This is a way to assign common attributes to multiple routes at the same time.

BGP Attributes

Routes learned via BGP have associated properties that are used to determine the best route to a destination when multiple paths exist to a particular destination. These properties are referred to as BGP attributes, and an understanding of how BGP attributes influence route selection is required for the design of robust networks. This section describes the attributes that BGP uses in the route selection process:

- Weight
- Local Preference
- Multi-Exit Discriminators (MEDs)
- Origin
- AS Path
- Next Hop

Best Path Selection Criteria

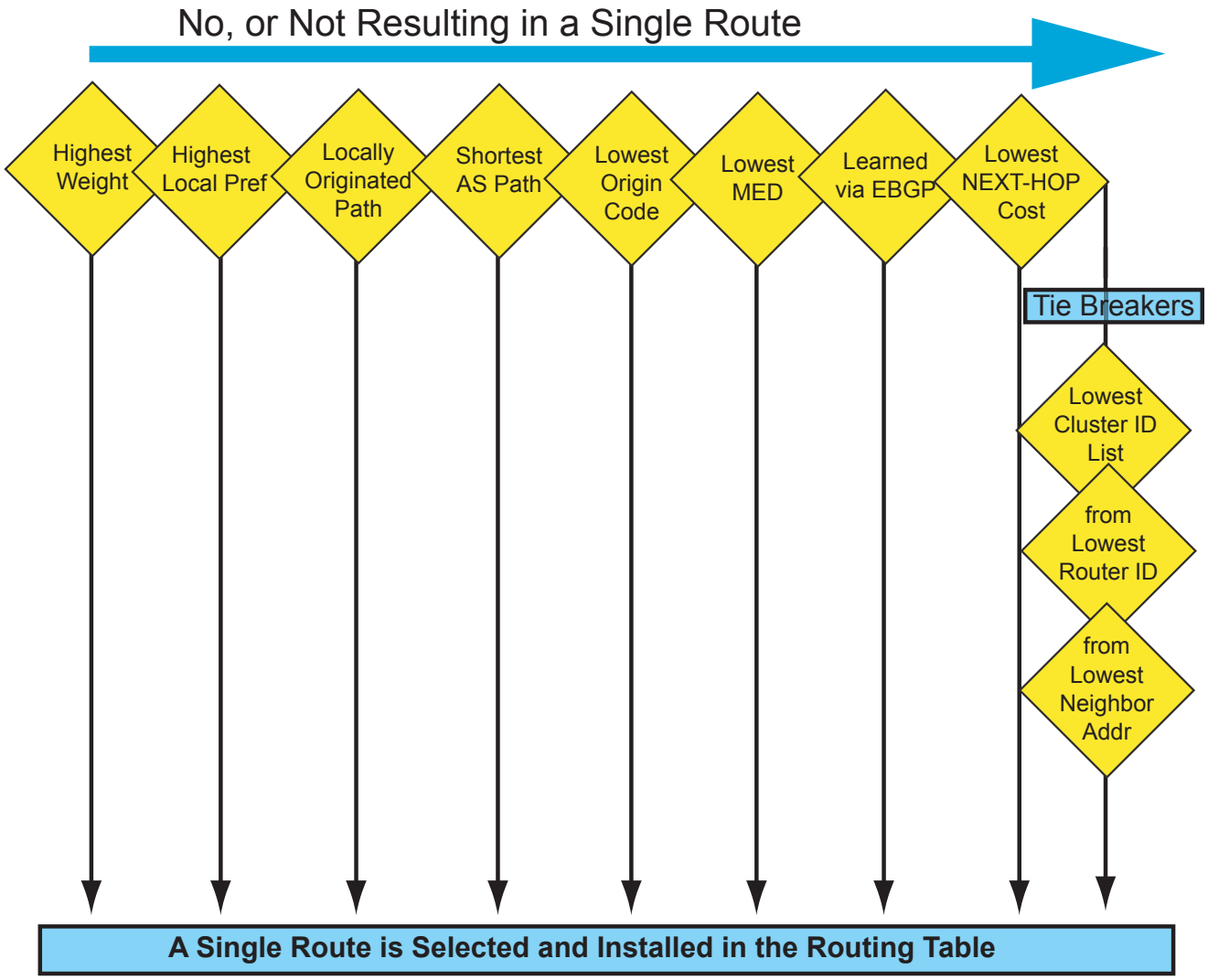
Paths for active routes are grouped in ascending order according to their neighboring external AS number (BGP best path selection is deterministic by default, which means the **bgp non-deterministic-med** command is NOT applied).

The best path in each group is selected based on specific criteria. Only one “best path” is selected at a time. If any of the criteria results in more than one path, BGP moves on to the next option in the list. For example, two paths may have the same weights, but different local preferences. BGP sees that the Weight criteria results in two potential “best paths” and moves to local preference to reduce the options. If a number of best paths is determined, this selection criteria is applied to group’s best to determine the ultimate best path.

In non-deterministic mode (the **bgp non-deterministic-med** command is applied), paths are compared in the order in which they arrive. This method can lead to FTOS choosing different best paths from a set of paths, depending on the order in which they were received from the neighbors, since MED may or may not get compared between adjacent paths. In deterministic mode, FTOS compares MED between adjacent paths within an AS group since all paths in the AS group are from the same AS.

Figure 10-4 illustrates the decisions BGP goes through to select the best path. The list following the illustration details the path selection criteria.

Figure 10-4. BGP Best Path Selection



Best Path selection details

1. Prefer the path with the largest WEIGHT attribute.
2. Prefer the path with the largest LOCAL_PREF attribute.
3. Prefer the path that was locally Originated via a **network** command, **redistribute** command or **aggregate-address** command.
   - Routes originated with the **network** or **redistribute** commands are preferred over routes originated with the **aggregate-address** command.
4. Prefer the path with the shortest AS_PATH (unless the **bgp bestpath as-path ignore** command is configured, then AS_PATH is not considered). The following criteria apply:
   - An AS_SET has a path length of 1, no matter how many ASs are in the set.
   - A path with no AS_PATH configured has a path length of 0.
   - AS_CONFED_SET is not included in the AS_PATH length.

- AS_CONFED_SEQUENCE has a path length of 1, no matter how many ASs are in the AS_CONFED_SEQUENCE.

5. Prefer the path with the lowest ORIGIN type (IGP is lower than EGP, and EGP is lower than INCOMPLETE).
6. Prefer the path with the lowest Multi-Exit Discriminator (MED) attribute. The following criteria apply:
   - This comparison is only done if the first (neighboring) AS is the same in the two paths; the MEDs are compared only if the first AS in the AS_SEQUENCE is the same for both paths.
   - If the **bgp always-compare-med** command is entered, MEDs are compared for all paths.
   - Paths with no MED are treated as "worst" and assigned a MED of 4294967295.
7. Prefer external (EBGP) to internal (IBGP) paths or confederation EBGP paths.
8. Prefer the path with the lowest IGP metric to the BGP next-hop.
9. FTOS deems the paths as equal and does not perform steps 9 through 11 listed below, if the following criteria is met:
   - the IBGP multipath or EBGP multipath are configured (**maximum-path** command)
   - the paths being compared were received from the same AS with the same number of ASs in the AS Path but with different NextHops
   - the paths were received from IBGP or EBGP neighbor respectively
10. If the **bgp bestpath router-id ignore** command is enabled and:
    - If the Router-ID is the same for multiple paths (because the routes were received from the same route) skip this step.
    - If the Router-ID is NOT the same for multiple paths, Prefer the path that was first received as the Best Path. The path selection algorithm should return without performing any of the checks outlined below.
11. Prefer the path originated from the BGP router with the lowest router ID. For paths containing a Route Reflector (RR) attribute, the originator ID is substituted for the router ID.
12. If two paths have the same router ID, prefer the path with the lowest cluster ID length. Paths without a cluster ID length are set to a 0 cluster ID length.
13. Prefer the path originated from the neighbor with the lowest address. (The neighbor address is used in the BGP neighbor configuration, and corresponds to the remote peer used in the TCP connection with the local router.)

After a number of best paths is determined, this selection criteria is applied to group's best to determine the ultimate best path.

In non-deterministic mode (the **bgp non-deterministic-med** command is applied), paths are compared in the order in which they arrive. This method can lead to FTOS choosing different best paths from a set of paths, depending on the order in which they were received from the neighbors since MED may or may not get compared between adjacent paths. In deterministic mode, FTOS compares MED between adjacent paths within an AS group since all paths in the AS group are from the same AS.

Weight

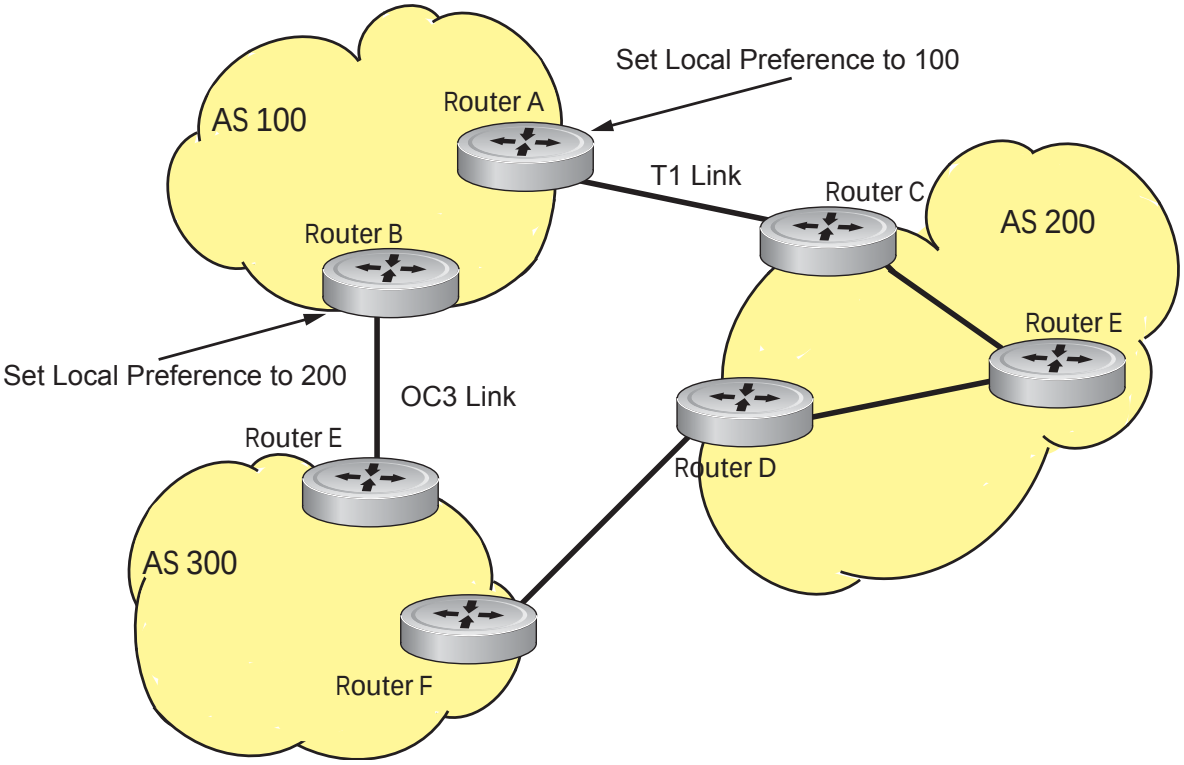
The Weight attribute is local to the router and is not advertised to neighboring routers. If the router learns about more than one route to the same destination, the route with the highest weight will be preferred. The route with the highest weight is installed in the IP routing table.

Local Preference

Local Preference (LOCAL_PREF) represents the degree of preference within the entire AS. The higher the number, the greater the preference for the route.

The Local Preference (LOCAL_PREF) is one of the criteria used to determine the best path, so keep in mind that other criteria may impact selection, as shown in Figure 10-4. For this example, assume that LOCAL_PREF is the only attribute applied. In Figure 10-5, AS100 has two possible paths to AS 200. Although the path through the Router A is shorter (one hop instead of two) the LOCAL_PREF settings have the preferred path go through Router B and AS300. This is advertised to all routers within AS100 causing all BGP speakers to prefer the path through Router B.

Figure 10-5. LOCAL_PREF Example



Multi-Exit Discriminators (MEDs)

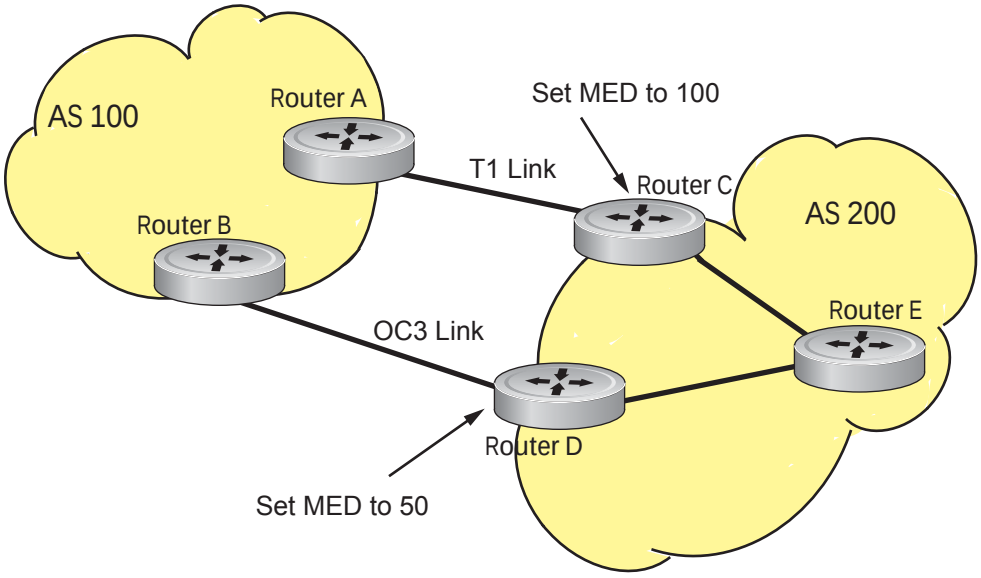
If two Autonomous Systems (AS) connect in more than one place, a Multi-Exit Discriminator (MED) can be used to assign a preference to a preferred path. The MED is one of the criteria used to determine the best path, so keep in mind that other criteria may impact selection, as shown in Figure 10-4.

One AS assigns the MED a value and the other AS uses that value to decide the preferred path. For this example, assume the MED is the only attribute applied. In Figure 10-6, AS100 and AS200 connect in two places. Each connection is a BGP session. AS200 sets the MED for its T1 exit point to 100 and the MED for its OC3 exit point to 50. This sets up a path preference through the OC3 link. The MEDs are advertised to AS100 routers so they know which is the preferred path.

An MED is a non-transitive attribute. If AS100 sends an MED to AS200, AS200 does not pass it on to AS300 or AS400. The MED is a locally relevant attribute to the two participating Autonomous Systems (AS100 and AS200).

Note that the MEDs are advertised across both links, so that if a link goes down AS 1 still has connectivity to AS300 and AS400.

Figure 10-6. MED Route Example



Note: With FTOS Release 8.3.1.0, configuring the **set metric-type internal** command in a route-map advertises the IGP cost as MED to outbound EBGP peers when redistributing routes. The configured **set metric** value overwrites the default IGP cost.

Origin

The Origin indicates the origin of the prefix, or how the prefix came into BGP. There are three Origin codes: IGP, EGP, INCOMPLETE.

- IGP indicated the prefix originated from information learned through an interior gateway protocol.
- EGP indicated the prefix originated from information learned from an EGP protocol, which NGP replaced.
- INCOMPLETE indicates that the prefix originated from an unknown source.

Generally, an IGP indicator means that the route was derived inside the originating AS. EGP generally means that a route was learned from an external gateway protocol. An INCOMPLETE origin code generally results from aggregation, redistribution or other indirect ways of installing routes into BGP.

In FTOS, these origin codes appear as shown in Figure 10-7. The question mark (?) indicates an Origin code of INCOMPLETE. The lower case letter (i) indicates an Origin code of IGP.

Figure 10-7. Origin attribute reported

```
FTOS#show ip bgp
BGP table version is 0, local router ID is 10.101.15.13
Status codes: s suppressed, d damped, h history, * valid, > best
Path source: I - internal, a - aggregate, c - confed-external, r - redistributed, n - network
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric     LocPrf Weight Path
*>  7.0.0.0/29       10.114.8.33              0          0 18508  ?
*>  7.0.0.0/30       10.114.8.33              0          0 18508  ?
*>  9.2.0.0/16       10.114.8.33             10          0 18508  701 i
```

AS Path

The AS Path is the list of all Autonomous Systems that all the prefixes listed in the update have passed through. The local AS number is added by the BGP speaker when advertising to a eBGP neighbor.

In FTOS the AS Path is shown in Figure 10-8. Note that the Origin attribute is shown following the AS Path information.

Figure 10-8. AS Path attribute reported

```
FTOS#show ip bgp paths
Total 30655 Paths
Address       Hash Refcount Metric Path
0x4014154        0         3 18508  701 3549 19421 i
0x4013914        0         3 18508  701 7018 14990 i
0x5166d6c        0         3 18508  209 4637 1221 9249 9249 i
0x5e62df4        0         2 18508  701 17302 i
0x3a1814c        0        26 18508  209 22291 i
0x567ea9c        0        75 18508  209 3356 2529 i
0x6cc1294        0         2 18508  209 1239 19265 i
0x6cc18d4        0         1 18508  701 2914 4713 17935 i
0x5982e44        0       162 18508  209 i
0x67d4a14        0         2 18508  701 19878 ?
0x559972c        0        31 18508  209 18756 i
0x59cd3b4        0         2 18508  209 7018 15227 i
0x7128114        0        10 18508  209 3356 13845 i
0x536a914        0         3 18508  209 701 6347 7781 i
0x2ffe884        0         1 18508  701 3561 9116 21350 i
```

Next Hop

The Next Hop is the IP address used to reach the advertising router. For EBGP neighbors, the Next-Hop address is the IP address of the connection between the neighbors. For IBGP, the EBGP Next-Hop address is carried into the local AS. A Next Hop attribute is set when a BGP speaker advertises itself to another BGP speaker outside its local AS. It can also be set when advertising routes within an AS. The Next Hop attribute also serves as a way to direct traffic to another BGP speaker, rather than waiting for a speaker to advertise.

FTOS allows you to set the Next Hop attribute in the CLI. Setting the Next Hop attribute lets you determine a router as the next hop for a BGP neighbor.

Multiprotocol BGP

MBGP for IPv6 unicast is supported on platforms E C

MBGP for IPv4 Multicast is supported on platform C E S

Multiprotocol Extensions for BGP (MBGP) is defined in IETF RFC 2858. MBGP allows different types of address families to be distributed in parallel. This allows information about the topology of IP Multicast-capable routers to be exchanged separately from the topology of normal IPv4 and IPv6 unicast routers. It allows a multicast routing topology different from the unicast routing topology.

Note: It is possible to configure BGP peers that exchange both unicast and multicast network layer reachability information (NLRI), but you cannot connect Multiprotocol BGP with BGP. Therefor, You cannot redistribute Multiprotocol BGP routes into BGP.

Implementing BGP with FTOS

Advertise IGP cost as MED for redistributed routes

When using multipath connectivity to an external AS, you can advertise the MED value selectively to each peer for redistributed routes. For some peers you can set the internal/IGP cost as the MED while setting others to a constant pre-defined metric as MED value.

FTOS 8.3.1.0 and later support configuring the **set metric-type internal** command in a route-map to advertise the IGP cost as the MED to outbound EBGP peers when redistributing routes. The configured **set metric** value overwrites the default IGP cost.

By using the **redistribute** command in conjunction with the **route-map** command, you can specify whether a peer advertises the standard MED or uses the IGP cost as the MED.

Note the following when configuring this functionality:

- If the **redistribute** command does not have any **metric** configured and BGP Peer out-bound route-map does have **metric-type internal** configured, BGP advertises the IGP cost as MED.
- If the **redistribute** command has **metric** configured (**route-map set metric** or **redistribute** *route-type* **metric)** and the BGP Peer out-bound route-map has **metric-type internal** configured, BGP advertises the metric configured in the redistribute command as MED.
- If BGP peer out-bound route-map has **metric** configured, then all other metrics are overwritten by this.

Note: When redistributing static, connected or OSPF routes, there is no metric option. Simply assign the appropriate route-map to the redistributed route.

Table 10-1gives some examples of these rules.

Table 10-1. Example MED advertisement

Command Settings	BGP Local Routing Information Base	MED Advertised to Peer	
		WITH route-map metric-type internal	WITHOUT route-map metric-type internal
redistribute *isis* (IGP cost = 20)	MED: IGP cost 20	MED = 20	MED = 0
redistribute *isis* route-map set metric 50	MED: IGP cost 50	MED: 50	MED: 50
redistribute *isis* metric 100	MED: IGP cost 100	MED: 100	MED: 100

Ignore Router-ID for some best-path calculations

FTOS 8.3.1.0 and later allow you to avoid unnecessary BGP best-path transitions between external paths under certain conditions. The **bgp bestpath router-id ignore** command reduces network disruption caused by routing and forwarding plane changes and allows for faster convergence.

4-Byte AS Numbers

FTOS Version 7.7.1 and later support 4-Byte (32-bit) format when configuring Autonomous System Numbers (ASNs). The 4-Byte support is advertised as a new BGP capability (4-BYTE-AS) in the OPEN message. If a 4-Byte BGP speaker has sent and received this capability from another speaker, all the messages will be 4-octet. The behavior of a 4-Byte BGP speaker will be different with the peer depending on whether the peer is 4-Byte or 2-Byte BGP speaker.

Where the 2-Byte format is 1-65535, the 4-Byte format is 1-4294967295. Enter AS Numbers using the traditional format. If the ASN is greater than 65535, the dot format is shown when using the **show ip bgp** commands. For example, an ASN entered as 3183856184 will appear in the show commands as 48581.51768; an ASN of 65123 is shown as 65123. To calculate the comparable dot format for an ASN from a traditional format, use **ASN/65536. ASN%65536.**

Table 10-2. 4-Byte ASN Dot Format Examples

Traditional Format		Dot Format
65001	**Is**	0.65501
65536	**The**	1.0
100000	**Same As**	1.34464
4294967295		65535.65535

When creating Confederations, all the routers in a Confederation must be either 4-Byte or 2-Byte identified routers. You cannot mix them.

Configure the 4-byte AS numbers with the **four-octect-support** command.

AS4 Number Representation

FTOS version 8.2.1.0 supports multiple representations of an 4-byte AS Numbers: **asplain**, **asdot+**, and **asdot**.

Note: The ASDOT and ASDOT+ representations are supported only in conjunction with the 4-Byte AS Numbers feature. If 4-Byte AS Numbers are not implemented, only ASPLAIN representation is supported.

ASPLAIN is the method FTOS has used for all previous FTOS versions.It remains the default method with FTOS 8.2.1.0 and later. With the ASPLAIN notation, a 32 bit binary AS number is translated into a decimal value.

- All AS Numbers between 0-65535 are represented as a decimal number when entered in the CLI as well as when displayed in the show command outputs.
- AS Numbers larger than 65535 are represented using ASPLAIN notation as well. 65546 is represented as 65546.

ASDOT+ representation splits the full binary 4-byte AS number into two words of 16 bits separated by a decimal point (.): <high-order 16 bit value>.<low-order 16 bit value>. Some examples are shown in Table 10-2.

- All AS Numbers between 0-65535 are represented as a decimal number, when entered in the CLI as well as when displayed in the show command outputs.
- AS Numbers larger than 65535 is represented using ASDOT notation as <higher 2 bytes in decimal>.<lower 2 bytes in decimal>. For example: AS 65546 is represented as 1.10.

ASDOT representation combines the ASPLAIN and ASDOT+ representations. AS Numbers less than 65536 appear in integer format (asplain); AS Numbers equal to or greater than 65536 appear using the decimal method (asdot+). For example, the AS Number 65526 appears as 65526, and the AS Number 65546 appears as 1.10.

Dynamic AS Number Notation application

FTOS 8.3.1.0 applies the ASN Notation type change dynamically to the running-config statements. When you apply or change an asnotation, the type selected is reflected immediately in the running-configuration and the show commands (Figure 10-9 and Figure 10-10).

Figure 10-9. Dynamic changes of the bgp asnotation command in the show running config

ASDOT

```
FTOS(conf-router_bgp)#bgp asnotation asdot
FTOS(conf-router_bgp)#show conf
!
router bgp 100
 bgp asnotation asdot
 bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>

FTOS(conf-router_bgp)#do show ip bgp
BGP table version is 24901, local router ID is 172.30.1.57
<output truncated>
```

ASDOT+

```
FTOS(conf-router_bgp)#bgp asnotation asdot+
FTOS(conf-router_bgp)#show conf
!
router bgp 100
 bgp asnotation asdot+
 bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>

FTOS(conf-router_bgp)#do show ip bgp
BGP table version is 31571, local router ID is 172.30.1.57
<output truncated>
```

AS-PLAIN

```
FTOS(conf-router_bgp)#bgp asnotation asplain
FTOS(conf-router_bgp)#sho conf
!
router bgp 100
 bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>

FTOS(conf-router_bgp)#do sho ip bgp
BGP table version is 34558, local router ID is 172.30.1.57
<output truncated>
```

Figure 10-10. Dynamic changes when bgp asnotation command is disabled in the show running config

AS NOTATION DISABLED

```
FTOS(conf-router_bgp)#no bgp asnotation
FTOS(conf-router_bgp)#sho conf
!
router bgp 100
 bgp four-octet-as-support
neighbor 172.30.1.250 local-as 65057
<output truncated>

FTOS(conf-router_bgp)#do sho ip bgp
BGP table version is 28093, local router ID is 172.30.1.57
```

AS4 SUPPORT DISABLED

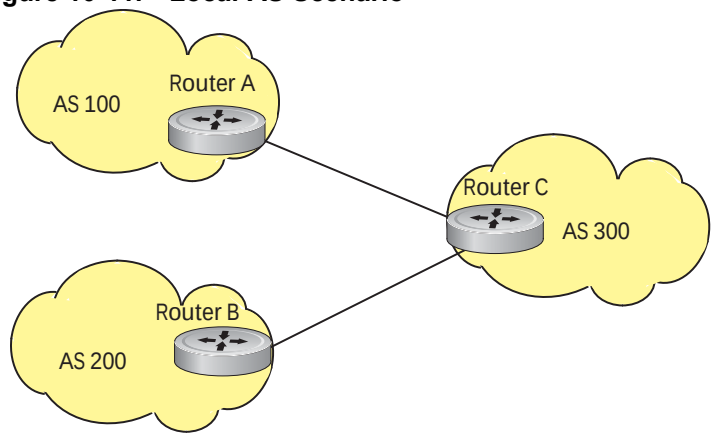
```
FTOS(conf-router_bgp)#no bgp four-octet-as-support
FTOS(conf-router_bgp)#sho conf
!
router bgp 100
neighbor 172.30.1.250 local-as 65057

FTOS(conf-router_bgp)#do show ip bgp
BGP table version is 28093, local router ID is 172.30.1.57
```

AS Number Migration

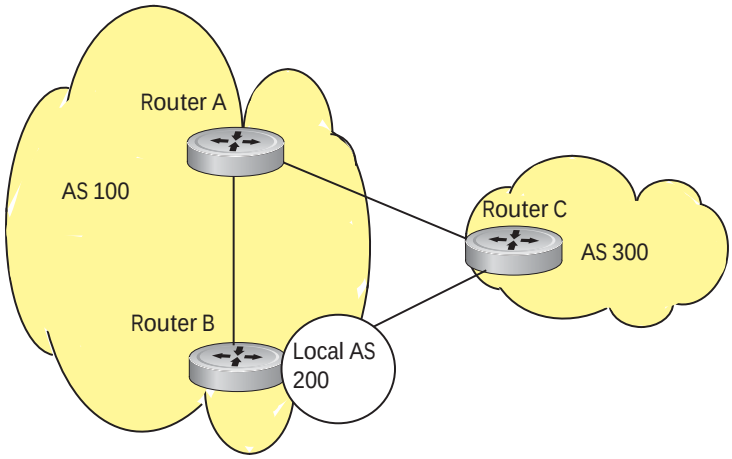
When migrating one AS to another, perhaps combining ASs, an eBGP network may lose its routing to an iBGP if the ASN changes. Migration can be difficult as all the iBGP and eBGP peers of the migrating network need to be updated to maintain network reachability. With this feature you can transparently change the AS number of entire BGP network and ensure that the routes are propagated throughout the network while the migration is in progress. Essentially, **Local-AS** provides a capability to the BGP speaker to operate as if it belongs to “virtual” AS network besides its physical AS network.

Figure 10-11 shows a scenario where Router A, Router B and Router C belong to AS 100, 200, 300 respectively. Router A acquired Router B; Router B has Router C as its customer. When Router B is migrating to Router A, it needs to maintain the connection with Router C without immediately updating Router C's configuration. **Local-AS** allows this to happen by allowing Router B to appear as if it still belongs to Router B's old network (AS 200) as far as communicating with Router C is concerned.

Figure 10-11. Local-AS Scenario



Before Migration



After Migration, with Local-AS enabled

When you complete your migration, and you have reconfigured your network with the new information you must disable this feature.

If the “no prepend” option is used, the local-as will not be prepended to the updates received from the eBGP peer. If “no prepend” is not selected (the default), the local-as is added to the first AS segment in the AS-PATH. If an inbound route-map is used to prepend the as-path to the update from the peer, the local-as is added first. For example, consider the topology described in Figure 10-11. If Router B has an inbound route-map applied on Router C to prepend "65001 65002" to the as-path, the following events will take place on Router B

1. Receive and validate the update
2. Prepend local-as 200 to as-path
3. Prepend "65001 65002" to as-path

Local-as is prepended before the route-map to give an impression that update passed thru a router in AS 200 before it reached Router B.

BGP4 Management Information Base (MIB)

The FORCE10-BGP4-V2-MIB enhances FTOS BGP Management Information Base (MIB) support with many new SNMP objects and notifications (traps) defined in the *draft-ietf-idr-bgp4-mibv2-05*. To see these enhancements, download the MIB from the Dell Force10 website, www.force10networks.com.

Note: See the Dell Force10 iSupport webpage for the *Force10-BGP4-V2-MIB* and other MIB documentation.

Important Points to Remember

- In f10BgpM2AsPathTableEntry table, f10BgpM2AsPathSegmentIndex, and f10BgpM2AsPathElementIndex are used to retrieve a particular ASN from the AS path. These indices are assigned to the AS segments and individual ASN in each segment starting from 0. For example, an AS path list of {200 300 400} 500 consists of two segments: {200 300 400} with segment index 0 and 500 with segment index 1. ASN 200, 300, and 400 are be assigned 0, 1, and 2 element indices in that order.
- Unknown optional transitive attributes within a given path attribute (PA) are assigned indices in order. These indices correspond to f10BgpM2PathAttrUnknownIndex field in the f10BgpM2PathAttrUnknownEntry table.
- Negotiation of multiple instances of the same capability is not supported. F10BgpM2PeerCapAnnouncedIndex and f10BgpM2PeerCapReceivedIndex are ignored in the peer capability lookup.
- Inbound BGP soft-reconfiguration must be configured on a peer for f10BgpM2PrefixInPrefixesRejected to display the number of prefixes filtered due to a policy. If BGP soft-reconfig is not enabled, the denied prefixes are not accounted for.
- F10BgpM2AdjRibsOutRoute stores the pointer to the NLRI in the peer's Adj-Rib-Out.
- PA Index (f10BgpM2PathAttrIndex field in various tables) is used to retrieve specific attributes from the PA table. The Next-Hop, RR Cluster-list, Originator ID attributes are not stored in the PA Table and cannot be retrieved using the index passed in. These fields are not populated in f10BgpM2PathAttrEntry, f10BgpM2PathAttrClusterEntry, f10BgpM2PathAttrOriginatorIdEntry.
- F10BgpM2PathAttrUnknownEntry contains the optional-transitive attribute details.
- Query for f10BgpM2LinkLocalNextHopEntry returns default value for Link-local Next-hop.
- RFC 2545 and the f10BgpM2Rfc2545Group are not supported.
- An SNMP query will display up to 89 AS paths. A query for a larger AS path count will display as "…" at the end of the output.
- SNMP set for BGP is not supported. For all peer configuration tables (f10BgpM2PeerConfigurationGroup, f10BgpM2PeerRouteReflectorCfgGroup, and f10BgpM2PeerAsConfederationCfgGroup), an SNMP set operation will return an error. Only SNMP queries are supported. In addition, the f10BgpM2CfgPeerError, f10BgpM2CfgPeerBgpPeerEntry, and f10BgpM2CfgPeerRowEntryStatus fields are to hold the SNMP set status and are ignored in SNMP query.

- The AFI/SAFI is not used as an index to the f10BgpM2PeerCountersEntry table. The BGP peer's AFI/SAFI (IPv4 Unicast or IPv6 Multicast) is used for various outbound counters. Counters corresponding to IPv4 Multicast cannot be queried.
- The f10BgpM2[Cfg]PeerReflectorClient field is populated based on the assumption that route-reflector clients are not in a full mesh if BGP client-2-client reflection is enabled and that the BGP speaker acting as reflector will advertise routes learned from one client to another client. If disabled, it is assumed that clients are in a full mesh, and there is no need to advertise prefixes to the other clients.
- High CPU utilization may be observed during an SNMP walk of a large BGP Loc-RIB.
- To avoid SNMP timeouts with a large-scale configuration (large number of BGP neighbors and a large BGP Loc-RIB), Dell Force10 recommends setting the timeout and retry count values to a relatively higher number. e.g. t = 60 or r = 5.
- To return all values on an snmpwalk for the f10BgpM2Peer sub-OID, use the -C c option, such as snmpwalk -v 2c -C c -c public <IP_address> <OID>.
- An SNMP walk may terminate pre-maturely if the index does not increment lexicographically. Dell Force10 recommends using options to ignore such errors.
- Multiple BPG process instances are not supported. Thus, the F10BgpM2PeerInstance field in various tables is not used to locate a peer.
- Multiple instances of the same NLRI in the BGP RIB are not supported and are set to zero in the SNMP query response.
- F10BgpM2NlriIndex and f10BgpM2AdjRibsOutIndex fields are not used.
- Carrying MPLS labels in BGP is not supported. F10BgpM2NlriOpaqueType and f10BgpM2NlriOpaquePointer fields are set to zero.
- 4-byte ASN is supported. f10BgpM2AsPath4byteEntry table contains 4-byte ASN-related parameters based on the configuration.

Traps (notifications) specified in the BGP4 MIB draft <draft-ietf-idr-bgp4-mibv2-05.txt> are not supported. Such traps (bgpM2Established and bgpM2BackwardTransition) are supported as part of RFC 1657.

Configuration Information

The software supports BGPv4 as well as the following:

- deterministic multi-exit discriminator (MED) (default)
- a path with a missing MED is treated as worst path and assigned an MED value of `(0xffffffff)`
- the community format follows RFC 1998
- delayed configuration (the software at system boot reads the entire configuration file prior to sending messages to start BGP peer sessions)

The following are not yet supported:

- auto-summarization (the default is no auto-summary)
- synchronization (the default is no synchronization)

BGP Configuration

To enable the BGP process and begin exchanging information, you must assign an AS number and use commands in the ROUTER BGP mode to configure a BGP neighbor.

Defaults

By default, BGP is disabled.

By default, FTOS compares the MED attribute on different paths from within the same AS (the **bgp always-compare-med** command is not enabled).

Note: In FTOS, all newly configured neighbors and peer groups are disabled. You must enter the **neighbor** {*ip-address* | *peer-group-name*} **no shutdown** command to enable a neighbor or peer group.

Table 10-3 displays the default values for BGP on FTOS.

Table 10-3. FTOS BGP Defaults

Item	Default
BGP Neighbor Adjacency changes	All BGP neighbor changes are logged.
Fast External Fallover feature	Enabled
graceful restart feature	Disabled
Local preference	100
MED	0
Route Flap Damping Parameters	half-life = 15 minutes reuse = 750 suppress = 2000 max-suppress-time = 60 minutes
Distance	external distance = 20 internal distance = 200 local distance = 200
Timers	keepalive = 60 seconds holdtime = 180 seconds

Configuration Task List for BGP

The following list includes the configuration tasks for BGP:

- Enable BGP
- Configure AS4 Number Representations
- Configure Peer Groups
- BGP fast fall-over

- Configure passive peering
- Maintain existing AS numbers during an AS migration
- Allow an AS number to appear in its own AS path
- Enable graceful restart
- Filter on an AS-Path attribute
- Configure IP community lists
- Manipulate the COMMUNITY attribute
- Change MED attribute
- Change LOCAL_PREFERENCE attribute
- Change NEXT_HOP attribute
- Change WEIGHT attribute
- Enable multipath
- Filter BGP routes
- Redistribute routes on page 246
- Configure BGP route reflectors
- Aggregate routes
- Configure BGP confederations
- Enable route flap dampening
- Change BGP timers
- BGP neighbor soft-reconfiguration
- Route map continue

Enable BGP

By default, BGP is not enabled on the system. FTOS supports one Autonomous System (AS) and you must assign the AS Number (ASN). To establish BGP sessions and route traffic, you must configure at least one BGP neighbor or peer.

In BGP, routers with an established TCP connection are called neighbors or peers. Once a connection is established, the neighbors exchange full BGP routing tables with incremental updates afterwards. In addition, neighbors exchange KEEPALIVE messages to maintain the connection.

In BGP, neighbor routers or peers can be classified as internal or external. External BGP peers must be connected physically to one another (unless you enable the EBGP multihop feature), while internal BGP peers do not need to be directly connected. The IP address of an EBGP neighbor is usually the IP address of the interface directly connected to the router. First, the BGP process determines if all internal BGP peers are reachable, and then it determines which peers outside the AS are reachable.

Note: Sample Configurations for enabling BGP routers are found at the end of this chapter.

Use these commands in the following sequence, starting in the CONFIGURATION mode to establish BGP sessions on the router.

Step	Command Syntax	Command Mode	Purpose
1	**router bgp** *as-number*	CONFIGURATION	Assign an AS number and enter the ROUTER BGP mode. AS Number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte) or 0.1-65535.65535 (Dotted format) Only one AS is supported per system If you enter a 4-Byte AS Number, 4-Byte AS Support is enabled automatically.
1a	**bgp four-octet-as-support**	CONFIG-ROUTER-BGP	Enable 4-Byte support for the BGP process. **Note:** This is an OPTIONAL command. Enable if you want to use 4-Byte AS numbers or if you support AS4 Number Representation. Use it only if you support 4-Byte AS Numbers or if you support AS4 Number Representation. If you are supporting 4-Byte ASNs, this command must be enabled first. Disable 4-Byte support and return to the default 2-Byte format by using the **no bgp four-octet-as-support** command. You cannot disable 4-Byte support if you currently have a 4-Byte ASN configured. Disabling 4-Byte AS Numbers also disables ASDOT and ASDOT+ number representation. All AS Numbers will be displayed in ASPLAIN format.
1b	**address-family [ipv4 | ipv6}**	CONFIG-ROUTER-BGP	Enable IPv4 multicast or IPv6 mode. Use this command to enter BGP for IPv6 mode (CONF-ROUTER_BGPv6_AF).
2	**neighbor** {*ip-address* | *peer-group name*} **remote-as** *as-number*	CONFIG-ROUTER-BGP	Add a neighbor as a remote AS. Formats: IP Address A.B.C.D Peer-Group Name: 16 characters AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte) or 0.1-65535.65535 (Dotted format) You must Configure Peer Groups *before* assigning it a remote AS.
3	**neighbor** {*ip-address* | *peer-group-name*} **no shutdown**	CONFIG-ROUTER-BGP	Enable the BGP neighbor.

Note: When you change the configuration of a BGP neighbor, always reset it by entering the **clear ip bgp** command in EXEC Privilege mode.

Enter **show config** in CONFIGURATION ROUTER BGP mode to view the BGP configuration. Use the **show ip bgp summary** command in EXEC Privilege mode to view the BGP status. Figure 10-12 shows the summary with a 2-Byte AS Number displayed; Figure 10-13 shows the summary with a 4-Byte AS Number displayed.

Figure 10-12. Command example: show ip bgp summary (2-Byte AS Number displayed)

```
R2#show ip bgp summary
BGP router identifier 192.168.10.2, local AS number 65123
BGP table version is 1, main routing table version 1
1 network entrie(s) using 132 bytes of memory
1 paths using 72 bytes of memory
BGP-RIB over all using 73 bytes of memory
1 BGP path attribute entrie(s) using 72 bytes of memory
1 BGP AS-PATH entrie(s) using 47 bytes of memory
5 neighbor(s) using 23520 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx

10.10.21.1      65123               0        0          0    0     0 never    Active
10.10.32.3      65123               0        0          0    0     0 never    Active
100.10.92.9     65192               0        0          0    0     0 never    Active
192.168.10.1    65123               0        0          0    0     0 never    Active
192.168.12.2    65123               0        0          0    0     0 never    Active
R2#
```

2-Byte AS Number

Figure 10-13. Command example: show ip bgp summary (4-Byte AS Number displayed)

```
R2#show ip bgp summary
BGP router identifier 192.168.10.2, local AS number 48735.59224
BGP table version is 1, main routing table version 1
1 network entrie(s) using 132 bytes of memory
1 paths using 72 bytes of memory
BGP-RIB over all using 73 bytes of memory
1 BGP path attribute entrie(s) using 72 bytes of memory
1 BGP AS-PATH entrie(s) using 47 bytes of memory
5 neighbor(s) using 23520 bytes of memory

Neighbor        AS            MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx

10.10.21.1      65123               0        0          0    0     0 never    Active
10.10.32.3      65123               0        0          0    0     0 never    Active
100.10.92.9     65192               0        0          0    0     0 never    Active
192.168.10.1    65123               0        0          0    0     0 never    Active
192.168.12.2    65123               0        0          0    0     0 never    Active
R2#
```

4-Byte AS Number

For the router's identifier, FTOS uses the highest IP address of the Loopback interfaces configured. Since Loopback interfaces are virtual, they cannot go down, thus preventing changes in the router ID. If no Loopback interfaces are configured, the highest IP address of any interface is used as the router ID.

To view the status of BGP neighbors, use the **show ip bgp neighbors** (Figure 10-14) command in EXEC Privilege mode. For BGP neighbor configuration information, use the **show running-config bgp** command in EXEC Privilege mode (Figure 10-15). Note that the **showconfig** command in CONFIGURATION ROUTER BGP mode gives the same information as thew **show running-config bgp**.

Figure 10-14 displays two neighbors, one is an external and the second one is an internal BGP neighbor. The first line of the output for each neighbor displays the AS number and states whether the link is an external or internal.

The third line of the **show ip bgp neighbors** output contains the BGP State. If anything other than ESTABLISHED is listed, the neighbor is not exchanging information and routes. For more details on using the **show ip bgp neighbors** command, refer to the *FTOS Command Line Interface Reference*.

Figure 10-14. Command example: show ip bgp neighbors

```
FTOS#show ip bgp neighbors

BGP neighbor is 10.114.8.60, remote AS 18508, external link
  BGP version 4, remote router ID 10.20.20.20
  BGP state ESTABLISHED, in this state for 00:01:58
  Last read 00:00:14, hold time is 90, keepalive interval is 30 seconds
  Received 18552 messages, 0 notifications, 0 in queue
  Sent 11568 messages, 0 notifications, 0 in queue
  Received 18549 updates, Sent 11562 updates
  Minimum time between advertisement runs is 30 seconds


  For address family: IPv4 Unicast
  BGP table version 216613, neighbor version 201190
  130195 accepted prefixes consume 520780 bytes
  Prefix advertised 49304, rejected 0, withdrawn 36143

  Connections established 1; dropped 0
  Last reset never
Local host: 10.114.8.39, Local port: 1037
Foreign host: 10.114.8.60, Foreign port: 179


BGP neighbor is 10.1.1.1, remote AS 65535, internal link
  Administratively shut down
  BGP version 4, remote router ID 10.0.0.0
  BGP state IDLE, in this state for 17:12:40
  Last read 17:12:40, hold time is 180, keepalive interval is 60 seconds
  Received 0 messages, 0 notifications, 0 in queue
  Sent 0 messages, 0 notifications, 0 in queue
  Received 0 updates, Sent 0 updates
  Minimum time between advertisement runs is 5 seconds


  For address family: IPv4 Unicast
  BGP table version 0, neighbor version 0
  0 accepted prefixes consume 0 bytes
  Prefix advertised 0, rejected 0, withdrawn 0

  Connections established 0; dropped 0
  Last reset never
  No active TCP connection
FTOS#
```

External BGP neighbor

Internal BGP neighbor

Figure 10-15. Command example: show running-config bgp

```
R2#show running-config bgp
!
router bgp 65123
 bgp router-id 192.168.10.2
 network 10.10.21.0/24
 network 10.10.32.0/24
 network 100.10.92.0/24
 network 192.168.10.0/24
 bgp four-octet-as-support
 neighbor 10.10.21.1 remote-as 65123
 neighbor 10.10.21.1 filter-list ISP1in
 neighbor 10.10.21.1 no shutdown
 neighbor 10.10.32.3 remote-as 65123
 neighbor 10.10.32.3 no shutdown
 neighbor 100.10.92.9 remote-as 65192
 neighbor 100.10.92.9 no shutdown
 neighbor 192.168.10.1 remote-as 65123
 neighbor 192.168.10.1 update-source Loopback 0
 neighbor 192.168.10.1 no shutdown
 neighbor 192.168.12.2 remote-as 65123
 neighbor 192.168.12.2 update-source Loopback 0
 neighbor 192.168.12.2 no shutdown
R2#
```

Configure AS4 Number Representations

Enable one type of AS Number Representation: ASPLAIN, ASDOT+, or ASDOT.

- ASPLAIN is the method FTOS has used for all previous FTOS versions. It remains the default method with FTOS 8.2.1.0 and later. With the ASPLAIN notation, a 32 bit binary AS number is translated into a decimal value.
- ASDOT+ representation splits the full binary 4-byte AS number into two words of 16 bits separated by a decimal point (.): <high-order 16 bit value>.<low-order 16 bit value>.
- ASDOT representation combines the ASPLAIN and ASDOT+ representations. AS Numbers less than 65536 appear in integer format (asplain); AS Numbers equal to or greater than 65536 appear using the decimal method (asdot+). For example, the AS Number 65526 appears as 65526, and the AS Number 65546 appears as 1.10.

Note: The ASDOT and ASDOT+ representations are supported only in conjunction with the 4-Byte AS Numbers feature. If 4-Byte AS Numbers are not implemented, only ASPLAIN representation is supported.

Only one form of AS Number Representation is supported at a time. You cannot combine the types of representations within an AS.

Task	Command Syntax	Command Mode
Enable ASPLAIN AS Number representation. Figure 10-16	**bgp asnotation asplain**	CONFIG-ROUTER-BGP
	Note: ASPLAIN is the default method FTOS uses and does not appear in the configuration display.	

Task	Command Syntax	Command Mode
Enable ASDOT AS Number representation. Figure 10-17	**bgp asnotation asdot**	CONFIG-ROUTER-BGP
Enable ASDOT+ AS Number representation.Figure 10-18	**bgp asnotation asdot+**	CONFIG-ROUTER-BGP

Figure 10-16. Command example and output: bgp asnotation asplain

```
FTOS(conf-router_bgp)#bgp asnotation asplain
FTOS(conf-router_bgp)#sho conf
!
router bgp 100
 bgp four-octet-as-support
 neighbor 172.30.1.250 remote-as 18508
 neighbor 172.30.1.250 local-as 65057
 neighbor 172.30.1.250 route-map rmap1 in
 neighbor 172.30.1.250 password 7
5ab3eb9a15ed02ff4f0dfd4500d6017873cfd9a267c04957
 neighbor 172.30.1.250 no shutdown
5332332 9911991 65057 18508 12182 7018 46164 i
```

Figure 10-17. Command example and output: bgp asnotation asdot

```
FTOS(conf-router_bgp)#bgp asnotation asdot
FTOS(conf-router_bgp)#sho conf
!
router bgp 100
 bgp asnotation asdot
bgp four-octet-as-support
 neighbor 172.30.1.250 remote-as 18508
 neighbor 172.30.1.250 local-as 65057
 neighbor 172.30.1.250 route-map rmap1 in
 neighbor 172.30.1.250 password 7
5ab3eb9a15ed02ff4f0dfd4500d6017873cfd9a267c04957
 neighbor 172.30.1.250 no shutdown
5332332 9911991 65057 18508 12182 7018 46164 i
```

Figure 10-18. Command example and output: bgp asnotation asdot+

```
FTOS(conf-router_bgp)#bgp asnotation asdot+
FTOS(conf-router_bgp)#sho conf
!
router bgp 100
 bgp asnotation asdot+
bgp four-octet-as-support
 neighbor 172.30.1.250 remote-as 18508
 neighbor 172.30.1.250 local-as 65057
 neighbor 172.30.1.250 route-map rmap1 in
 neighbor 172.30.1.250 password 7
5ab3eb9a15ed02ff4f0dfd4500d6017873cfd9a267c04957
 neighbor 172.30.1.250 no shutdown
5332332 9911991 65057 18508 12182 7018 46164 i
```

Configure Peer Groups

To configure multiple BGP neighbors at one time, create and populate a BGP peer group. Another advantage of peer groups is that members of a peer groups inherit the configuration properties of the group and share same update policy.

A *maximum* of 256 Peer Groups are allowed on the system.

You create a peer group by assigning it a name, then adding members to the peer group. Once a peer group is created, you can configure route policies for it. Refer to Filter BGP routes for information on configuring route policies for a peer group.

Note: Sample Configurations for enabling Peer Groups are found at the end of this chapter.

Use these commands in the following sequence starting in the CONFIGURATION ROUTER BGP mode to create a peer group

Step	Command Syntax	Command Mode	Purpose
1	**neighbor** *peer-group-name* **peer-group**	CONFIG-ROUTER-BGP	Create a peer group by assigning a name to it.
2	**neighbor** *peer-group-name* **no shutdown**	CONFIG-ROUTER-BGP	Enable the peer group. By default, all peer groups are disabled
3	**neighbor** *ip-address* **remote-as** *as-number*	CONFIG-ROUTER-BGP	Create a BGP neighbor.
4	**neighbor** *ip-address* **no shutdown**	CONFIG-ROUTER-BGP	Enable the neighbor.
5	**neighbor** *ip-address* **peer-group** *peer-group-name*	CONFIG-ROUTER-BGP	Add an enabled neighbor to the peer group.
6	**neighbor** {*ip-address* | *peer-group name*} **remote-as** *as-number*	CONFIG-ROUTER-BGP	Add a neighbor as a remote AS. Formats: IP Address A.B.C.D Peer-Group Name16 characters AS-number: 0-65535 (2-Byte) or 1-4294967295 | 0.1- 65535.65535 (4-Byte) or 0.1-65535.65535 (Dotted format) To add an external BGP (EBGP) neighbor, configure the *as-number* parameter with a number *different* from the BGP *as-number* configured in the **router bgp** *as-number* command. To add an internal BGP (IBGP neighbor, configure the *as-number* parameter with the *same* BGP *as-number* configured in the **router bgp** *as-number* command.

After you create a peer group, you can use any of the commands beginning with the keyword **neighbor** to configure that peer group.

When you add a peer to a peer group, it inherits all the peer group's configured parameters.

A neighbor *cannot* become part of a peer group if it has any of the following commands are configured:

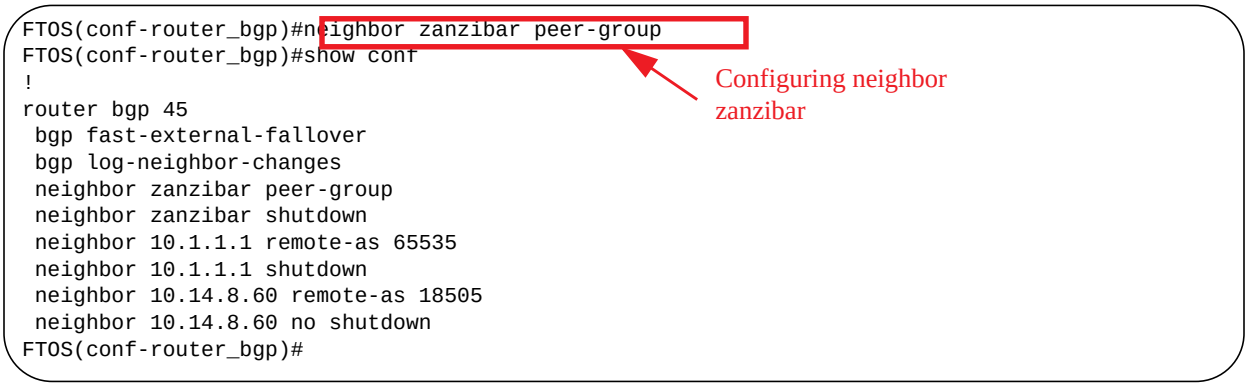
- neighbor advertisement-interval
- neighbor distribute-list out
- neighbor filter-list out
- neighbor next-hop-self
- neighbor route-map out
- neighbor route-reflector-client
- neighbor send-community

A neighbor may keep its configuration after it was added to a peer group if the neighbor's configuration is more specific than the peer group's, and the neighbor's configuration does not affect outgoing updates.

Note: When you configure a new set of BGP policies for a peer group, ***always*** reset the peer group by entering the **clear ip bgp peer-group** *peer-group-name* command in EXEC Privilege mode.

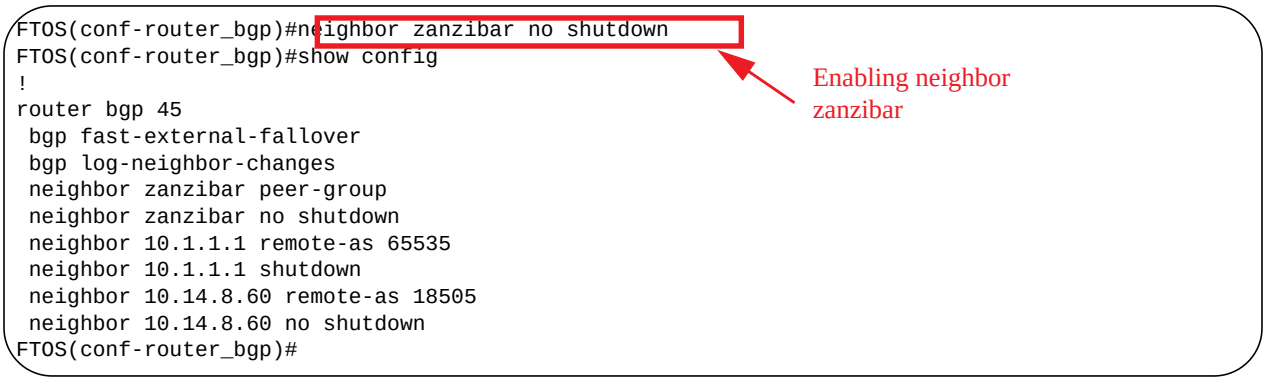
Use the **show config** command in the CONFIGURATION ROUTER BGP mode to view the configuration. When you create a peer group, it is disabled (**shutdown**). Figure 10-19 shows the creation of a peer group (zanzibar).

Figure 10-19. Command example: show config (creating peer-group)



Use the **neighbor peer-group-name no shutdown** command in the CONFIGURATION ROUTER BGP mode to enable a peer group.

Figure 10-20. Command example: show config (peer-group enabled



To disable a peer group, use the **neighbor** *peer-group-name* **shutdown** command in the CONFIGURATION ROUTER BGP mode. The configuration of the peer group is maintained, but it is not applied to the peer group members. When you disable a peer group, all the peers within the peer group that are in ESTABLISHED state are moved to IDLE state.

Use the show **ip bgp peer-group** command in EXEC Privilege mode (Figure 10-21) to view the status of peer groups.

Figure 10-21. Command example: show ip bgp peer-group

```
FTOS>show ip bgp peer-group


Peer-group zanzibar, remote AS 65535
BGP version 4
Minimum time between advertisement runs is 5 seconds

For address family: IPv4 Unicast
BGP neighbor is zanzibar, peer-group internal,
Number of peers in this group 26
Peer-group members (* - outbound optimized):
  10.68.160.1
  10.68.161.1
  10.68.162.1
  10.68.163.1
  10.68.164.1
  10.68.165.1
  10.68.166.1
  10.68.167.1
  10.68.168.1
  10.68.169.1
  10.68.170.1
  10.68.171.1
  10.68.172.1
  10.68.173.1
  10.68.174.1
  10.68.175.1
  10.68.176.1
  10.68.177.1
  10.68.178.1
  10.68.179.1
  10.68.180.1
  10.68.181.1
  10.68.182.1
  10.68.183.1
  10.68.184.1
  10.68.185.1
FTOS>
```

BGP fast fall-over

By default, a BGP session is governed by the hold time. BGP routers typically carry large routing tables, so frequent session resets are not desirable. The BGP fast fall-over feature reduces the convergence time while maintaining stability. The connection to a BGP peer is immediately reset if a link to a directly connected external peer fails.

When fall-over is enabled, BGP tracks IP reachability to the peer remote address and the peer local address. Whenever either address becomes unreachable (for example, no active route exists in the routing table for peer IPv6 destinations/local address), BGP brings down the session with the peer.

The BGP fast fall-over feature is configured on a per-neighbor or peer-group basis and is disabled by default.

Command Syntax	Command Mode	Purpose
neighbor {*ip-address* | *peer-group-name*} **fall-over**	CONFIG-ROUTER-BGP	Enable BGP Fast Fall-Over

To disable Fast Fall-Over, use the **[no] neighbor [neighbor | peer-group] fall-over** command in CONFIGURATION ROUTER BGP mode

Use the **show ip bgp neighbors** command as shown in Figure 10-22 to verify that fast fall-over is enabled on a particular BGP neighbor. Note that since Fast Fall-Over is disabled by default, it will appear only if it has been enabled

Figure 10-22. Command example: show ip bgp neighbors

```
FTOS#sh ip bgp neighbors

BGP neighbor is 100.100.100.100, remote AS 65517, internal link
  Member of peer-group test for session parameters
  BGP version 4, remote router ID 30.30.30.5
  BGP state ESTABLISHED, in this state for 00:19:15
  Last read 00:00:15, last write 00:00:06
  Hold time is 180, keepalive interval is 60 seconds
  Received 52 messages, 0 notifications, 0 in queue
  Sent 45 messages, 5 notifications, 0 in queue
  Received 6 updates, Sent 0 updates
  Route refresh request: received 0, sent 0
  Minimum time between advertisement runs is 5 seconds
  Minimum time before advertisements start is 0 seconds

  Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Capabilities advertised to neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Fall-over enabled          <-- Fast Fall-Over Indicator

  Update source set to Loopback 0

  Peer active in peer-group outbound optimization

  For address family: IPv4 Unicast
  BGP table version 52, neighbor version 52
  4 accepted prefixes consume 16 bytes
  Prefix advertised 0, denied 0, withdrawn 0

  Connections established 6; dropped 5
  Last reset 00:19:37, due to Reset by peer

  Notification History
   'Connection Reset' Sent : 5  Recv: 0

Local host: 200.200.200.200, Local port: 65519
Foreign host: 100.100.100.100, Foreign port: 179

FTOS#
```

Use the **show ip bgp peer-group** command to verify that fast fall-over is enabled on a peer-group.

Figure 10-23. Command example: show ip bgp peer-group

```
FTOS#sh ip bgp peer-group

Peer-group test
  Fall-over enabled
  BGP version 4
  Minimum time between advertisement runs is 5 seconds

  For address family: IPv4 Unicast
  BGP neighbor is test
  Number of peers in this group 1
  Peer-group members (* - outbound optimized):
    100.100.100.100*

FTOS#

router bgp 65517
 neighbor test peer-group
 neighbor test fall-over
 neighbor test no shutdown
 neighbor 100.100.100.100 remote-as 65517
 neighbor 100.100.100.100 fall-over
 neighbor 100.100.100.100 update-source Loopback 0
 neighbor 100.100.100.100 no shutdown
FTOS#
```

Fast Fall-Over Indicator

Configure passive peering

When you enable a peer-group, the software sends an OPEN message to initiate a TCP connection. If you enable passive peering for the peer group, the software does not send an OPEN message, but it will respond to an OPEN message.

When a BGP neighbor connection with authentication configured is rejected by a passive peer-group, FTOS does not allow another passive peer-group on the same subnet to connect with the BGP neighbor. To work around this, change the BGP configuration or change the order of the peer group configuration.

Use these commands in the following sequence, starting in the CONFIGURATION ROUTER BGP mode to configure passive peering.

Step	Command Syntax	Command Mode	Purpose
1	**neighbor** *peer-group-name* **peer-group passive [match-af]**	CONFIG-ROUTER-BGP	Configure a peer group that does not initiate TCP connections with other peers. (Optional) Enter the **match-af** keyword to restrict the peer adjacency established in the passive peer group. **match-af** requires that a peer's address family matches the address family of the subnet assigned to the peer group (Step 2) before a peering session is brought up.
2	**neighbor** *peer-group-name* **subnet** *subnet-number mask*	CONFIG-ROUTER-BGP	Assign a subnet to the peer group. The peer group will respond to OPEN messages sent on this subnet.

Step	Command Syntax	Command Mode	Purpose
3	**neighbor** *peer-group-name* **no shutdown**	CONFIG-ROUTER-BGP	Enable the peer group.
4	**neighbor** *peer-group-name* **remote-as** *as-number*	CONFIG-ROUTER-BGP	Create and specify a remote peer as a BGP neighbor.

Only after the peer group responds to an OPEN message sent on the subnet does its BGP state change to ESTABLISHED. Once the peer group is ESTABLISHED, the peer group is the same as any other peer group.

For more information on peer groups, refer to Configure Peer Groups on page 232.

Maintain existing AS numbers during an AS migration

The **local-as** feature smooths out the BGP network migration operation and allows you to maintain existing ASNs during a BGP network migration.

When you complete your migration, be sure to reconfigure your routers with the new information and disable this feature.

Command Syntax	Command Mode	Purpose
neighbor {*IP address* | *peer-group-name* **local-as** *as number* [no prepend]	CONFIG-ROUTER-BGP	Allow external routes from this neighbor. Format: IP Address: A.B.C.D Peer Group Name: 16 characters AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte) or 0.1-65535.65535 (Dotted format) No Prepend specifies that local AS values are not prepended to announcements from the neighbor.
	You must Configure Peer Groups *before* assigning it to an AS. This feature is not supported on passive peer groups.	

Disable this feature, using the **no neighbor local-as** command in CONFIGURATION ROUTER BGP mode.

Figure 10-24. Local-as information shown

```
R2(conf-router_bgp)#show conf
!
router bgp 65123
 bgp router-id 192.168.10.2
 network 10.10.21.0/24
 network 10.10.32.0/24
 network 100.10.92.0/24
 network 192.168.10.0/24
 bgp four-octet-as-support
 neighbor 10.10.21.1 remote-as 65123
 neighbor 10.10.21.1 filter-list Laura in
 neighbor 10.10.21.1 no shutdown
 neighbor 10.10.32.3 remote-as 65123
 neighbor 10.10.32.3 no shutdown
 neighbor 100.10.92.9 remote-as 65192
 neighbor 100.10.92.9 local-as 6500
 neighbor 100.10.92.9 no shutdown
 neighbor 192.168.10.1 remote-as 65123
 neighbor 192.168.10.1 update-source Loopback 0
 neighbor 192.168.10.1 no shutdown
 neighbor 192.168.12.2 remote-as 65123
 neighbor 192.168.12.2 update-source Loopback 0
 neighbor 192.168.12.2 no shutdown
R2(conf-router_bgp)#
```

Actual AS Number

Local-AS Number 6500
Maintained During Migration

Allow an AS number to appear in its own AS path

This command allows you to set the number of times a particular AS number can occur in the AS path. The **allow-as** feature permits a BGP speaker to allow the ASN to be present for specified number of times in the update received from the peer, even if that ASN matches its own. The AS-PATH loop is detected if the local ASN is present more than the specified number of times in the command.

Command Syntax	Command Mode	Purpose
neighbor {*IP address* | *peer-group-name}* **allowas-in** *number*	CONFIG-ROUTER-BGP	Allow this neighbor ID to use the AS path the specified number of times. Format: IP Address: A.B.C.D Peer Group Name: 16 characters Number: 1-10 You must Configure Peer Groups *before* assigning it to an AS.

To disable this feature, use the **no neighbor allow-as in** *number* command in the CONFIGURATION ROUTER BGP mode.

Figure 10-25. Allowas-in information shown

```
R2(conf-router_bgp)#show conf
!
router bgp 65123
 bgp router-id 192.168.10.2
 network 10.10.21.0/24
 network 10.10.32.0/24
 network 100.10.92.0/24
 network 192.168.10.0/24
 bgp four-octet-as-support
 neighbor 10.10.21.1 remote-as 65123
 neighbor 10.10.21.1 filter-list Laura in
 neighbor 10.10.21.1 no shutdown
 neighbor 10.10.32.3 remote-as 65123
 neighbor 10.10.32.3 no shutdown
 neighbor 100.10.92.9 remote-as 65192
 neighbor 100.10.92.9 local-as 6500
 neighbor 100.10.92.9 no shutdown
 neighbor 192.168.10.1 remote-as 65123
 neighbor 192.168.10.1 update-source Loopback 0
 neighbor 192.168.10.1 no shutdown
 neighbor 192.168.12.2 remote-as 65123
 neighbor 192.168.12.2 allowas-in 9
 neighbor 192.168.12.2 update-source Loopback 0
 neighbor 192.168.12.2 no shutdown
R2(conf-router_bgp)#R2(conf-router_bgp)#
```

Number of Times ASN 65123
Can Appear in AS PATH

Enable graceful restart

Use this feature to lessen the negative effects of a BGP restart. FTOS advertises support for this feature to BGP neighbors through a capability advertisement. You can enable graceful restart by router and/or by peer or peer group.

Note: By default, BGP graceful restart is disabled.

The default role for BGP on is as a receiving or restarting peer. If you enable BGP, when a peer that supports graceful restart resumes operating, FTOS performs the following tasks:

- Continues saving routes received from the peer if the peer advertised it had graceful restart capability. Continues forwarding traffic to the peer.
- Flags routes from the peer as Stale and sets a timer to delete them if the peer does not perform a graceful restart.
- Deletes all routes from the peer if forwarding state information is not saved.
- Speeds convergence by advertising a special update packet known as an end-of-RIB marker. This marker indicates the peer has been updated with all routes in the local RIB.

If you configure your system to do so, FTOS can perform the following actions during a hot failover:

- Save all FIB and CAM entries on the line card and continue forwarding traffic while the secondary RPM is coming online.

- Advertise to all BGP neighbors and peer-groups that the forwarding state of all routes has been saved. This prompts all peers to continue saving the routes they receive from your E-Series and to continue forwarding traffic.
- Bring the secondary RPM online as the primary and re-open sessions with all peers operating in "no shutdown" mode.
- Defer best path selection for a certain amount of time. This helps optimize path selection and results in fewer updates being sent out.

Enable graceful restart using the **configure router bgp graceful-restart** command. The table below shows the command and its available options:

Command Syntax	Command Mode	Usage
bgp graceful-restart	CONFIG-ROUTER-BGP	Enable graceful restart for the BGP node.
bgp graceful-restart [**restart-time** *time-in-seconds*]	CONFIG-ROUTER-BGP	Set maximum restart time for all peers. Default is 120 seconds.
bgp graceful-restart [**stale-path-time** *time-in-seconds*]	CONFIG-ROUTER-BGP	Set maximum time to retain the restarting peer's stale paths. Default is 360 seconds.
bgp graceful-restart [**role receiver-only**]	CONFIG-ROUTER-BGP	Local router supports graceful restart as a receiver only.

BGP graceful restart is active only when the neighbor becomes established. Otherwise it is disabled. Graceful-restart applies to all neighbors with established adjacency.

With the graceful restart feature, FTOS enables the receiving/restarting mode by default. In receiver-only mode, graceful restart saves the advertised routes of peers that support this capability when they restart. However, the E-Series does not advertise that it saves these forwarding states when it restarts. This option provides support for remote peers for their graceful restart without supporting the feature itself.

You can implement BGP graceful restart either by neighbor or by BGP peer-group. For more information, please see the following table or the *FTOS Command Line Interface Reference*.

Command Syntax	Command Mode	Purpose
neighbor {*ip-address* | *peer-group-name*} **graceful-restart**	CONFIG-ROUTER-BGP	Add graceful restart to a BGP neighbor or peer-group.
neighbor {*ip-address* | *peer-group-name*} **graceful-restart** [**restart-time** *time-in-seconds*]	CONFIG-ROUTER-BGP	Set maximum restart time for the neighbor or peer-group. Default is 120 seconds.
neighbor {*ip-address* | *peer-group-name*} **graceful-restart** [**role receiver-only**]	CONFIG-ROUTER-BGP	Local router supports graceful restart for this neighbor or peer-group as a receiver only.
neighbor {*ip-address* | *peer-group-name*} **graceful-restart** [**stale-path-time** *time-in-seconds*]	CONFIG-ROUTER-BGP	Set maximum time to retain the restarting neighbor's or peer-group's stale paths. Default is 360 seconds.

Filter on an AS-Path attribute

The BGP attribute, AS_PATH, can be used to manipulate routing policies. The AS_PATH attribute contains a sequence of AS numbers representing the route's path. As the route traverses an Autonomous System, the AS number is prepended to the route. You can manipulate routes based on their AS_PATH to affect interdomain routing. By identifying certain AS numbers in the AS_PATH, you can permit or deny routes based on the number in its AS_PATH.

To view all BGP path attributes in the BGP database, use the **show ip bgp paths** command in EXEC Privilege mode (Figure 10-26).

Figure 10-26. Command example: show ip bgp paths

```
FTOS#show ip bgp paths
Total 30655 Paths
Address       Hash Refcount Metric Path
0x4014154       0          3        18508 701 3549 19421 i
0x4013914       0          3        18508 701 7018 14990 i
0x5166d6c       0          3        18508 209 4637 1221 9249 9249 i
0x5e62df4       0          2        18508 701 17302 i
0x3a1814c       0         26        18508 209 22291 i
0x567ea9c       0         75        18508 209 3356 2529 i
0x6cc1294       0          2        18508 209 1239 19265 i
0x6cc18d4       0          1        18508 701 2914 4713 17935 i
0x5982e44       0        162        18508 209 i
0x67d4a14       0          2        18508 701 19878 ?
0x559972c       0         31        18508 209 18756 i
0x59cd3b4       0          2        18508 209 7018 15227 i
0x7128114       0         10        18508 209 3356 13845 i
0x536a914       0          3        18508 209 701 6347 7781 i
0x2ffe884       0          1        18508 701 3561 9116 21350 i
0x2ff7284       0         99        18508 701 1239 577 855 ?
0x2ff7ec4       0          4        18508 209 3561 4755 17426 i
0x2ff8544       0          3        18508 701 5743 2648 i
0x736c144       0          1        18508 701 209 568 721 1494 i
0x3b8d224       0         10        18508 209 701 2019 i
0x5eb1e44       0          1        18508 701 8584 16158 i
0x5cd891c       0          9        18508 209 6453 4759 i
--More--
```

AS-PATH ACLs use regular expressions to search AS_PATH values. AS-PATH ACLs have an "implicit deny." This means that routes that do not meet a deny or match filter are dropped.

Use these commands in the following sequence, starting in the CONFIGURATION mode to configure an AS-PATH ACL to filter a specific AS_PATH value.

Step	Command Syntax	Command Mode	Purpose
1	**ip as-path access-list** *as-path-name*	CONFIGURATION	Assign a name to a AS-PATH ACL and enter AS-PATH ACL mode.

Step	Command Syntax	Command Mode	Purpose
2	{**deny** | **permit**} *filter parameter*	CONFIG-AS-PATH	Enter the parameter to match BGP AS-PATH for filtering. This is the filter that will be used to match the AS-path. The entries can be any format, letters, numbers, or regular expressions. This command can be entered multiple times if multiple filters are desired. See Table 10-4 for accepted expressions.
3	**exit**	AS-PATH ACL	Return to CONFIGURATION mode
4	**router bgp** *as-number*	CONFIGURATION	Enter ROUTER BGP mode.
5	**neighbor** {*ip-address* | *peer-group-name*} **filter-list** *as-path-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Use a configured AS-PATH ACL for route filtering and manipulation. If you assign an non-existent or empty AS-PATH ACL, the software allows all routes.

Regular Expressions as filters

Regular expressions are used to filter AS paths or community lists. A regular expression is a special character used to define a pattern that is then compared with an input string.

For an AS-path access list as shown in the commands above, if the AS path matches the regular expression in the access list, then the route matches the access list.

Figure 10-27 applies access list Eagle to routes inbound from BGP peer 10.5.5.2. Access list Eagle uses a regular expression to deny routes originating in AS 32.

Figure 10-27. Filtering with Regular Expression

```
FTOS(config)#router bgp 99
FTOS(conf-router_bgp)#neigh AAA peer-group
FTOS(conf-router_bgp)#neigh AAA no shut
FTOS(conf-router_bgp)#show conf
!
router bgp 99
 neighbor AAA peer-group
 neighbor AAA no shutdown
 neighbor 10.155.15.2 remote-as 32
 neighbor 10.155.15.2 shutdown
FTOS(conf-router_bgp)#neigh 10.155.15.2 filter-list 1 in
FTOS(conf-router_bgp)#ex

FTOS(conf)#ip as-path access-list Eagle
FTOS(config-as-path)#deny 32$
FTOS(config-as-path)#ex
FTOS(conf)#router bgp 99
FTOS(conf-router_bgp)#neighbor AAA filter-list Eagle in
FTOS(conf-router_bgp)#show conf
!
router bgp 99
 neighbor AAA peer-group
 neighbor AAA filter-list Eaglein
 neighbor AAA no shutdown
 neighbor 10.155.15.2 remote-as 32
 neighbor 10.155.15.2 filter-list 1 in
 neighbor 10.155.15.2 shutdown
FTOS(conf-router_bgp)#ex
FTOS(conf)#ex

FTOS#show ip as-path-access-lists
ip as-path access-list Eagle
 deny 32$
FTOS#
```

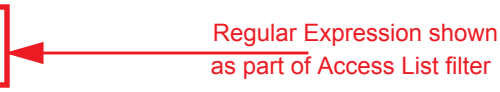


Table 10-4 lists the Regular Expressions accepted in FTOS.

Table 10-4. Regular Expressions

Regular Expression	Definition
^ (carrot)	Matches the beginning of the input string. Alternatively, when used as the first character within brackets [^] matches any number except the ones specified within the brackets.
$ (dollar)	Matches the end of the input string.
. (period)	Matches any single character, including white space.
* (asterisk)	Matches 0 or more sequences of the immediately previous character or pattern.
+ (plus)	Matches 1 or more sequences of the immediately previous character or pattern.
? (question)	Matches 0 or 1 sequence of the immediately previous character or pattern.

Table 10-4. Regular Expressions

Regular Expression	Definition
( ) (parenthesis)	Specifies patterns for multiple use when followed by one of the multiplier metacharacters: asterisk *, plus sign +, or question mark ?
[ ] (brackets)	Matches any enclosed character; specifies a range of single characters
- (hyphen)	Used within brackets to specify a range of AS or community numbers.
_ (underscore)	Matches a ^, a $, a comma, a space, a {, or a }. Placed on either side of a string to specify a literal and disallow substring matching. Numerals enclosed by underscores can be preceded or followed by any of the characters listed above.
| (pipe)	Matches characters on either side of the metacharacter; logical OR.

As seen in Figure 10-27, the expressions are displayed when using the **show** commands. Use the **show config** command in the CONFIGURATION AS-PATH ACL mode and the **show ip as-path-access-list** command in EXEC Privilege mode to view the AS-PATH ACL configuration.

For more information on this command and route filtering, refer to Filter BGP routes.

Redistribute routes

In addition to filtering routes, you can add routes from other routing instances or protocols to the BGP process. With the **redistribute** command syntax, you can include ISIS, OSPF, static, or directly connected routes in the BGP process.

Use any of the following commands in ROUTER BGP mode to add routes from other routing instances or protocols.

Command Syntax	Command Mode	Purpose
redistribute {**connected** | **static**} [**route-map** *map-name*]	ROUTER BGP or CONF-ROUTER_BGPv6_AF	Include, directly connected or user-configured (static) routes in BGP. Configure the following parameters: • *map-name*: name of a configured route map.
redistribute isis [**level-1** | **level-1-2** | **level-2**] [**metric** *value*] [**route-map** *map-name*]	ROUTER BGP or CONF-ROUTER_BGPv6_AF	Include specific ISIS routes in BGP. Configure the following parameters: • **level-1**, l**evel-1-2**, or **level-2**: Assign all redistributed routes to a level. Default is **level-2**. • *metric* range: 0 to 16777215. Default is 0. • *map-name*: name of a configured route map.

Command Syntax	Command Mode	Purpose
redistribute ospf *process-id* [**match external** {**1 | 2**} | **match internal**] [**metric-type** {**external** | **internal**}] [**route-map** *map-name*]	ROUTER BGP or CONF-ROUTER_BGPv6_ AF	Include specific OSPF routes in IS-IS. Configure the following parameters: • *process-id* range: 1 to 65535 • **match external** range: 1 or 2 • **match internal** • **metric-type**: external or internal. • *map-name*: name of a configured route map.

Configure IP community lists

Within FTOS, you have multiple methods of manipulating routing attributes. One attribute you can manipulate is the COMMUNITY attribute. This attribute is an optional attribute that is defined for a group of destinations. In FTOS, you can assign a COMMUNITY attribute to BGP routers by using an IP Community list. After you create an IP Community list, you can apply routing decisions to all routers meeting the criteria in the IP Community list.

IETF RFC 1997 defines the COMMUNITY attribute and the pre-defined communities of INTERNET, NO_EXPORT_SUBCONFED, NO_ADVERTISE, and NO_EXPORT. All BGP routes belong to the INTERNET community. In the RFC, the other communities are defined as follows:

- All routes with the NO_EXPORT_SUBCONFED (0xFFFFFF03) community attribute are not sent to CONFED-EBGP or EBGP peers, but are sent to IBGP peers within CONFED-SUB-AS.
- All routes with the NO_ADVERTISE (0xFFFFFF02) community attribute must not be advertised.
- All routes with the NO_EXPORT (0xFFFFFF01) community attribute must not be advertised outside a BGP confederation boundary, but are sent to CONFED-EBGP and IBGP peers.

FTOS also supports BGP Extended Communities as described in RFC 4360—BGP Extended Communities Attribute.

Use these commands in the following sequence, starting in the CONFIGURATION mode to configure an IP community list.

Step	Command Syntax	Command Mode	Purpose
1	**ip community-list** *community-list-name*	CONFIGURATION	Create a Community list and enter the COMMUNITY-LIST mode.
2	{**deny** | **permit**} {*community-number* | **local-AS** | **no-advertise** | **no-export** | **quote-regexp** *regular-expression-list* | **regexp** *regular-expression*}	CONFIG-COMMUNITY-LIST	Configure a Community list by denying or permitting specific community numbers or types of community • *community-number:* use AA:NN format where AA is the AS number (2 or 4 Bytes) and NN is a value specific to that autonomous system. • **local-AS**: routes with the COMMUNITY attribute of NO_EXPORT_SUBCONFED. • **no-advertise:** routes with the COMMUNITY attribute of NO_ADVERTISE. • **no-export:** routes with the COMMUNITY attribute of NO_EXPORT. • **quote-regexp:** followed by any number of regular expressions. The software applies all regular expressions in the list. • **regexp:** followed by a regular expression.

Use these commands in the following sequence, starting in the CONFIGURATION mode to configure an IP extended community list.

Step	Command Syntax	Command Mode	Purpose
1	**ip extcommunity-list** *extcommunity-list-name*	CONFIGURATION	Create a extended community list and enter the EXTCOMMUNITY-LIST mode.
2	{**permit** | **deny**} {{**rt** | **soo**} {*ASN:NN* | *IPADDR:N*} | **regex** *REGEX-LINE*}	CONFIG-COMMUNITY-LIST	Two types of extended communities are supported. Filter routes based on the type of extended communities they carry using one of the following keywords: • **rt**: Route Target • **soo**: Route Origin or Site-of-Origin. Support for matching extended communities against regular expression is also supported. Match against a regular expression using the following keyword: • **regexp**: regular expression

To set or modify an extended community attribute, use the **set extcommunity** {**rt** | **soo**} {*ASN:NN* | *IPADDR:NN*} command.

To view the configuration, use the **show config** command in the CONFIGURATION COMMUNITY-LIST or CONFIGURATION EXTCOMMUNITY LIST mode or the **show ip {community-lists | extcommunity-list} command** in EXEC Privilege mode (Figure 10-28).

Figure 10-28. Command example: show ip community-lists

```
FTOS#show ip community-lists
ip community-list standard 1
 deny 701:20
 deny 702:20
 deny 703:20
 deny 704:20
 deny 705:20
 deny 14551:20
 deny 701:112
 deny 702:112
 deny 703:112
 deny 704:112
 deny 705:112
 deny 14551:112
 deny 701:667
 deny 702:667
 deny 703:667
```

Use these commands in the following sequence, starting in the CONFIGURATION mode, To use an IP Community list or Extended Community List to filter routes, you must apply a **match community** filter to a route map and then apply that route map to a BGP neighbor or peer group.

Step	Command Syntax	Command Mode	Purpose
1	**route-map** *map-name* [**permit** | **deny**] [*sequence-number*]	CONFIGURATION	Enter the ROUTE-MAP mode and assign a name to a route map.
2	**match** {**community** *community-list-name* [**exact**] **| extcommunity** *extcommunity-list-name* [**exact**]}	CONFIG-ROUTE-MAP	Configure a match filter for all routes meeting the criteria in the IP Community or Extended Community list.
3	**exit**	CONFIG-ROUTE-MAP	Return to the CONFIGURATION mode.
4	**router bgp** *as-number*	CONFIGURATION	Enter the ROUTER BGP mode. AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte) or 0.1-65535.65535 (Dotted format)
5	**neighbor** {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Apply the route map to the neighbor or peer group's incoming or outgoing routes.

To view the BGP configuration, use the **show config** command in CONFIGURATION ROUTER BGP mode. To view a route map configuration, use the **show route-map** command in EXEC Privilege mode.

To view which BGP routes meet an IP Community or Extended Community list's criteria, use the **show ip bgp** {**community-list | extcommunity-list} command** in EXEC Privilege mode.

Manipulate the COMMUNITY attribute

In addition to permitting or denying routes based on the values of the COMMUNITY attributes, you can manipulate the COMMUNITY attribute value and send the COMMUNITY attribute with the route information.

By default, FTOS does not send the COMMUNITY attribute.

Use the following command in the CONFIGURATION ROUTER BGP mode to send the COMMUNITY attribute to BGP neighbors.

Command Syntax	Command Mode	Purpose
neighbor {*ip-address* | *peer-group-name*} **send-community**	CONFIG-ROUTER-BGP	Enable the software to send the router's COMMUNITY attribute to the BGP neighbor or peer group specified.

To view the BGP configuration, use the **show config** command in the CONFIGURATION ROUTER BGP mode.

If you want to remove or add a specific COMMUNITY number from a BGP path, you must create a route map with one or both of the following statements in the route map. Then apply that route map to a BGP neighbor or peer group. Use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**route-map** *map-name* [**permit** | **deny**] [*sequence-number*]	CONFIGURATION	Enter the ROUTE-MAP mode and assign a name to a route map.
2	**set comm-list** *community-list-name* **delete**	CONFIG-ROUTE-MAP	Configure a set filter to delete all COMMUNITY numbers in the IP Community list.
	set community {*community-number* | **local-as** | **no-advertise** | **no-export** | **none**}	CONFIG-ROUTE-MAP	Configure a Community list by denying or permitting specific community numbers or types of community • *community-number:* use AA:NN format where AA is the AS number (2 or 4 Bytes) and NN is a value specific to that autonomous system. • **local-AS**: routes with the COMMUNITY attribute of NO_EXPORT_SUBCONFED and are not sent to EBGP peers. • **no-advertise:** routes with the COMMUNITY attribute of NO_ADVERTISE and are not advertised. • **no-export:** routes with the COMMUNITY attribute of NO_EXPORT. • **none:** remove the COMMUNITY attribute. • **additive:** add the communities to already existing communities.

Step	Command Syntax	Command Mode	Purpose
3	**exit**	CONFIG-ROUTE-MAP	Return to the CONFIGURATION mode.
4	**router bgp** *as-number*	CONFIGURATION	Enter the ROUTER BGP mode.
5	**neighbor** {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Apply the route map to the neighbor or peer group's incoming or outgoing routes.

To view the BGP configuration, use the **show config** command in CONFIGURATION ROUTER BGP mode. To view a route map configuration, use the **show route-map** command in EXEC Privilege mode.

Use the **show ip bgp community** command in EXEC Privilege mode (Figure 10-29) to view BGP routes matching a certain community number or pre-defined BGP community.

Figure 10-29. Command example: show ip bgp community (Partial)

```
FTOS>show ip bgp community
BGP table version is 3762622, local router ID is 10.114.8.48
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

    Network          Next Hop            Metric        LocPrf    Weight  Path
* i 3.0.0.0/8        195.171.0.16                        100         0  209 701 80 i
*>i 4.2.49.12/30     195.171.0.16                        100         0  209 i
* i 4.21.132.0/23    195.171.0.16                        100         0  209 6461 16422 i
*>i 4.24.118.16/30   195.171.0.16                        100         0  209 i
*>i 4.24.145.0/30    195.171.0.16                        100         0  209 i
*>i 4.24.187.12/30   195.171.0.16                        100         0  209 i
*>i 4.24.202.0/30    195.171.0.16                        100         0  209 i
*>i 4.25.88.0/30     195.171.0.16                        100         0  209 3561 3908 i
*>i 6.1.0.0/16       195.171.0.16                        100         0  209 7170 1455 i
*>i 6.2.0.0/22       195.171.0.16                        100         0  209 7170 1455 i
*>i 6.3.0.0/18       195.171.0.16                        100         0  209 7170 1455 i
*>i 6.4.0.0/16       195.171.0.16                        100         0  209 7170 1455 i
*>i 6.5.0.0/19       195.171.0.16                        100         0  209 7170 1455 i
*>i 6.8.0.0/20       195.171.0.16                        100         0  209 7170 1455 i
*>i 6.9.0.0/20       195.171.0.16                        100         0  209 7170 1455 i
*>i 6.10.0.0/15      195.171.0.16                        100         0  209 7170 1455 i
```

Change MED attribute

By default, FTOS uses the MULTI_EXIT_DISC or MED attribute when comparing EBGP paths from the same AS.

Use any or all of the following commands in the CONFIGURATION ROUTER BGP mode to change how the MED attribute is used.

Command Syntax	Command Mode	Purpose
bgp always-compare-med	CONFIG-ROUTER-BGP	Enable MED comparison in the paths from neighbors with different ASs. By default, this comparison is not performed.
bgp bestpath med {confed | missing-as-best}	CONFIG-ROUTER-BGP	Change the bestpath MED selection to one of the following: **confed**: Chooses the bestpath MED comparison of paths learned from BGP confederations. **missing-as-bes**t: Treat a path missing an MED as the most preferred one

Use the **show config** command in the CONFIGURATION ROUTER BGP mode to view the nondefault values.

Change LOCAL_PREFERENCE attribute

In FTOS, you can change the value of the LOCAL_PREFERENCE attribute.

Use the following command in the CONFIGURATION ROUTER BGP mode to change the default values of this attribute for all routes received by the router.

Command Syntax	Command Mode	Purpose
bgp default local-preference *value*	CONFIG-ROUTER-BGP	Change the LOCAL_PREF value. • *value* range: 0 to 4294967295 • Default is 100.

Use the **show config** command in CONFIGURATION ROUTER BGP mode or the **show running-config bgp** command in EXEC Privilege mode to view BGP configuration.

A more flexible method for manipulating the LOCAL_PREF attribute value is to use a route map.

Use these commands in the following sequence, starting CONFIGURATION mode to change the default value of the LOCAL_PREF attribute for specific routes.

Step	Command Syntax	Command Mode	Purpose
1	**route-map** *map-name* [**permit** | **deny**] [*sequence-number*]	CONFIGURATION	Enter the ROUTE-MAP mode and assign a name to a route map.
2	**set local-preference** *value*	CONFIG-ROUTE-MAP	Change LOCAL_PREF value for routes meeting the criteria of this route map.
3	**exit**	CONFIG-ROUTE-MAP	Return to the CONFIGURATION mode.

Step	Command Syntax	Command Mode	Purpose
4	**router bgp** *as-number*	CONFIGURATION	Enter the ROUTER BGP mode.
5	**neighbor** {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Apply the route map to the neighbor or peer group's incoming or outgoing routes.

To view the BGP configuration, use the **show config** command in the CONFIGURATION ROUTER BGP mode. To view a route map configuration, use the **show route-map** command in EXEC Privilege mode.

Change NEXT_HOP attribute

You can change how the NEXT_HOP attribute is used.

Use the following command in the CONFIGURATION ROUTER BGP mode to change the how the NEXT_HOP attribute is used.

Command Syntax	Command Mode	Purpose
neighbor {*ip-address* | *peer-group-name*} **next-hop-self**	CONFIG-ROUTER-BGP	Disable next hop processing and configure the router as the next hop for a BGP neighbor.

Use the **show config** command in CONFIGURATION ROUTER BGP mode or the **show running-config bgp** command in EXEC Privilege mode to view BGP configuration.

You can also use route maps to change this and other BGP attributes. For example, you can include the following command in a route map to specify the next hop address:

Command Syntax	Command Mode	Purpose
set next-hop *ip-address*	CONFIG-ROUTE-MAP	Sets the next hop address.

Change WEIGHT attribute

Use the following command in CONFIGURATION ROUTER BGP mode to change the how the WEIGHT attribute is used.

Command Syntax	Command Mode	Purpose
neighbor {*ip-address* | *peer-group-name*} **weight** *weight*	CONFIG-ROUTER-BGP	Assign a weight to the neighbor connection. • *weight* range: 0 to 65535 • Default is 0

Use the **show config** command in CONFIGURATION ROUTER BGP mode or the **show running-config bgp** command in EXEC Privilege mode to view BGP configuration.

You can also use route maps to change this and other BGP attributes. For example, you can include the following command in a route map to specify the next hop address:

Command Syntax	Command Mode	Purpose
set weight *weight*	CONFIG-ROUTE-MAP	Sets weight for the route. • *weight* range: 0 to 65535

Enable multipath

By default, the software allows one path to a destination. You can enable multipath to allow up to 16 parallel paths to a destination.

Use the following command in the CONFIGURATION ROUTER BGP mode to allow more than one path.

Command Syntax	Command Mode	Purpose
maximum-paths {**ebgp** | **ibgp**} *number*	CONFIG-ROUTER-BGP	Enable multiple parallel paths. • *number* range: 1 to 16 • Default is 1

The **show ip bgp** *network* command includes multipath information for that network.

Filter BGP routes

Filtering routes allows you to implement BGP policies. You can use either IP prefix lists, route maps, AS-PATH ACLs or IP Community lists (via a route map) to control which routes are accepted and advertised by the BGP neighbor or peer group. Prefix lists filter routes based on route and prefix length, while AS-Path ACLs filter routes based on the Autonomous System number. Route maps can filter and set conditions, change attributes, and assign update policies.

Note: FTOS supports up to 255 characters in a set community statement inside a route map.

Note: With FTOS, you can create inbound and outbound policies. Each of the commands used for filtering, has **in** and **out** parameters that must be applied. In FTOS, the order of preference varies depending on whether the attributes are applied for inbound updates or outbound updates.

For inbound and outbound updates the order of preference is:

- prefix lists (using **neighbor distribute-list** command)
- AS-PATH ACLs (using **neighbor filter-list** command)
- route maps (using **neighbor route-map** command)

Prior to filtering BGP routes, you must create the prefix list, AS-PATH ACL, or route map to be used.

Refer to Chapter 8, "IP Access Control Lists (ACL), Prefix Lists, and Route-maps," on page 133 for configuration information on prefix lists, AS-PATH ACLs, and route maps.

Note: When you configure a new set of BGP policies, always reset the neighbor or peer group by entering the **clear ip bgp** command in EXEC Privilege mode.

Use these commands in the following sequence, starting in the CONFIGURATION mode to filter routes using prefix lists.

Step	Command Syntax	Command Mode	Purpose
1	**ip prefix-list** *prefix-name*	CONFIGURATION	Create a prefix list and assign it a name.
2	**seq** *sequence-number* {**deny** | **permit**} {*any* | *ip-prefix* [**ge | le**] }	CONFIG-PREFIX LIST	Create multiple prefix list filters with a deny or permit action. **ge**: Minimum prefix length to be matched **le**: maximum prefix length to me matched Refer to Chapter 8, "IP Access Control Lists (ACL), Prefix Lists, and Route-maps," on page 133 for information on configuring prefix lists.
3	**exit**	CONFIG-PREFIX LIST	Return to the CONFIGURATION mode.
4	**router bgp** *as-number*	CONFIGURATION	Enter ROUTER BGP mode.
5	**neighbor** {*ip-address* | *peer-group-name*} **distribute-list** *prefix-list-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Filter routes based on the criteria in the configured prefix list. Configure the following parameters: • *ip-address* or *peer-group-name:* enter the neighbor's IP address or the peer group's name. • *prefix-list-name:* enter the name of a configured prefix list. • **in:** apply the prefix list to inbound routes. • **out:** apply the prefix list to outbound routes.

As a reminder, below are some rules concerning prefix lists:

- If the prefix list contains no filters, all routes are permitted.
- If none of the routes match any of the filters in the prefix list, the route is denied. This action is called an implicit deny. (If you want to forward all routes that do not match the prefix list criteria, you must configure a prefix list filter to permit all routes. For example, you could have the following filter as the last filter in your prefix list **permit 0.0.0.0/0 le 32**).
- Once a route matches a filter, the filter's action is applied. No additional filters are applied to the route.

To view the BGP configuration, use the **show config** command in the ROUTER BGP mode. To view a prefix list configuration, use the **show ip prefix-list detail** or **show ip prefix-list summary** commands in EXEC Privilege mode.

Use these commands in the following sequence, starting in the CONFIGURATION mode to filter routes using a route map.

Step	Command Syntax	Command Mode	Purpose
1	**route-map** *map-name* [**permit** | **deny**] [*sequence-number*]	CONFIGURATION	Create a route map and assign it a name.
2	{**match** | **set**}	CONFIG-ROUTE-MAP	Create multiple route map filters with a match or set action. Refer to Chapter 8, "IP Access Control Lists (ACL), Prefix Lists, and Route-maps," on page 133 for information on configuring route maps.
3	**exit**	CONFIG-ROUTE-MAP	Return to the CONFIGURATION mode.
4	**router bgp** *as-number*	CONFIGURATION	Enter ROUTER BGP mode.
	neighbor {*ip-address* | *peer-group-name*} **route-map** *map-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Filter routes based on the criteria in the configured route map. Configure the following parameters: • *ip-address* or *peer-group-name:* enter the neighbor's IP address or the peer group's name. • *map-name:* enter the name of a configured route map. • **in:** apply the route map to inbound routes. • **out:** apply the route map to outbound routes.

Use the **show config** command in CONFIGURATION ROUTER BGP mode to view the BGP configuration. Use the **show route-map** command in EXEC Privilege mode to view a route map configuration.

Use these commands in the following sequence, beginning in the CONFIGURATION mode to filter routes based on AS-PATH information.

Step	Command Syntax	Command Mode	Purpose
1	**ip as-path access-list** *as-path-name*	CONFIGURATION	Create a AS-PATH ACL and assign it a name.
2	{**deny** | **permit**} *as-regular-expression*	AS-PATH ACL	Create a AS-PATH ACL filter with a deny or permit action.
3	**exit**	AS-PATH ACL	Return to the CONFIGURATION mode.
4	**router bgp** *as-number*	CONFIGURATION	Enter ROUTER BGP mode.

Step	Command Syntax	Command Mode	Purpose
5	**neighbor** {*ip-address* | *peer-group-name*} **filter-list** *as-path-name* {**in** | **out**}	CONFIG-ROUTER-BGP	Filter routes based on the criteria in the configured route map. Configure the following parameters: • *ip-address* or *peer-group-name:* enter the neighbor's IP address or the peer group's name. • *as-path-name:* enter the name of a configured AS-PATH ACL. • **in:** apply the AS-PATH ACL map to inbound routes. • **out:** apply the AS-PATH ACL to outbound routes.

Use the **show config** command in CONFIGURATION ROUTER BGP mode and **show ip as-path-access-list** command in EXEC Privilege mode to view which commands are configured.

Include this filter **permit .*** in your AS-PATH ACL to forward all routes not meeting the AS-PATH ACL criteria.

Configure BGP route reflectors

BGP route reflectors are intended for Autonomous Systems with a large mesh and they reduce the amount of BGP control traffic. With route reflection configured properly, IBGP routers are not fully meshed within a cluster but all receive routing information.

Configure clusters of routers where one router is a concentration router and others are clients who receive their updates from the concentration router.

Use the following commands in the CONFIGURATION ROUTER BGP mode to configure a route reflector.

Command Syntax	Command Mode	Purpose
bgp cluster-id *cluster-id*	CONFIG-ROUTER-BGP	Assign an ID to a router reflector cluster. You can have multiple clusters in an AS.
neighbor {*ip-address* | *peer-group-name*} **route-reflector-client**	CONFIG-ROUTER-BGP	Configure the local router to function as the route reflector and a specified neighbor or peer group to be the route-reflector clients in the cluster.

To view a route reflector configuration, use the **show config** command in the CONFIGURATION ROUTER BGP mode or **show running-config bgp** in EXEC Privilege mode.

When you enable a route reflector, FTOS automatically enables route reflection to all clients. To disable route reflection between all clients in this reflector, use the **no bgp client-to-client reflection** command in CONFIGURATION ROUTER BGP mode. All clients should be fully meshed before you disable route reflection.

Aggregate routes

FTOS provides multiple ways to aggregate routes in the BGP routing table. At least one specific route of the aggregate must be in the routing table for the configured aggregate to become active.

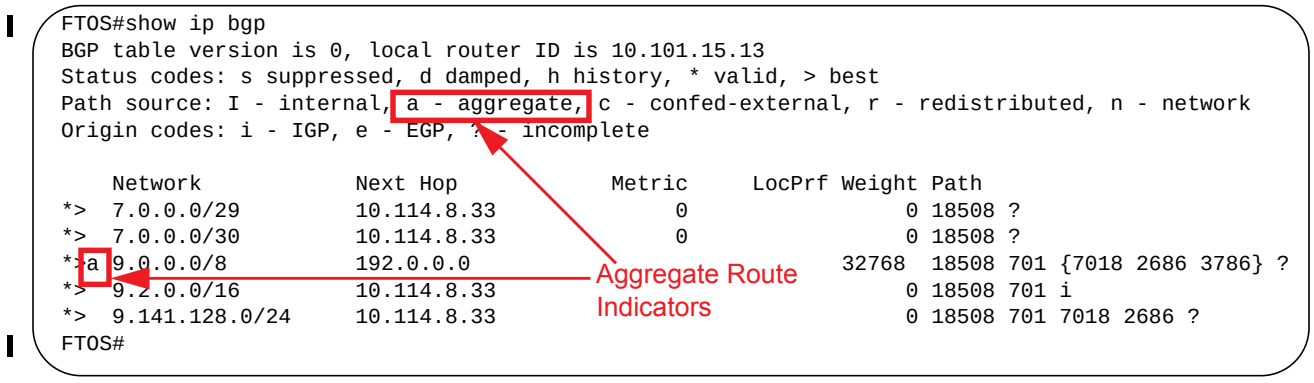
Use the following command in the CONFIGURATION ROUTER BGP mode to aggregate routes.

Command Syntax	Command Mode	Purpose
aggregate-address *ip-address mask* [**advertise-map** *map-name*] [**as-set**] [**attribute-map** *map-name*] [**summary-only**] [**suppress-map** *map-name*]	CONFIG-ROUTER-BGP	Assign the IP address and mask of the prefix to be aggregated. Optional parameters are: • **advertise-map** *map-name*: set filters for advertising an aggregate route • **as-set**: generate path attribute information and include it in the aggregate. • **attribute-map** *map-name: modify* attributes of the aggregate, except for the AS_PATH and NEXT_HOP attributes • **summary-only**: advertise only the aggregate address. Specific routes will not be advertised • **suppress-map** *map-name*: identify which more-specific routes in the aggregate are suppressed

AS_SET includes AS_PATH and community information from the routes included in the aggregated route.

In the **show ip bgp** command, aggregates contain an ‘a’ in the first column and routes suppressed by the aggregate contain an ‘s’ in the first column.

Figure 10-30. Command Example: show ip bgp



Configure BGP confederations

Another way to organize routers within an AS and reduce the mesh for IBGP peers is to configure BGP confederations. As with route reflectors, BGP confederations are recommended only for IBGP peering involving a large number of IBGP peering sessions per router. Basically, when you configure BGP confederations, you break the AS into smaller sub-AS, and to those outside your network, the confederations appear as one AS. Within the confederation sub-AS, the IBGP neighbors are fully meshed and the MED, NEXT_HOP, and LOCAL_PREF attributes are maintained between confederations.

Use the following commands in the CONFIGURATION ROUTER BGP mode to configure BGP confederations.

Command Syntax	Command Mode	Purpose
bgp confederation identifier *as-number*	CONFIG-ROUTER-BGP	Specifies the confederation ID. AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte)
bgp confederation peers *as-number* [... *as-number*]	CONFIG-ROUTER-BGP	Specifies which confederation sub-AS are peers. AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte) All Confederation routers must be either 4-Byte or 2-Byte. You cannot have a mix of router ASN support,

Use the **show config** command in the CONFIGURATION ROUTER BGP mode to view the configuration.

Enable route flap dampening

When EBGP routes become unavailable, they “flap” and the router issues both WITHDRAWN and UPDATE notices. A flap is when a route

- is withdrawn
- is readvertised after being withdrawn
- has an attribute change

The constant router reaction to the WITHDRAWN and UPDATE notices causes instability in the BGP process. To minimize this instability, you may configure penalties, a numeric value, for routes that flap. When that penalty value reaches a configured limit, the route is not advertised, even if the route is up. In FTOS, that penalty value is 1024. As time passes and the route does not flap, the penalty value decrements or is decayed. However, if the route flaps again, it is assigned another penalty.

The penalty value is cumulative and penalty is added under following cases:

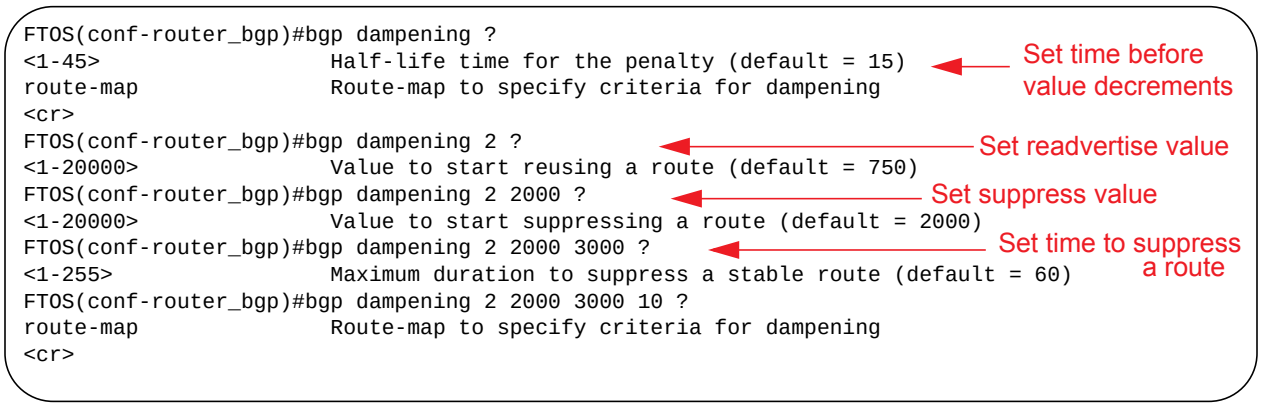
- Withdraw
- Readvertise
- Attribute change

When dampening is applied to a route, its path is described by one of the following terms:

- history entry—an entry that stores information on a downed route
- dampened path—a path that is no longer advertised
- penalized path—a path that is assigned a penalty

The CLI example below shows configuring values to start reusing or restarting a route, as well as their default values.

Figure 10-31. Setting Reuse and Restart Route Values



Use the following command in the CONFIGURATION ROUTER BGP mode to configure route flap dampening parameters.

Command Syntax	Command Mode	Purpose
bgp dampening [*half-life* | *reuse* | *suppress max-suppress-time*] [**route-map** *map-name*]	CONFIG-ROUTER-BGP	Enable route dampening. Enter the following optional parameters to configure route dampening parameters: • *half-life* range: 1 to 45. Number of minutes after which the Penalty is decreased. After the router assigns a Penalty of 1024 to a route, the Penalty is decreased by half after the half-life period expires. (Default: 15 minutes) • *reuse* range: 1 to 20000. This number is compared to the flapping route's Penalty value. If the Penalty value is less than the reuse value, the flapping route is once again advertised (or no longer suppressed). Withdrawn routes are removed from history state. (Default: 750) • *suppress* range: 1 to 20000. This number is compared to the flapping route's Penalty value. If the Penalty value is greater than the suppress value, the flapping route is no longer advertised (that is, it is suppressed). (Default: 2000.) • *max-suppress-time* range: 1 to 255. The maximum number of minutes a route can be suppressed. The default is four times the half-life value. (Default: 60 minutes.) • **route-map** *map-name:* name of a configured route map. Only match commands in the configured route map are supported. Use this parameter to apply route dampening to selective routes.

To view the BGP configuration, use **show config** in the CONFIGURATION ROUTER BGP mode or **show running-config bgp** in EXEC Privilege mode.

To set dampening parameters via a route map, use the following command in CONFIGURATION ROUTE-MAP mode:

Command Syntax	Command Mode	Purpose
set dampening *half-life reuse suppress max-suppress-time*	CONFIG-ROUTE-MAP	Enter the following optional parameters to configure route dampening parameters: • *half-life* range: 1 to 45. Number of minutes after which the Penalty is decreased. After the router assigns a Penalty of 1024 to a route, the Penalty is decreased by half after the half-life period expires. (Default: 15 minutes) • *reuse* range: 1 to 20000. This number is compared to the flapping route's Penalty value. If the Penalty value is less than the reuse value, the flapping route is once again advertised (or no longer suppressed). (Default: 750) • *suppress* range: 1 to 20000. This number is compared to the flapping route's Penalty value. If the Penalty value is greater than the suppress value, the flapping route is no longer advertised (that is, it is suppressed). (Default: 2000.) • *max-suppress-time* range: 1 to 255. The maximum number of minutes a route can be suppressed. The default is four times the half-life value. (Default: 60 minutes.)

To view a count of dampened routes, history routes and penalized routes when route dampening is enabled, look at the seventh line of the **show ip bgp** summary command output (Figure 10-32).

Figure 10-32. Command example: show ip bgp summary

```
FTOS>show ip bgp summary
BGP router identifier 10.114.8.131, local AS number 65515
BGP table version is 855562, main routing table version 780266
122836 network entrie(s) and 221664 paths using 29697640 bytes of memory
34298 BGP path attribute entrie(s) using 1920688 bytes of memory
29577 BGP AS-PATH entrie(s) using 1384403 bytes of memory
184 BGP community entrie(s) using 7616 bytes of memory
Dampening enabled. 0 history paths, 0 dampened paths, 0 penalized paths

Neighbor        AS    MsgRcvd MsgSent   TblVer  InQ   OutQ Up/Down  State/PfxRcd

10.114.8.34     18508   82883   79977   780266    0      2 00:38:51       118904
10.114.8.33     18508  117265   25069   780266    0     20 00:38:50       102759
FTOS>
```

Dampening Information

To view which routes are dampened (non-active), use the **show ip bgp dampened-routes** command in EXEC Privilege mode.

Use the following command in EXEC Privilege mode to clear information on route dampening and return suppressed routes to active state.

Command Syntax	Command Mode	Purpose
clear ip bgp dampening [*ip-address mask*]	EXEC Privilege	Clear all information or only information on a specific route.

Use the following command in EXEC and EXEC Privilege mode to view statistics on route flapping.

Command Syntax	Command Mode	Purpose
show ip bgp flap-statistics [*ip-address* [*mask*]] [**filter-list** *as-path-name*] [**regexp** *regular-expression*]	EXEC EXEC Privilege	View all flap statistics or for specific routes meeting the following criteria: • *ip-address* [*mask*]: enter the IP address and mask • **filter-list** *as-path-name:* enter the name of an AS-PATH ACL. • **regexp** *regular-expression:* enter a regular express to match on.

By default, the path selection in FTOS is deterministic, that is, paths are compared irrespective of the order of their arrival. You can change the path selection method to non-deterministic, that is, paths are compared in the order in which they arrived (starting with the most recent). Furthermore, in non-deterministic mode, the software may not compare MED attributes though the paths are from the same AS.

Use the following command in CONFIGURATION ROUTER BGP mode to change the path selection from the default mode (deterministic) to non-deterministic.

Command Syntax	Command Mode	Purpose
bgp non-deterministic-med	CONFIG-ROUTER-BGP	Change the best path selection method to non-deterministic.

Note: When you change the best path selection method, path selection for existing paths remains unchanged until you reset it by entering the **clear ip bgp** command in EXEC Privilege mode.

Change BGP timers

Use either or both of the following commands in the CONFIGURATION ROUTER BGP mode to configure BGP timers.

Command Syntax	Command Mode	Purpose
neighbors {*ip-address* | *peer-group-name*} **timers** *keepalive holdtime*	CONFIG-ROUTER-BGP	Configure timer values for a BGP neighbor or peer group. • *keepalive* range: 1 to 65535. Time interval, in seconds, between keepalive messages sent to the neighbor routers. (Default: 60 seconds) • *holdtime* range: 3 to 65536. Time interval, in seconds, between the last keepalive message and declaring the router dead. (Default: 180 seconds)
timers bgp *keepalive holdtime*	CONFIG-ROUTER-BGP	Configure timer values for all neighbors. • *keepalive* range: 1 to 65535. Time interval, in seconds, between keepalive messages sent to the neighbor routers. (Default: 60 seconds) • *holdtime* range: 3 to 65536. Time interval, in seconds, between the last keepalive message and declaring the router dead. (Default: 180 seconds)

Use the **show config** command in CONFIGURATION ROUTER BGP mode or the **show running-config bgp** command in EXEC Privilege mode to view non-default values.

Timer values configured with the **neighbor timers** command override the timer values configured with the **timers bgp** command.

When two neighbors, configured with different *keepalive* and *holdtime* values, negotiate for new values, the resulting values will be as follows:

- the lower of the *holdtime* values is the new *holdtime* value, and
- whichever is the lower value; one-third of the new *holdtime* value, or the configured *keepalive* value is the new *keepalive* value.

BGP neighbor soft-reconfiguration

Changing routing policies typically requires a reset of BGP sessions (the TCP connection) for the policies to take effect. This type of reset causes undue interruption to traffic due to the hard reset of the BGP cache and the time it takes to re-establish the session.

BGP soft reconfiguration allows you to re-apply policies to a session without resetting the BGP session. You can perform soft reconfiguration on a per-neighbor basis for either inbound or outbound policies. BGP soft reconfiguration clears and reapplies policies without resetting the TCP connection.

Use the **clear ip bgp** command in EXEC Privilege mode to reset a BGP connection using BGP soft reconfiguration.

Command Syntax	Command Mode	Purpose
neighbor {*ipv4-address* | *ipv6-address* | *peer-group-name*} **soft-reconfiguration inbound**	CONFIG-ROUTER-BGP	Enable inbound soft-reconfiguration for the specified BGP neighbor. BGP stores all updates received by the neighbor but does not reset the peer session. Entering this command starts the storage of updates on inbound routes, which is required for inbound soft reconfiguration. Outbound BGP soft reconfiguration does not require inbound soft reconfiguration to be enabled.
clear ip bgp { * | *as-number* | *ipv4-neighbor-addr* | *ipv6-neighbor-addr* **| peer-group** *name*} **{ipv4 unicast | ipv4 multicast | ipv6 unicast} soft** [**in | out**]	EXEC Privilege	Clears and reapplies policies on: *: All peers *as-number*: BGP routers that belong to the specified AS *neighbor-addr*: BGP neighbor with specified IP address Type of routes to be reapplied: **ipv4 unicast**, **ipv4 multicast**, or **ipv6 unicast** **in**: Reapplies only inbound policies **out**: Reapplies only outbound policies

When inbound soft reconfiguration is enabled on a neighbor and you enter the **clear ip bgp soft in** command, the update database stored in the router is replayed and updates are reevaluated. The replay and update process is triggered only if a route-refresh request is not negotiated with the peer. If a route-refresh request is negotiated, BGP sends a request to the neighbor and receives all of the peer's updates.

To use soft reconfiguration (or soft reset) without preconfiguration, both BGP peers must support the soft route refresh capability, which is advertised in the open message sent when the peers establish a TCP session. To determine whether a BGP router supports this capability, use the **show ip bgp neighbors** command. If a router supports the route refresh capability, the following message is displayed:

```
Received route refresh capability from peer.
```

If you specify a BGP peer group by using the *peer-group-name* argument, all members of the peer group inherit the characteristic configured with this command.

The following example (Figure 10-33) shows how to enable inbound soft reconfiguration for the neighbor 10.108.1.1. All updates received from this neighbor are stored unmodified, regardless of the inbound policy. When inbound soft reconfiguration is performed later, the stored information is used to generate a new set of inbound updates.

Figure 10-33. Command example: router bgp

```
FTOS>router bgp 100
      neighbor 10.108.1.1 remote-as 200
      neighbor 10.108.1.1 soft-reconfiguration inbound
```

Route map continue

The BGP route map **continue** feature (in ROUTE-MAP mode) allows movement from one route-map entry to a specific route-map entry (the *sequence number*). If the sequence number is not specified, the continue feature moves to the next sequence number (also known as an implied continue). If a match clause exists, the **continue** feature executes only after a successful match occurs. If there are no successful matches, **continue** is ignored.

continue [*sequence-number***]**

Match Clause with a Continue Clause

The **continue** feature can exist without a match clause. Without a match clause, the continue clause executes and jumps to the specified route-map entry. With a match clause and a continue clause, the match clause executes first and the continue clause next in a specified route map entry. The continue clause launches only after a successful match. The behavior is:

- A successful match with a continue clause—the route map executes the set clauses and then goes to the specified route map entry upon execution of the continue clause.
- If the next route map entry contains a continue clause, the route map executes the continue clause if a successful match occurs.
- If the next route map entry does not contain a continue clause, the route map evaluates normally. If a match does not occur, the route map does not continue and falls-through to the next sequence number, if one exists.

Set Clause with a Continue Clause

If the route-map entry contains sets with the continue clause, then the set actions operation is performed first followed by the continue clause jump to the specified route map entry.

- If a set actions operation occurs in the first route map entry and then the same set action occurs with a different value in a subsequent route map entry, the last set of actions overrides the previous set of actions with the same **set** command.
- If the **set community additive** and **set as-path prepend** commands are configured, the communities and AS numbers are prepended.

MBGP Configuration

MBGP for IPv6 unicast is supported on platforms E(T) C

MBGP for IPv4 Multicast is supported on platform C E(T) S

MBGP is *not* supported on the E-Series ExaScale E(X) platform.

Multiprotocol BGP (MBGP) is an enhanced BGP that carries IP multicast routes. BGP carries two sets of routes: one set for unicast routing and one set for multicast routing. The routes associated with multicast routing are used by the Protocol Independent Multicast (PIM) to build data distribution trees.

FTOS MBGP is implemented as per RFC 1858. The MBGP feature can be enabled per router and/or per peer/peer-group.

Default is IPv4 Unicast routes.

Command Syntax	Command Mode	Purpose
address family ipv4 multicast	CONFIG-ROUTER-BGP	Enables support for the IPv4 Multicast family on the BGP node
neighbor [*ip-address* | *peer-group-name*] **activate**	CONFIG-ROUTER-BGP-AF (Address Family)	Enable IPv4 Multicast support on a BGP neighbor/peer group

When a peer is configured to support IPv4 Multicast, FTOS takes the following actions:

- Send a capacity advertisement to the peer in the BGP Open message specifying IPv4 Multicast as a supported AFI/SAFI (Subsequent Address Family Identifier).
- If the corresponding capability is received in the peer's Open message, BGP will mark the peer as supporting the AFI/SAFI.
- When exchanging updates with the peer, BGP sends and receives IPv4 Multicast routes if the peer is marked as supporting that AFI/SAFI.
- Exchange of IPv4 Multicast route information occurs through the use of two new attributes called MP_REACH_NLRI and MP_UNREACH_NLRI, for feasible and withdrawn routes, respectively.
- If the peer has not been activated in any AFI/SAFI, the peer remains in Idle state.

Most FTOS BGP IPv4 Unicast commands are extended to support the IPv4 Multicast RIB using extra options to the command. See the *FTOS Command Line Interface Reference* for a detailed description of the MBGP commands.

BGP Regular Expression Optimization

BGP policies that contain regular expressions to match against as-paths and communities might take a lot of CPU processing time, thus affect BGP routing convergence. Also, **show bgp** commands that get filtered through regular expressions can to take a lot of CPU cycles, especially when the database is large. FTOS optimizes processing time when using regular expressions by caching and re-using regular expression evaluated results, at the expense of some memory in RP1 processor. This feature is turned on by default. Use the command **bgp regex-eval-optz-disable** in CONFIGURATION ROUTER BGP mode to disable it if necessary.

Retain NH in BGP Advertisement

BGP does not update the NEXT_HOP attribute if it is a Route-Reflector. **bgp retain-ibgp-nexthop** can be configured to retain the NEXT_HOP attribute when advertising to internal BGP peer.

Debugging BGP

Use any of the commands in EXEC Privilege mode to enable BGP debugging.

Command Syntax	Command Mode	Purpose
debug ip bgp [*ip-address* | **peer-group** *peer-group-name*] [**in** | **out**]	EXEC Privilege	View all information on BGP, including BGP events, keepalives, notifications, and updates.
debug ip bgp dampening [**in | out**]	EXEC Privilege	View information on BGP route being dampened.
debug ip bgp [*ip-address* | **peer-group** *peer-group-name*] **events** [**in** | **out**]	EXEC Privilege	View information on local BGP state changes and other BGP events.
debug ip bgp [*ip-address* | **peer-group** *peer-group-name*] **keepalive** [**in** | **out**]	EXEC Privilege	View information about BGP KEEPALIVE messages.
debug ip bgp [*ip-address* | **peer-group** *peer-group-name*] **notifications** [**in** | **out**]	EXEC Privilege	View information about BGP notifications received from or sent to neighbors.
debug ip bgp [*ip-address* | **peer-group** *peer-group-name*] **updates** [**in** | **out**] [**prefix-list name**]	EXEC Privilege	View information about BGP updates and filter by prefix name

Command Syntax	Command Mode	Purpose
debug ip bgp {*ip-address* | *peer-group-name*} **soft-reconfiguration**	EXEC Privilege	Enable soft-reconfiguration debug. Enable soft-reconfiguration debug. To enhance debugging of soft reconfig, use the following command only when route-refresh is not negotiated to avoid the peer from resending messages: **bgp soft-reconfig-backup** In-BGP is shown via the **show ip protocols** command.

FTOS displays debug messages on the console. To view which debugging commands are enabled, use the **show debugging** command in EXEC Privilege mode.

Use the keyword no followed by the debug command To disable a specific debug command. For example, to disable debugging of BGP updates, enter **no debug ip bgp updates** command.

Use **no debug ip bgp** to disable all BGP debugging.

Use **undebug all** to disable all debugging.

Storing Last and Bad PDUs

FTOS stores the last notification sent/received, and the last bad PDU received on per peer basis. The last bad PDU is the one that causes a notification to be issued. These PDUs are shown in the output of the command **show ip bgp neighbor**, as shown in Figure 10-34.

Figure 10-34. Viewing the Last Bad PDU from BGP Peers

```
FTOS(conf-router_bgp)#do show ip bgp neighbors 1.1.1.2

BGP neighbor is 1.1.1.2, remote AS 2, external link
  BGP version 4, remote router ID 2.4.0.1
  BGP state ESTABLISHED, in this state for 00:00:01
  Last read 00:00:00, last write 00:00:01
  Hold time is 90, keepalive interval is 30 seconds
  Received 1404 messages, 0 in queue
    3 opens, 1 notifications, 1394 updates
    6 keepalives, 0 route refresh requests
  Sent 48 messages, 0 in queue
    3 opens, 2 notifications, 0 updates
    43 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 30 seconds
  Minimum time before advertisements start is 0 seconds

  Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Capabilities advertised to neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  For address family: IPv4 Unicast
  BGP table version 1395, neighbor version 1394
  Prefixes accepted 1 (consume 4 bytes), 0 withdrawn by peer
  Prefixes advertised 0, rejected 0, 0 withdrawn from peer

  Connections established 3; dropped 2
  Last reset 00:00:12, due to Missing well known attribute

  Notification History
   'UPDATE error/Missing well-known attr' Sent : 1  Recv: 0
   'Connection Reset' Sent : 1  Recv: 0

   Last notification (len 21) sent 00:26:02 ago                    Last PDUs
    ffffffff ffffffff ffffffff ffffffff 00160303 03010000
   Last notification (len 21) received 00:26:20 ago
    ffffffff ffffffff ffffffff ffffffff 00150306 00000000
   Last PDU (len 41) received 00:26:02 ago that caused notification to be issued
    ffffffff ffffffff ffffffff ffffffff 00290200 00000e01 02040201 00024003 04141414
02186da8
    01000000
Local host: 1.1.1.1, Local port: 179
Foreign host: 1.1.1.2, Foreign port: 41758
```

Capturing PDUs

Capture incoming and outgoing PDUs on a per-peer basis using the command **capture bgp-pdu neighbor direction.** Disable capturing using the no form of this command.

The buffer size supports a maximum value between 40 MB (the default) and 100 MB. The capture buffers are cyclic and reaching the limit prompts the system to overwrite the oldest PDUs when new ones are received for a given neighbor or direction. Setting the buffer size to a value lower than the current max, might cause captured PDUs to be freed to set the new limit.

Note: Memory on RP1 is not pre-allocated, and is allocated only when a PDU needs to be captured.

Use the command **capture bgp-pdu max-buffer-size** (Figure 10-35) to change the maximum buffer size. View the captured PDUs using the command **show capture bgp-pdu neighbor**.

Figure 10-35. Viewing Captured PDUs

```
FTOS#show capture bgp-pdu neighbor 20.20.20.2

Incoming packet capture enabled for BGP neighbor 20.20.20.2
Available buffer size 40958758, 26 packet(s) captured using 680 bytes
  PDU[1] : len 101, captured 00:34:51 ago
    ffffffff ffffffff ffffffff ffffffff 00650100 00000013 00000000 00000000 419ef06c
00000000
    00000000 00000000 00000000 00000000 0181a1e4 0181a25c 41af92c0 00000000 00000000
00000000
    00000000 00000001 0181a1e4 0181a25c 41af9400 00000000
  PDU[2] : len 19, captured 00:34:51 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
  PDU[3] : len 19, captured 00:34:51 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
  PDU[4] : len 19, captured 00:34:22 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
[. . .]

Outgoing packet capture enabled for BGP neighbor 20.20.20.2
Available buffer size 40958758, 27 packet(s) captured using 562 bytes
  PDU[1] : len 41, captured 00:34:52 ago
    ffffffff ffffffff ffffffff ffffffff 00290104 000100b4 14141401 0c020a01 04000100
01020080
    00000000
  PDU[2] : len 19, captured 00:34:51 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
  PDU[3] : len 19, captured 00:34:50 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
  PDU[4] : len 19, captured 00:34:20 ago
    ffffffff ffffffff ffffffff ffffffff 00130400
[. . .]
```

The buffers storing the PDU free memory when:

- BGP is disabled
- A neighbor is unconfigured
- **clear ip bgp** is issued
- New PDU are captured and there is no more space to store them
- The max buffer size is reduced. (This may cause PDUs to be cleared depending upon the buffer space consumed and the new limit.)

With full internet feed (205K) captured, approximately 11.8MB is required to store all of the PDUs, as shown in Figure 10-36.

Figure 10-36. Required Memory for Captured PDUs

```
FTOS(conf-router_bgp)#do show capture bgp-pdu neighbor 172.30.1.250

Incoming packet capture enabled for BGP neighbor 172.30.1.250
Available buffer size 29165743, 192991 packet(s) captured using 11794257 bytes
 [. . .]

FTOS(conf-router_bgp)#do sho ip bg s
BGP router identifier 172.30.1.56, local AS number 65056
BGP table version is 313511, main routing table version 313511
207896 network entrie(s) and 207896 paths using 42364576 bytes of memory
59913 BGP path attribute entrie(s) using 2875872 bytes of memory
59910 BGP AS-PATH entrie(s) using 2679698 bytes of memory
3 BGP community entrie(s) using 81 bytes of memory

Neighbor        AS     MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx

1.1.1.2         2           17    18966          0    0     0 00:08:19 Active
172.30.1.250    18508   243295       25     313511    0     0 00:12:46    207896
```

PDU Counters

FTOS version 7.5.1.0 introduces additional counters for various types of PDUs sent and received from neighbors. These are seen in the output of the command **show ip bgp neighbor**.

Sample Configurations

The following configurations are examples for enabling BGP and setting up some peer groups. These are not comprehensive directions. They are intended to give you a some guidance with typical configurations.

You can copy and paste from these examples to your CLI. Be sure you make the necessary changes to support your own IP Addresses, Interfaces, Names, etc.

Figure 10-37 is a graphic illustration of the configurations shown on the following pages. These configurations show how to create BGP areas using physical and virtual links. They include setting up the interfaces and peers groups with each other.

Figure 10-37. Sample Configuration Illustration

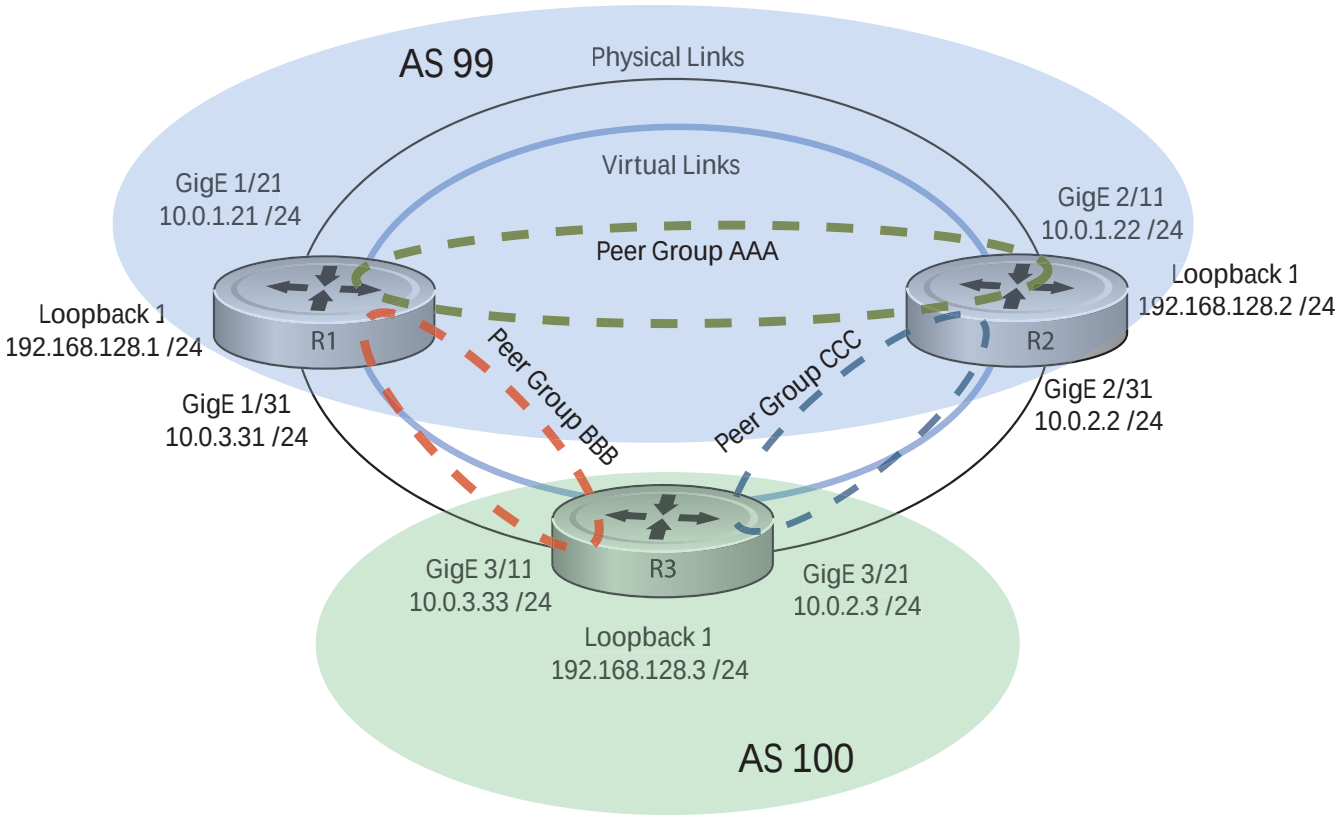


Figure 10-38. Enable BGP - Router 1

```
R1# conf
R1(conf)#int loop 0
R1(conf-if-lo-0)#ip address 192.168.128.1/24
R1(conf-if-lo-0)#no shutdown
R1(conf-if-lo-0)#show config
!
interface Loopback 0
 ip address 192.168.128.1/24
 no shutdown
R1(conf-if-lo-0)#int gig 1/21
R1(conf-if-gi-1/21)#ip address 10.0.1.21/24
R1(conf-if-gi-1/21)#no shutdown
R1(conf-if-gi-1/21)#show config
!
interface GigabitEthernet 1/21
 ip address 10.0.1.21/24
 no shutdown
R1(conf-if-gi-1/21)#int gig 1/31
R1(conf-if-gi-1/31)#ip address 10.0.3.31/24
R1(conf-if-gi-1/31)#no shutdown
R1(conf-if-gi-1/31)#show config
!
interface GigabitEthernet 1/31
 ip address 10.0.3.31/24
 no shutdown
R1(conf-if-gi-1/31)#router bgp 99
R1(conf-router_bgp)#network 192.168.128.0/24
R1(conf-router_bgp)#neighbor 192.168.128.2 remote 99
R1(conf-router_bgp)#neighbor 192.168.128.2 no shut
R1(conf-router_bgp)#neighbor 192.168.128.2 update-source loop 0
R1(conf-router_bgp)#neighbor 192.168.128.3 remote 100
R1(conf-router_bgp)#neighbor 192.168.128.3 no shut
R1(conf-router_bgp)#neighbor 192.168.128.3 update-source loop 0
R1(conf-router_bgp)#show config
!
router bgp 99
 network 192.168.128.0/24
 neighbor 192.168.128.2 remote-as 99
 neighbor 192.168.128.2 update-source Loopback 0
 neighbor 192.168.128.2 no shutdown
 neighbor 192.168.128.3 remote-as 100
 neighbor 192.168.128.3 update-source Loopback 0
 neighbor 192.168.128.3 no shutdown
R1(conf-router_bgp)#end
R1#
R1#show ip bgp summary
BGP router identifier 192.168.128.1, local AS number 99
BGP table version is 4, main routing table version 4
4 network entrie(s) using 648 bytes of memory
6 paths using 408 bytes of memory
BGP-RIB over all using 414 bytes of memory
3 BGP path attribute entrie(s) using 144 bytes of memory
2 BGP AS-PATH entrie(s) using 74 bytes of memory
2 neighbor(s) using 8672 bytes of memory

Neighbor        AS     MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx

192.168.128.2   99           4        5          4    0     0 00:00:32         1
192.168.128.3   100          5        4          1    0     0 00:00:09         4
R1#
```

Figure 10-39. Enable BGP - Router 2

```
R2# conf
R2(conf)#int loop 0
R2(conf-if-lo-0)#ip address 192.168.128.2/24
R2(conf-if-lo-0)#no shutdown
R2(conf-if-lo-0)#show config
!
interface Loopback 0
 ip address 192.168.128.2/24
 no shutdown
R2(conf-if-lo-0)#int gig 2/11
R2(conf-if-gi-2/11)#ip address 10.0.1.22/24
R2(conf-if-gi-2/11)#no shutdown
R2(conf-if-gi-2/11)#show config
!
interface GigabitEthernet 2/11
 ip address 10.0.1.22/24
 no shutdown
R2(conf-if-gi-2/11)#int gig 2/31
R2(conf-if-gi-2/31)#ip address 10.0.2.2/24
R2(conf-if-gi-2/31)#no shutdown
R2(conf-if-gi-2/31)#show config
!
interface GigabitEthernet 2/31
 ip address 10.0.2.2/24
 no shutdown
R2(conf-if-gi-2/31)#

R2(conf-if-gi-2/31)#router bgp 99
R2(conf-router_bgp)#network 192.168.128.0/24
R2(conf-router_bgp)#neighbor 192.168.128.1 remote 99
R2(conf-router_bgp)#neighbor 192.168.128.1 no shut
R2(conf-router_bgp)#neighbor 192.168.128.1 update-source loop 0
R2(conf-router_bgp)#neighbor 192.168.128.3 remote 100
R2(conf-router_bgp)#neighbor 192.168.128.3 no shut
R2(conf-router_bgp)#neighbor 192.168.128.3 update loop 0
R2(conf-router_bgp)#show config
!
router bgp 99
 bgp router-id 192.168.128.2
 network 192.168.128.0/24
 bgp graceful-restart
 neighbor 192.168.128.1 remote-as 99
 neighbor 192.168.128.1 update-source Loopback 0
 neighbor 192.168.128.1 no shutdown
 neighbor 192.168.128.3 remote-as 100
 neighbor 192.168.128.3 update-source Loopback 0
 neighbor 192.168.128.3 no shutdown
R2(conf-router_bgp)#end

R2#show ip bgp summary
BGP router identifier 192.168.128.2, local AS number 99
BGP table version is 1, main routing table version 1
1 network entrie(s) using 132 bytes of memory
3 paths using 204 bytes of memory
BGP-RIB over all using 207 bytes of memory
2 BGP path attribute entrie(s) using 128 bytes of memory
2 BGP AS-PATH entrie(s) using 90 bytes of memory
2 neighbor(s) using 9216 bytes of memory

Neighbor        AS     MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
192.168.128.1   99          40       35          1    0     0 00:01:05         1
192.168.128.3   100          4        4          1    0     0 00:00:16         1
R2#
```

Figure 10-40. Enable BGP - Router 3

```
R3# conf
R3(conf)#
R3(conf)#int loop 0
R3(conf-if-lo-0)#ip address 192.168.128.3/24
R3(conf-if-lo-0)#no shutdown
R3(conf-if-lo-0)#show config
!
interface Loopback 0
 ip address 192.168.128.3/24
 no shutdown
R3(conf-if-lo-0)#int gig 3/11
R3(conf-if-gi-3/11)#ip address 10.0.3.33/24
R3(conf-if-gi-3/11)#no shutdown
R3(conf-if-gi-3/11)#show config
!
interface GigabitEthernet 3/11
 ip address 10.0.3.33/24
 no shutdown

R3(conf-if-lo-0)#int gig 3/21
R3(conf-if-gi-3/21)#ip address 10.0.2.3/24
R3(conf-if-gi-3/21)#no shutdown
R3(conf-if-gi-3/21)#show config
!
interface GigabitEthernet 3/21
 ip address 10.0.2.3/24
 no shutdown

R3(conf-if-gi-3/21)#
R3(conf-if-gi-3/21)#router bgp 100
R3(conf-router_bgp)#show config
!
router bgp 100
R3(conf-router_bgp)#network 192.168.128.0/24
R3(conf-router_bgp)#neighbor 192.168.128.1 remote 99
R3(conf-router_bgp)#neighbor 192.168.128.1 no shut
R3(conf-router_bgp)#neighbor 192.168.128.1 update-source loop 0
R3(conf-router_bgp)#neighbor 192.168.128.2 remote 99
R3(conf-router_bgp)#neighbor 192.168.128.2 no shut
R3(conf-router_bgp)#neighbor 192.168.128.2 update loop 0
R3(conf-router_bgp)#show config
!
router bgp 100
 network 192.168.128.0/24
 neighbor 192.168.128.1 remote-as 99
 neighbor 192.168.128.1 update-source Loopback 0
 neighbor 192.168.128.1 no shutdown
 neighbor 192.168.128.2 remote-as 99
 neighbor 192.168.128.2 update-source Loopback 0
 neighbor 192.168.128.2 no shutdown
R3(conf)#end
R3#show ip bgp summary
BGP router identifier 192.168.128.3, local AS number 100
BGP table version is 1, main routing table version 1
1 network entrie(s) using 132 bytes of memory
3 paths using 204 bytes of memory
BGP-RIB over all using 207 bytes of memory
2 BGP path attribute entrie(s) using 128 bytes of memory
2 BGP AS-PATH entrie(s) using 90 bytes of memory
2 neighbor(s) using 9216 bytes of memory
Neighbor        AS     MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
192.168.128.1   99          24       25          1    0     0 00:14:20         1
192.168.128.2   99          14       14          1    0     0 00:10:22         1
```

Figure 10-41. Enable Peer Group - Router 1

```
R1#conf
R1(conf)#router bgp 99
R1(conf-router_bgp)# network 192.168.128.0/24
R1(conf-router_bgp)# neighbor AAA peer-group
R1(conf-router_bgp)# neighbor AAA no shutdown
R1(conf-router_bgp)# neighbor BBB peer-group
R1(conf-router_bgp)# neighbor BBB no shutdown
R1(conf-router_bgp)# neighbor 192.168.128.2 peer-group AAA
R1(conf-router_bgp)# neighbor 192.168.128.3 peer-group BBB
R1(conf-router_bgp)#
R1(conf-router_bgp)#show config
!
router bgp 99
 network 192.168.128.0/24
 neighbor AAA peer-group
 neighbor AAA no shutdown
 neighbor BBB peer-group
 neighbor BBB no shutdown
 neighbor 192.168.128.2 remote-as 99
 neighbor 192.168.128.2 peer-group AAA
 neighbor 192.168.128.2 update-source Loopback 0
 neighbor 192.168.128.2 no shutdown
 neighbor 192.168.128.3 remote-as 100
 neighbor 192.168.128.3 peer-group BBB
 neighbor 192.168.128.3 update-source Loopback 0
 neighbor 192.168.128.3 no shutdown
R1#
R1#show ip bgp summary
BGP router identifier 192.168.128.1, local AS number 99
BGP table version is 1, main routing table version 1
1 network entrie(s) using 132 bytes of memory
3 paths using 204 bytes of memory
BGP-RIB over all using 207 bytes of memory
2 BGP path attribute entrie(s) using 96 bytes of memory
2 BGP AS-PATH entrie(s) using 74 bytes of memory
2 neighbor(s) using 8672 bytes of memory

Neighbor        AS     MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx

192.168.128.2   99          23       24          1    0   (0) 00:00:17         1
192.168.128.3   100         30       29          1    0   (0) 00:00:14         1
!
R1#show ip bgp neighbors

BGP neighbor is 192.168.128.2, remote AS 99, internal link
  Member of peer-group AAA for session parameters
  BGP version 4, remote router ID 192.168.128.2
  BGP state ESTABLISHED, in this state for 00:00:37
  Last read 00:00:36, last write 00:00:36
  Hold time is 180, keepalive interval is 60 seconds
  Received 23 messages, 0 in queue
    2 opens, 0 notifications, 2 updates
    19 keepalives, 0 route refresh requests
  Sent 24 messages, 0 in queue
    2 opens, 1 notifications, 2 updates
    19 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 5 seconds
  Minimum time before advertisements start is 0 seconds
```

Figure 10-42. Enable Peer Groups - Router 1 continued

```
Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)
Capabilities advertised to neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Update source set to Loopback 0
  Peer active in peer-group outbound optimization
For address family: IPv4 Unicast
  BGP table version 1, neighbor version 1
  Prefixes accepted 1 (consume 4 bytes), withdrawn 0 by peer
  Prefixes advertised 1, denied 0, withdrawn 0 from peer

  Connections established 2; dropped 1
  Last reset 00:00:57, due to user reset

  Notification History
   'Connection Reset' Sent : 1  Recv: 0
Last notification (len 21) sent 00:00:57 ago
    ffffffff ffffffff ffffffff ffffffff 00150306 00000000
Local host: 192.168.128.1, Local port: 179
Foreign host: 192.168.128.2, Foreign port: 65464

BGP neighbor is 192.168.128.3, remote AS 100, external link
  Member of peer-group BBB for session parameters
  BGP version 4, remote router ID 192.168.128.3
  BGP state ESTABLISHED, in this state for 00:00:37
  Last read 00:00:36, last write 00:00:36
  Hold time is 180, keepalive interval is 60 seconds
  Received 30 messages, 0 in queue
    4 opens, 2 notifications, 4 updates
    20 keepalives, 0 route refresh requests
  Sent 29 messages, 0 in queue
    4 opens, 1 notifications, 4 updates
    20 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 30 seconds
  Minimum time before advertisements start is 0 seconds

  Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)
Update source set to Loopback 0
  Peer active in peer-group outbound optimization
For address family: IPv4 Unicast
  BGP table version 1, neighbor version 1
  Prefixes accepted 1 (consume 4 bytes), withdrawn 0 by peer
  Prefixes advertised 1, denied 0, withdrawn 0 from peer
Connections established 4; dropped 3
  Last reset 00:00:54, due to user reset
R1#
```

Figure 10-43. Enable Peer Groups - Router 2

```
R2#conf
R2(conf)#router bgp 99
R2(conf-router_bgp)# neighbor CCC peer-group
R2(conf-router_bgp)# neighbor CC no shutdown
R2(conf-router_bgp)# neighbor BBB peer-group
R2(conf-router_bgp)# neighbor BBB no shutdown
R2(conf-router_bgp)# neighbor 192.168.128.1 peer AAA
R2(conf-router_bgp)# neighbor 192.168.128.1 no shut
R2(conf-router_bgp)# neighbor 192.168.128.3 peer BBB
R2(conf-router_bgp)# neighbor 192.168.128.3 no shut
R2(conf-router_bgp)#show conf
!
router bgp 99
 network 192.168.128.0/24
 neighbor AAA peer-group
 neighbor AAA no shutdown
 neighbor BBB peer-group
 neighbor BBB no shutdown
 neighbor 192.168.128.1 remote-as 99
 neighbor 192.168.128.1 peer-group CCC
 neighbor 192.168.128.1 update-source Loopback 0
 neighbor 192.168.128.1 no shutdown
 neighbor 192.168.128.3 remote-as 100
 neighbor 192.168.128.3 peer-group BBB
 neighbor 192.168.128.3 update-source Loopback 0
 neighbor 192.168.128.3 no shutdown
R2(conf-router_bgp)#end

R2#
R2#show ip bgp summary
BGP router identifier 192.168.128.2, local AS number 99
BGP table version is 2, main routing table version 2
1 network entrie(s) using 132 bytes of memory
3 paths using 204 bytes of memory
BGP-RIB over all using 207 bytes of memory
2 BGP path attribute entrie(s) using 128 bytes of memory
2 BGP AS-PATH entrie(s) using 90 bytes of memory
2 neighbor(s) using 9216 bytes of memory

Neighbor        AS     MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx
192.168.128.1   99         140      136          2    0   (0) 00:11:24         1
192.168.128.3   100        138      140          2    0   (0) 00:18:31         1

R2#show ip bgp neighbor

BGP neighbor is 192.168.128.1, remote AS 99, internal link
  Member of peer-group AAA for session parameters
  BGP version 4, remote router ID 192.168.128.1
  BGP state ESTABLISHED, in this state for 00:11:42
  Last read 00:00:38, last write 00:00:38
  Hold time is 180, keepalive interval is 60 seconds
  Received 140 messages, 0 in queue
    6 opens, 2 notifications, 19 updates
    113 keepalives, 0 route refresh requests
  Sent 136 messages, 0 in queue
    12 opens, 3 notifications, 6 updates
    115 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 5 seconds
  Minimum time before advertisements start is 0 seconds
```

Figure 10-44. Enable Peer Group - Router 3

```
R3#conf
R3(conf)#router bgp 100
R3(conf-router_bgp)# neighbor AAA peer-group
R3(conf-router_bgp)# neighbor AAA no shutdown
R3(conf-router_bgp)# neighbor CCC peer-group
R3(conf-router_bgp)# neighbor CCC no shutdown
R3(conf-router_bgp)# neighbor 192.168.128.2 peer-group BBB
R3(conf-router_bgp)# neighbor 192.168.128.2 no shutdown
R3(conf-router_bgp)# neighbor 192.168.128.1 peer-group BBB
R3(conf-router_bgp)# neighbor 192.168.128.1 no shutdown
R3(conf-router_bgp)#

R3(conf-router_bgp)#end

R3#show ip bgp summary
BGP router identifier 192.168.128.3, local AS number 100
BGP table version is 1, main routing table version 1
1 network entrie(s) using 132 bytes of memory
3 paths using 204 bytes of memory
BGP-RIB over all using 207 bytes of memory
2 BGP path attribute entrie(s) using 128 bytes of memory
2 BGP AS-PATH entrie(s) using 90 bytes of memory
2 neighbor(s) using 9216 bytes of memory

Neighbor        AS     MsgRcvd  MsgSent     TblVer  InQ  OutQ Up/Down  State/Pfx

192.168.128.1   99          93       99          1    0   (0) 00:00:15         1
192.168.128.2   99         122      120          1    0   (0) 00:00:11         1
R3#show ip bgp neighbor

BGP neighbor is 192.168.128.1, remote AS 99, external link
  Member of peer-group BBB for session parameters
  BGP version 4, remote router ID 192.168.128.1
  BGP state ESTABLISHED, in this state for 00:00:21
  Last read 00:00:09, last write 00:00:08
  Hold time is 180, keepalive interval is 60 seconds
  Received 93 messages, 0 in queue
    5 opens, 0 notifications, 5 updates
    83 keepalives, 0 route refresh requests
  Sent 99 messages, 0 in queue
    5 opens, 4 notifications, 5 updates
    85 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 30 seconds
  Minimum time before advertisements start is 0 seconds

  Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Capabilities advertised to neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Update source set to Loopback 0
  Peer active in peer-group outbound optimization

  For address family: IPv4 Unicast
  BGP table version 1, neighbor version 1
  Prefixes accepted 1 (consume 4 bytes), withdrawn 0 by peer
  Prefixes advertised 1, denied 0, withdrawn 0 from peer
```

Figure 10-45. Enable Peer Groups - Router 3 continued

```
Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Capabilities advertised to neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Update source set to Loopback 0
  Peer active in peer-group outbound optimization

  For address family: IPv4 Unicast
  BGP table version 2, neighbor version 2
  Prefixes accepted 1 (consume 4 bytes), withdrawn 0 by peer
  Prefixes advertised 1, denied 0, withdrawn 0 from peer

  Connections established 6; dropped 5
  Last reset 00:12:01, due to Closed by neighbor

  Notification History
   'HOLD error/Timer expired' Sent : 1  Recv: 0
   'Connection Reset' Sent : 2  Recv: 2

   Last notification (len 21) received 00:12:01 ago
    ffffffff ffffffff ffffffff ffffffff 00150306 00000000
Local host: 192.168.128.2, Local port: 65464
Foreign host: 192.168.128.1, Foreign port: 179

BGP neighbor is 192.168.128.3, remote AS 100, external link
  Member of peer-group BBB for session parameters
  BGP version 4, remote router ID 192.168.128.3
  BGP state ESTABLISHED, in this state for 00:18:51
  Last read 00:00:45, last write 00:00:44
  Hold time is 180, keepalive interval is 60 seconds
  Received 138 messages, 0 in queue
    7 opens, 2 notifications, 7 updates
    122 keepalives, 0 route refresh requests
  Sent 140 messages, 0 in queue
    7 opens, 4 notifications, 7 updates
    122 keepalives, 0 route refresh requests
  Minimum time between advertisement runs is 30 seconds
  Minimum time before advertisements start is 0 seconds
Capabilities advertised to neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
  Capabilities received from neighbor for IPv4 Unicast :
    MULTIPROTO_EXT(1)
    ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)
ROUTE_REFRESH(2)
    CISCO_ROUTE_REFRESH(128)

  Update source set to Loopback 0
  Peer active in peer-group outbound optimization
For address family: IPv4 Unicast
  BGP table version 2, neighbor version 2
  Prefixes accepted 1 (consume 4 bytes), withdrawn 0 by peer
  Prefixes advertised 1, denied 0, withdrawn 0 from peer
```

11

Content Addressable Memory

Content Addressable Memory is supported on platforms C E T S

Note: Different platforms support varying levels of CAM adjustment. Be sure to read this chapter carefully prior to changing any CAM parameters.

Content Addressable Memory

Content Addressable Memory (CAM) is a type of memory that stores information in the form of a lookup table. On Dell Force10 systems, the CAM stores Layer 2 and Layer 3 forwarding information, access-lists (ACL), flows, and routing policies. On Dell Force10 systems, there are one or two CAM (Dual-CAM) modules per port-pipe depending on the type of line card.

- The ExaScale EH and EJ series line cards are single-CAM line cards that support 10M and 40M CAM for storing the lookup information.
- The TeraScale EG-series line cards are dual-CAM and use two 18 Megabit CAM modules with a dedicated 512 IPv4 Forwarding Information Base (FIB), and flexible CAM allocations for Layer2, FIB, and ACLs.
- Either ExaScale 10G or 40G CAM line cards can be used in a system.

CAM Profiles

Dell Force10 systems partition each CAM module so that it can store the different types of information. The size of each partition is specified in the CAM profile. A CAM profile is stored on every card, including each RPM. The same profile must be on every line card and RPM in the chassis.

There is a default CAM profile and several other CAM profiles available so that you can partition the CAM according to your performance requirements. For example, the default profile has 1K Layer 2 ingress ACL entries. If you need more memory for Layer 2 ingress ACLs, select the profile *l2-ipv4-inacl*.

Table 11-1 describes the available profiles. The default profile is an all-purpose profile that allocates CAM space according to the way Dell Force10 systems are most commonly used. In general, non-default profiles allocate more space to particular regions to accommodate specific applications. The size of CAM partitions is measured in entries. The total CAM space is finite, therefor adding entries to one region necessarily decreases the number available to other regions.

Note: Not all CAM profiles and microcodes are available for all systems. Refer to the Command Line Interface Reference Guide for details regarding available profiles for each system.

Table 11-1. CAM Profile Descriptions

CAM Profile	Description
Default	An all-purpose profile that allocates CAM space according to the way Dell Force10 systems are most commonly used. Available Microcodes: default, lag-hash-align, lag-hash-mpls, l2-switched-pbr
eg-default	For EG-series line cards only. EG series line cards have two CAM modules per Port-pipe. Available Microcodes: default, ipv6-extacl
ipv4-320k	Provides 320K entries for the IPv4 Forwarding Information Base (FIB) and reduces the IPv4 Flow partition to 12K. Available Microcodes: default, lag-hash-mpls, l2-switched-pbr
ipv4-egacl-16k	Provides 16K entries for egress ACLs Available Microcodes: acl-group
ipv6-extacl	Provides IPv6 functionality. Available Microcodes: ipv6-extacl
l2-ipv4-inacl	Provides 32K entries for Layer 2 ingress ACLs and 28K entries for Layer 3 IPv4 ingress ACLs. Available Microcodes: default
unified-default	Maintains the CAM allocations for the and IPv4 FIB while allocating more CAM space for the Ingress and Egress Layer 2 ACL, and IPv4 ACL regions. Available Microcodes: ipv6-extacl
ipv4-VRF	Provides VRF functionality for IPv4. Available Microcodes:ipv4-vrf
ipv4-v6-VRF	Provides VRF functionality for both IPv4 and I.Pv6 Available Microcodes: ipv4-v6-vrf
ipv4-64k-ipv6	Provides IPv6 functionality; an alternate to ipv6-extacl that redistributes CAM space from the IPv4FIB to IPv4Flow and IPv6FIB. Available Microcodes: ipv6-extacl

The size of CAM partitions is measured in entries. Table 11-1 shows the number of entries available in each partition for all CAM profiles. The total CAM space is finite, therefor adding entries to one region necessarily decreases the number available to other regions.

Table 11-2. CAM entries per partition

Profile / Partition	L2FIB	L2ACL	IPv4FIB	IPv4ACL	IPv4Flow	EgL2ACL	EgIPv4ACL	Reserved	IPv6FIB	IPv6ACL	IPv6Flow	EgIPv6ACL
Default	32K	2K	256K	12K	24K	1K	1K	8K	0	0	0	0
eg-default	32K	2K	512K	12K	24K	1K	1K	8K	32K	3K	4K	1K
ipv4-320k	32K	2K	320K	12K	12K	1K	1K	4K	0	0	0	0
pv4-egacl-16k	32K	2K	192K	8K	24K	0	16K	8K	0	0	0	0
ipv6-extacl	32K	2K	192K	12K	8K	1K	1K	2K	6K	3K	4K	2K
l2-ipv4-inacl	32K	33K	64K	27K	8K	2K	2K	2K	0	0	0	0
unified-default	32K	3K	192K	9K	8K	2K	2K	2K	6K	2K	4K	2K
IPv4-VRF	32K	3K	160K	2K	12K	1K	12K	2K	0	0	0	0
IPv4-v6-VRF	32K	3K	64K	1K	12K	1K	11K	2K	18K	4K	3K	1K
ipv4-64k-ipv6	32K	2K	64K	12K	24K	1K	1K	8K	16K	3K	4K	1K

Microcode

Microcode is a compiled set of instructions for a CPU. On Dell Force10 systems, the microcode controls how packets are handled.

There is a default microcode, and several other microcodes are available, so that you can adjust packet handling according to your application. Specifying a microcode is mandatory when selecting a CAM profile (though you are not required to change it).

Note: Not all CAM profiles and microcodes are available for all systems. Refer to the Command Line Interface Reference Guide for details regarding available profiles for each system.

Table 11-3. Microcode Descriptions

Microcode	Description
default	Distributes CAM space for a typical deployment
lag-hash-align	For applications that require the same hashing for bi-directional traffic (for example, VoIP call or P2P file sharing). For port-channels, this microcode maps both directions of a bi-directional flow to the same output link.

Table 11-3. Microcode Descriptions

Microcode	Description
lag-hash-mpls	For hashing based on MPLS labels (up to five labels deep). With the default microcode, MPLS packets are distributed over a port-channel based on the MAC source and destination address. With the lag-hash-mpls microcode, MPLS packets are distributed across the port-channel based on IP source and destination address and IP protocol. This is applicable for MPLS packets with up to five labels. When the IP header is not available after the 5th label, hashing for default load-balance is based on MPLS labels. For packets with more than 5 labels, hashing is always based on the MAC source and destination address.
ipv6-extacl	Use this microcode when IPv6 is enabled.
acl-group	For applications that need 16k egress IPv4 ACLs (for example, the VLAN ACL Group feature, which permits group VLANs IP egress ACLs.
ipv4-vrf	Apply to IPv4 VRF CAM profile.
ipv4-v6-vrf	Enable IPv4 and IPv6 CAM profiles for VRF.
l2-switched-pbr	**E-Series TeraScale only**: If you apply a PBR redirect list (using the **ip re-direct group** command) to a VLAN interface, Layer 2 traffic is redirected and dropped by default. To avoid having Layer 2 traffic affected by PBR, configure a CAM profile that supports l2-switched-pbr (IPv4-LDA) microcode. l2-switched-pbr microcode allows only Layer 3 traffic to be redirected while Layer 2 traffic is switched within the VLAN.

CAM Profiling for ACLs

CAM Profiling for ACLs is supported on platform E(T) only.

Refer to Content Addressable Memory for ExaScale for E-Series ExaScale E(X) CAM descriptions.

The default CAM profile has 1K Layer 2 ingress ACL entries. If you need more memory for Layer 2 ingress ACLs, select the profile *l2-ipv4-inacl*.

When budgeting your CAM allocations for ACLs and QoS configurations, remember that ACL and QoS rules might consume more than one CAM entry depending on complexity. For example, TCP and UDP rules with port range options might require more than one CAM entry.

The Layer 2 ACL CAM partition has sub-partitions for several types of information. Table 11-4 lists the sub-partition and the percentage of the Layer 2 ACL CAM partition that FTOS allocates to each by default.

Table 11-4. Layer 2 ACL CAM Sub-partition Sizes

Partition	% Allocated
Sysflow	6
L2ACL	14
*PVST	50

Table 11-4. Layer 2 ACL CAM Sub-partition Sizes

Partition	% Allocated
QoS	12
L2PT	13
FRRP	5

You can re-configure the amount of space, in percentage, allocated to each sub-partition. As with the IPv4Flow partition, you can configure the Layer 2 ACL partition from EXEC Privilege mode or CONFIGURATION mode.

The amount of space that you can distribute to the sub-partitions is equal to the amount of CAM space that the selected CAM profile allocates to the Layer 2 ACL partition. FTOS requires that you specify the amount of CAM space for all sub-partitions and that the sum of all sub-partitions is 100%. FTOS displays the following message if the total allocated space is not correct:

```
% Error: Sum of all regions does not total to 100%.
```

Boot Behavior

The profile and microcode loaded on the primary RPM determines the profile and microcode that is required on all other chassis components and is called the “chassis profile.” A profile mismatch condition exists if either the CAM profile or the microcode does not match. The following points describe line card boot behavior when the line card profile does not match the chassis profile.

- A microcode mismatch constitutes a profile mismatch.
- When the line card profile and chassis profile are of the same type (single-CAM or dual-CAM), but their CAM profiles do not match, the line card must load a new profile and therefore takes longer to come online.
- If you insert a single-CAM line card into a chassis with a dual-CAM profile, the system displays Message 1. The line card boots with the default (single-CAM) profile and remains in a problem state (Figure 11-1). The line card cannot forward traffic in a problem state.
- If you insert a dual-CAM line card into a chassis with a single-CAM profile, the line card boots with a matching profile, but operates with a lower capability.

Message 1 EF Line Card with EG Chassis Profile Error

```
# Before reload:
01:09:56: %RPM0-P:CP %CHMGR-4-EG_PROFILE_WARN: If EG CAM profile is selected, non-EG cards
will be in problem state after reload
# After reload:
00:04:46: %RPM0-P:CP %CHMGR-3-PROFILE_MISMATCH: Mismatch: line card 1 has mismatch CAM
profile or microcode
```

Message 2 EH Line Card with EG Chassis Profile Error

```
# Before reload:
01:09:56: %RPM0-P:CP %CHMGR-4-EH_PROFILE_WARN: If EH CAM profile is selected, non-EJ cards
will be in problem state after reload
# After reload:
00:04:46: %RPM0-P:CP %CHMGR-3-PROFILE_MISMATCH: Mismatch: line card 1 has mismatch CAM
profile or microcode
```

Figure 11-1. EF Line Card with EG Chassis Profile—Card Problem

```
R1#show linecard 1 brief

--  Line card 1 --
Status        : card problem - mismatch cam profile
Next Boot     : online
Required Type : E48TF - 48-port 10/100/1000Base-T line card with RJ-45 interfaces (EF)
Current Type  : E48TF - 48-port 10/100/1000Base-T line card with RJ-45 interfaces (EF)
Hardware Rev  : Base  - 1.1  PP0 - 1.1  PP1 - 1.1
Num Ports     : 48
Up Time       : 0 sec
FTOS Version  : 7.6.1.0
Jumbo Capable : yes
```

Figure 11-2. EH Line Card with EG Chassis Profile—Card Problem

```
R1#show linecard 1 brief

--  Line card 1 --
Status        : card problem - mismatch cam profile
Next Boot     : online
Required Type : E90MH - 90-port 10/100/1000Base-T line card with mini RJ-21 interfaces (EH)
Current Type  : E90MH - 90-port 10/100/1000Base-T line card with mini RJ-21 interfaces (EH)
Hardware Rev  : Base  - 0.3 PP0 - 1.1  PP0 - PP1 -
Num Ports     : 90
Up Time       : 0 sec
FTOS Version  : 8.1.1.0
Jumbo Capable : yes
```

When to Use CAM Profiling

The CAM profiling feature enables you to partition the CAM to best suit your application. For example:

- Configure more Layer 2 FIB entries when the system is deployed as a switch.
- Configure more Layer 3 FIB entries when the system is deployed as a router.
- Configure more ACLs (when IPv6 is not employed).
- Hash MPLS packets based on source and destination IP addresses for LAGs. See LAG Hashing on page 298.
- Hash based on bidirectional flow for LAGs. See LAG Hashing based on Bidirectional Flow on page 299.

- Optimize the VLAN ACL Group feature, which permits group VLANs for IP egress ACLs. See CAM profile for the VLAN ACL group feature on page 299.

Important Points to Remember

- CAM Profiling is available on the E-Series TeraScale with FTOS versions 6.3.1.1 and later.
- All line cards within a single system must have the same CAM profile; this profile must match the system CAM profile (the profile on the primary RPM).
  - FTOS automatically reconfigures the CAM profile on line cards and the secondary RPM to match the system CAM profile by saving the correct profile on the card and then rebooting it.
- The CAM configuration is applied to entire system when you use CONFIGURATION mode commands. You must save the running-configuration to affect the change.
- All CAM configuration commands require you to reboot the system.
- When budgeting your CAM allocations for ACLs and QoS configurations, remember that ACL and QoS rules might consume more than one CAM entry depending on complexity. For example, TCP and UDP rules with port range options might require more than one CAM entry. See Pre-calculating Available QoS CAM Space on page 874.
- After you install a secondary RPM, copy the running-configuration to the startup-configuration so that the new RPM has the correct CAM profile.

Differences Between EtherScale and TeraScale

- Only one CAM profile and microcode is available on EtherScale systems.
- Only EtherScale systems can sub-partition the IPv4ACL partition.
- Both EtherScale and TeraScale systems can sub-partition the IPv4Flow CAM partition.

Select CAM Profiles

A CAM profile is selected in CONFIGURATION mode. The CAM profile is applied to entire system, however, you must save the running-configuration to affect the change.

All components in the chassis must have the same CAM profile and microcode. The profile and microcode loaded on the primary RPM determines the profile that is required on all other chassis components.

- If a newly installed line card has a profile different from the primary RPM, the card reboots so that it can load the proper profile.
- If a the standby RPM has a profile different from the primary RPM, the card reboots so that it can load the proper profile.

To change the CAM profile on the entire system:

Step	Task	Command Syntax	Command Mode
1	Select a CAM profile.	**cam-profile** *profile* **microcode** *microcode*	CONFIGURATION
	Note: If selecting a cam-profile for VRF (**cam-profile ipv4-vrf or ipv4-v6-vrf**), implement the command in the CONFIGURATION mode only. If you use EXEC Privilege mode, the linecards may go into an error state.		
2	Save the running-configuration.	**copy running-config startup-config**	EXEC Privilege
3	Verify that the new CAM profile will be written to the CAM on the next boot.	**show cam-profile summary**	EXEC Privilege
4	Reload the system.	**reload**	EXEC Privilege

CAM Allocation

User Configurable CAM Allocations is available on platforms: C S

Allocate space for IPV4 ACLs and QoS regions, and IPv6 6 ACLs and QoS regions on the C-Series and S-Series by using the **cam-acl** command in CONFIGURATION mode.

The CAM space is allotted in FP blocks. The total space allocated must equal 13 FP blocks. The default CAM Allocation settings on a C-Series system are:

- L3 ACL (ipv4acl): 5
- L2 ACL(l2acl) : 6
- IPv6 L3 ACL (ipv6acl): 0
- L3 QoS (ipv4qos): 1
- L2 QoS (l2qos): 1
- L2PT (l2pt): 0
- MAC ACLs (ipmacacl): 0
- ECFMACL (ecfmacl): 0
- VMAN QoS (vman-qos): 0
- VMAN Dual QoS (vman-dual-qos): 0

Note: The ipmacacl region was introduced for Secure DHCP. These ACL are not created through CLI, but rather are system generated from the DHCP snooping table. Whenever a new DHCP client is assigned an IP, and ip dhcp snooping source-address-validation ipmac is configured on the interface connected to the client, a single ACL is installed on the interface to permit (only) the source IP and source MAC pair.

The **ipv6acl** and **vman-dual-qos** allocations must be entered as a factor of 2 (2, 4, 6, 8, 10). All other profile allocations can use either even or odd numbered ranges.

You must save the new CAM settings to the startup-config (**write-mem or copy run start**) then reload the system for the new settings to take effect.

To configure the IPv4 and IPv6 ACLs and Qos regions on the entire system:

Step	Task	Command Syntax	Command Mode
1	Select a cam-acl action	**cam-acl** [**default** | **l2acl**]	CONFIGURATION
	Note: Selecting **default** resets the CAM entries to the default settings. Select **l2acl** to allocate space for the ACLs, and QoS regions.		
2	Enter the number of FP blocks for each region.	**l2acl** *number* **ipv4acl** *number* **ipv6acl** *number*, **ipv4qos** *number* **l2qos** *number*, **l2pt** *number* **ipmacacl** *number* **ecfmacl** *number* [**vman-qos** | **vman-dual-qos** *number*	EXEC Privilege
3	Verify that the new settings will be written to the CAM on the next boot.	**show cam-acl**	EXEC Privilege
4	Reload the system.	**reload**	EXEC Privilege

Test CAM Usage

The **test cam-usage** command is supported on platforms C E S

This command applies to both IPv4 and IPv6 CAM profiles, but is best used when verifying QoS optimization for IPv6 ACLs.

Use this command to determine whether sufficient ACL CAM space is available to enable a service-policy. Create a Class Map with all required ACL rules, then execute the **test cam-usage** command in Privilege mode to verify the actual CAM space required. Figure 11-3 gives a sample of the output shown when executing the command. The status column indicates whether or not the policy can be enabled.

Figure 11-3. Command Example: test cam-usage (C-Series)

```
FTOS#test cam-usage service-policy input TestPolicy linecard all

Linecard | Portpipe | CAM Partition | Available CAM | Estimated CAM per Port | Status
--------------------------------------------------------------------------------
       2 |        1 | IPv4Flow      |           232 |                      0 | Allowed
       2 |        1 | IPv6Flow      |             0 |                      0 | Allowed
       4 |        0 | IPv4Flow      |           232 |                      0 | Allowed
       4 |        0 | IPv6Flow      |             0 |                      0 | Allowed
FTOS#
```

View CAM Profiles

View the current CAM profile for the chassis and each component using the command **show cam-profile**, as shown in Figure 11-4. This command also shows the profile that will be loaded upon the next chassis or component reload.

Figure 11-4. Viewing CAM Profiles on E-Series TeraScale

```
FTOS#show cam-profile

-- Chassis CAM Profile --

CamSize          : 18-Meg
                 : Current Settings : Next Boot
Profile Name     : Default          : Default
L2FIB            : 32K entries      : 32K entries
L2ACL            : 1K entries       : 1K entries
IPv4FIB          : 256K entries     : 256K entries
IPv4ACL          : 12K entries      : 12K entries
IPv4Flow         : 24K entries      : 24K entries
EgL2ACL          : 1K entries       : 1K entries
EgIPv4ACL        : 1K entries       : 1K entries
Reserved         : 8K entries       : 8K entries
FIB          : 0  entries       : 0  entries
ACL          : 0  entries       : 0  entries
Flow         : 0  entries       : 0  entries
EgACL        : 0  entries       : 0  entries
MicroCode Name   : Default          : Default
--More--
```

View a brief output of the command **show cam-profile** using the **summary** option.

The command **show running-config cam-profile** shows the current profile and microcode (Figure 11-5).

Note: If you select the CAM profile from CONFIGURATION mode, the output of this command does not reflect any changes until you save the running-configuration and reload the chassis.

Figure 11-5. Viewing CAM Profile Information in the Running-configuration

```
FTOS#show running-config cam-profile
!
cam-profile default microcode default

FTOS#
```

View CAM-ACL settings

View the current cam-acl settings for the C-Series and S-Series systems chassis and each component using the command **show cam-acl**, as shown in Figure 11-6.

Figure 11-6. View CAM-ACl settings on C-Series and S-Series

```
FTOS# show cam-acl

-- Chassis Cam ACL --
             Current Settings(in block sizes)
L2Acl        :         2
Ipv4Acl      :         2
Ipv6Acl      :         2
Ipv4Qos      :         2
L2Qos        :         2
L2PT         :         1
IpMacAcl     :         2
VmanQos      :         0
VmanDualQos  :         0

-- Line card 0 --
           Current Settings(in block sizes)
L2Acl        :         2
Ipv4Acl      :         2
Ipv6Acl      :         2
Ipv4Qos      :         2
L2Qos        :         2
L2PT         :         1
IpMacAcl     :         2
VmanQos      :         0
VmanDualQos  :         0

-- Line card 6 --
           Current Settings(in block sizes)
L2Acl        :         2
Ipv4Acl      :         2
Ipv6Acl      :         2
Ipv4Qos      :         2
L2Qos        :         2
L2PT         :         1
IpMacAcl     :         2
VmanQos      :         0
VmanDualQos  :         0
```

View CAM Usage

View the amount of CAM space available, used, and remaining in each partition (including IPv4Flow and Layer 2 ACL sub-partitions) using the command **show cam-usage** from EXEC Privilege mode, as shown in Figure 11-7.

Figure 11-7. Viewing CAM Usage Information

```
R1#show cam-usage
Linecard|Portpipe| CAM Partition   | Total CAM   |  Used CAM   |Available CAM
========|========|=================|=============|=============|==============
    1   |   0    | IN-L2 ACL       |      1008   |       320   |       688
        |        | IN-L2 FIB       |     32768   |      1132   |     31636
        |        | IN-L3 ACL       |     12288   |         2   |     12286
        |        | IN-L3 FIB       |    262141   |        14   |    262127
        |        | IN-L3-SysFlow   |      2878   |        45   |      2833
        |        | IN-L3-TrcList   |      1024   |         0   |      1024
        |        | IN-L3-McastFib  |      9215   |         0   |      9215
        |        | IN-L3-Qos       |      8192   |         0   |      8192
        |        | IN-L3-PBR       |      1024   |         0   |      1024
        |        | IN-V6 ACL       |         0   |         0   |         0
        |        | IN-V6 FIB       |         0   |         0   |         0
        |        | IN-V6-SysFlow   |         0   |         0   |         0
        |        | IN-V6-McastFib  |         0   |         0   |         0
        |        | OUT-L2 ACL      |      1024   |         0   |      1024
        |        | OUT-L3 ACL      |      1024   |         0   |      1024
        |        | OUT-V6 ACL      |         0   |         0   |         0
    1   |   1    | IN-L2 ACL       |       320   |         0   |       320
        |        | IN-L2 FIB       |     32768   |      1136   |     31632
        |        | IN-L3 ACL       |     12288   |         2   |     12286
        |        | IN-L3 FIB       |    262141   |        14   |    262127
        |        | IN-L3-SysFlow   |      2878   |        44   |      2834
--More--
```

Configure IPv4Flow Sub-partitions

IPv4Flow sub-partitions are supported on platform E

The IPv4Flow CAM partition has sub-partitions for several types of information. Table 11-5 lists the types of information stored in this partition and the number of entries that FTOS allocates to each type.

Table 11-5. IPv4Flow CAM Sub-partition Sizes

Partition	Space Allocated (EtherScale)	Space Allocated (TeraScale)	Space Allocated (ExaScale)
ACL	8K	—	—
Multicast FIB/ACL	9K	3K	3K
PBR	1K	1K	1K
QoS	8K	2K	2K
System Flow	5K	5K	5K
Trace Lists	1	1K	1K

You can re-configure the amount of space allocated for each type of entry. FTOS requires that you specify an amount of CAM space for all types and in the order shown in Table 11-5.

- The IPv4Flow configuration is applied to entire system when you enter the command **cam-ipv4flow** from CONFIGURATION mode, however, you must save the running-configuration to affect the change.

The amount of space that is allocated among the sub-partitions must be equal to the amount of CAM space allocated to IPv4Flow by the selected CAM profile (see Table 11-1.); Message 3 is displayed if the total allocated space is not correct.

Message 3 IPv4Flow Configuration Error

```
% Error: Total size must add up to match IPv4flow size of 24K required by the configured
profile.
```

The minimum amount of space that can be allocated to any sub-partition is 1K, except for System flow, for which the minimum is 4K.

To re-allocate CAM space within the IPv4Flow partition on the entire system:

Step	Task	Command Syntax	Command Mode
1	Re-allocate CAM space within the IPv4Flow partition.	**cam-ipv4flow**	CONFIGURATION
2	Save the running-configuration.	**copy running-config startup-config**	EXEC Privilege
3	Verify that the new CAM configuration will be written to the CAM on the next boot.	**show cam-ipv4flow**	EXEC Privilege
4	Reload the system.	**reload**	EXEC Privilege

Figure 11-8. Configuring IPv4Flow on the Entire System

```
FTOS(conf)#cam-ipv4flow default
FTOS#copy running-config startup-config
File with same name already exist.
Proceed to copy the file [confirm yes/no]: yes
!
3914 bytes successfully copied

FTOS#sh cam-ipv4flow
-- Chassis Cam Ipv4Flow --
                     Current Settings   Next Boot
Multicast Fib/Acl :   8K                 9K
Pbr               :   2K                 1K
Qos               :   7K                 8K
System Flow       :   6K                 5K
Trace Lists       :   1K                 1K

-- Line card 0 --

                     Current Settings   Next Boot
Multicast Fib/Acl :   8K                 9K
Pbr               :   2K                 1K
Qos               :   7K                 8K
System Flow       :   6K                 5K
Trace Lists       :   1K                 1K

-- Line card 1 --
                     Current Settings   Next Boot

Multicast Fib/Acl :   8K                 9K
Pbr               :   2K                 1K
Qos               :   7K                 8K
System Flow       :   6K                 5K
Trace Lists       :   1K                 1K
```

Configure Ingress Layer 2 ACL Sub-partitions

IPv4Flow sub-partitions are supported on platform E

The Ingress Layer 2 ACL CAM partition has sub-partitions for several types of information. Table 11-6 lists the sub-partition and the percentage of the Ingress Layer 2 ACL CAM partition that FTOS allocates to each by default.

Table 11-6. Layer 2 ACL CAM Sub-partition Sizes

Partition	% Allocated
Sysflow	6
L2ACL	14
*PVST	50
QoS	12

Table 11-6. Layer 2 ACL CAM Sub-partition Sizes (continued)

Partition	% Allocated
L2PT	13
FRRP	5

You can re-configure the amount of space, in percentage, allocated to each sub-partition.

- Apply the Ingress Layer 2 ACL configuration to entire system by entering the command **cam-l2acl** from CONFIGURATION mode, however, you must save the running-configuration to affect the change.

The amount of space that you can distribute to the sub-partitions is equal to the amount of CAM space that the selected CAM profile allocates to the Ingress Layer 2 ACL partition (see Table 11-1). FTOS requires that you specify the amount of CAM space for all sub-partitions and that the sum of all sub-partitions is 100%. FTOS displays message Message 4 if the total allocated space is not correct.

Message 4 Layer 2 ACL Configuration Error

% Error: Sum of all regions does not total to 100%.

Note: You must allocate at least (<*number of VLANs*> * <*Number of switching ports per port-pipe*>) entries at least when employing PVST+ . For example, the default CAM Profile allocates 1000 entries to the Ingress Layer 2 ACL CAM region, and a 48-port linecard has two port-pipes with 24 ports each. If you have 5 VLANs, then you must allocate at least 120 (5*24) entries to the PVST Ingress Layer 2 ACL CAM region, which is 12% of the total 1000 available entries.

To re-allocate CAM space within the Ingress Layer 2 ACL partition on the entire system (Figure 11-9):

Step	Task	Command Syntax	Command Mode
1	Re-allocate CAM space within the Ingress Layer 2 ACL partition.	**cam-l2acl**	CONFIGURATION
2	Save the running-configuration.	**copy running-config startup-config**	EXEC Privilege
3	Verify that FTOS will write the new CAM configuration to the CAM on the next boot.	**show cam-l2acl**	EXEC Privilege
4	Reload the system.	**reload**	EXEC Privilege

Figure 11-9. Configuring Ingress Layer 2 ACL on the Entire System

```
FTOS(conf)#do show cam-l2acl | find “Line card 1”
-- Line card 1 --
          Current Settings(in percent)
Sysflow  :         6
L2Acl    :        14
Pvst     :        50
Qos      :        12
L2pt     :        13
Frrp     :         5

[output omitted]
FTOS(conf)#cam-l2acl system-flow 100 l2acl 0 p 0 q 0 l 0 f 0
FTOS(conf)#do show cam-l2acl | find “Line card 1”
-- Line card 1 --
          Current Settings(in percent)
Sysflow  :         6
L2Acl    :        14
Pvst     :        50
Qos      :        12
L2pt     :        13
Frrp     :         5

[output omitted]

FTOS(conf)#do copy run start
File with same name already exist.
Proceed to copy the file [confirm yes/no]: yes
!
8676 bytes successfully copied
02:00:49: %RPM0-P:CP %FILEMGR-5-FILESAVED: Copied running-config to startup-config in flash by
default
FTOS(conf)#do show cam-l2acl | find “Line card 1”
-- Line card 1 --
          Current Settings(in percent)  Next Boot(in percent)
Sysflow  :         6                          100
L2Acl    :        14                            5
Pvst     :        50                            5
Qos      :        12                            5
L2pt     :        13                            5
Frrp     :         5                            5
```

Return to the Default CAM Configuration

Return to the default CAM Profile, microcode, IPv4Flow, or Layer 2 ACL configuration using the keyword **default** from EXEC Privilege mode or from CONFIGURATION mode, as shown in Figure 11-10.

Figure 11-10. Returning to the default Configuration

```
FTOS(conf)#cam-profile ?
default                 Enable default CAM profile
eg-default              Enable eg-default CAM profile
ipv4-320k               Enable 320K CAM profile
ipv4-egacl-16k          Enable CAM profile with 16K IPv4 egress ACL
ipv6-extacl             Enable CAM profile with  extended ACL
l2-ipv4-inacl           Enable CAM profile with 32K L2 and 28K IPv4 ingress ACL
unified-default         Enable default unified CAM profile
FTOS(conf)#cam-profile default microcode ?
default                 Enable default microcode
lag-hash-align          Enable microcode with LAG hash align
lag-hash-mpls           Enable microcode with LAG hash MPLS
FTOS(conf)#cam-profile default microcode default
FTOS(conf)#cam-ipv4flow ?
default                 Reset IPv4flow CAM entries to default setting
multicast-fib           Set multicast FIB entries
FTOS(conf)#cam-l2acl ?
default                 Reset L2-ACL CAM entries to default setting
system-flow             Set system flow entries
```

CAM Optimization

CAM optimization is supported on platforms C S

When this command is enabled, if a Policy Map containing classification rules (ACL and/or dscp/ip-precedence rules) is applied to more than one physical interface on the same port-pipe, only a single copy of the policy is written (only 1 FP entry will be used). When the command is disabled, the system behaves as described in this chapter.

Applications for CAM Profiling

LAG Hashing

FTOS includes a CAM profile and microcode that treats MPLS packets as non-IP packets. Normally, switching and LAG hashing is based on source and destination MAC addresses. Alternatively, you can base LAG hashing for MPLS packets on source and destination IP addresses. This type of hashing is allowed for MPLS packets with 5 labels or less.

MPLS packets are treated as follows:

- When MPLS IP packets are received, FTOS looks up to 5 labels deep for the IP header.
- When an IP header is present, hashing is based on IP 3 tuple (source IP address, destination IP address, and IP protocol).
- If an IP header is not found after the 5th label, hashing is based on the MPLS labels.

- If the packet has more than 5 MPLS labels, hashing is based on the source and destination MAC address.

To enable this type of hashing, use the default CAM profile with the microcode *lag-hash-mpls*.

LAG Hashing based on Bidirectional Flow

To hash LAG packets such that both directions of a bidirectional flow (for example, VoIP or P2P file sharing) are mapped to the same output link in the LAG bundle, use the default CAM profile with the microcode *lag-hash-align*.

CAM profile for the VLAN ACL group feature

IPv4Flow sub-partitions are supported on platform E-Series only.

To optimize for the VLAN ACL Group feature, which permits group VLANs for the IP egress ACL, use the CAM profile *ipv4-egacl-16k* with the default microcode.

Note: Do not use this CAM profile for Layer 2 egress ACLs.

Troubleshoot CAM Profiling

CAM Profile Mismatches

The CAM profile on all cards must match the system profile. In most cases, the system corrects mismatches by copying the correct profile to the card, and rebooting the card. If three resets do not bring up the card, or if the system is running an FTOS version prior to 6.3.1.1, the system presents an error message. In this case, manually adjust the CAM configuration on the card to match the system configuration.

Dell Force10 recommends the following to prevent mismatches:

- Use the eg-default CAM profile in a chassis that has only EG Series line cards. If this profile is used in a chassis with non-EG line cards, the non-EG line cards enter a problem state.
- Before moving a card to a new chassis, change the CAM profile on a card to match the new system profile.
- After installing a secondary RPM into a chassis, copy the running-configuration to the startup-configuration.
- Change to the default profile if downgrading to and FTOS version earlier than 6.3.1.1.
- Use the CONFIGURATION mode commands so that the profile is change throughout the system.
- Use the EXEC Privilege mode commands to match the profile of a component to the profile of the target system.

QoS CAM Region Limitation

The default CAM profile allocates a partition within the IPv4Flow region to store QoS service policies. If the QoS CAM space is exceeded, messages similar to the ones in Message 5 are displayed.

Message 5 QoS CAM Region Exceeded

```
%EX2YD:12 %DIFFSERV-2-DSA_QOS_CAM_INSTALL_FAILED: Not enough space in L3 Cam(PolicyQos) for
class 2 (Gi 12/20) entries on portpipe 1 for linecard 12
%EX2YD:12 %DIFFSERV-2-
DSA_QOS_CAM_INSTALL_FAILED: Not enough space in L3 Cam(PolicyQos) for class 5 (Gi 12/22)
entries on portpipe 1 for linecard 12
```

If you exceed the QoS CAM space:

Step	Task
1	Verify that you have configured a CAM profile that allocates 24K entries to the IPv4 system flow region. See View CAM Profiles on page 291.
2	Allocate more entries in the IPv4Flow region to QoS. See Configure IPv4Flow Sub-partitions on page 293.

FTOS version 7.4.1 introduced the ability to view the actual CAM usage before applying a service-policy. The command **test cam-usage service-policy** provides this test framework, see Pre-calculating Available QoS CAM Space on page 874.

Note: For troubleshooting other CAM issues see the *E-Series Network Operations Guide*.

12“

Configuration Replace and Rollback

Configuration Replace and Rollback is supported on platforms C E

The E-Series ExaScale platform is supported with FTOS 8.1.1.0 and later.

Configuration Replace and Rollback enables you to replace the current running-configuration with different configuration without restarting the chassis.

Without this feature, if you want to load a new running configuration, you must copy the desired configuration file to the startup-configuration (using the command **copy** *file* **startup-configuration**) and reboot the chassis (using the command **reload**). Copying the desired configuration file to the running-configuration file (using the command **copy** *file* **running-configuration**) merely appends the running configuration; any conflicts between the two files is reported to the console, but FTOS does not overwrite the running configuration, therefore the new configuration is not fully implemented.

The reboot process takes several minutes by default, and if your startup-configuration is extensive, the process can take several minutes more. As a result, when the Dell Force10 system is deployed in production environment, you must wait for a maintenance window to load a new configuration.

The Configuration Replace and Rollback feature allows you to archive your running configuration, and at a later time, replace your running configuration with the archived one without rebooting the chassis. During replacement FTOS calculates and applies only the difference between the archived file and the running-configuration, making the process faster. Once the archived configuration is loaded, you can confirm the replacement, or revert (roll back) to your previous configuration. Rolling back allows you to view and test a configuration before completing the change.

Archived Files

Archived files are stored on the internal flash in a hidden directory. The maximum number of archived files is configurable between 10 and 15. If you archive more than the configured maximum, the oldest archived file is deleted to create space. You can view the name, size, and date of creation of a file, but you cannot view the contents of the archived file directly (using the command **show file**). To view the contents of a file you can backup the archive file to another location and then use the command show file, or view the the differences between the archived file and another file using the **show diff** command.

Configuring Configuration Replace and Rollback

Configuring Configuration Replace and Rollback is a three-step process:

1. Enable the archive service. See page 302.
2. Archive a running-configuration. See page 303.
3. Replace the running-configuration with an archived configuration. See page 303.

Related Configuration Tasks

- Configuring an Archive File Maximum on page 305
- Configuring Auto-archive on page 306
- Copying and Deleting an Archive File on page 307
- Viewing and Editing the Contents of an Archive File on page 307

Important Points to Remember

- FTOS automatically locks CONFIGURATON mode during the replace and rollback operation; see Lock CONFIGURATION mode on page 72. Therefore, when using this feature, no other user may be in CONFIGURATION mode. The lock is released when the replace or rollback operation is complete.
- Configuration Replace and Rollback cannot remove some FTOS configuration statements. See the release notes for your FTOS version for details.

Enabling the Archive Service

Before you can archive a configuration, you must enter ARCHIVE mode using the command **archive** from CONFIGURATION mode, as shown in Figure 12-1. This enables the archiving service. If you do not enable the archive service, Message 1 appears when you attempt to archive a configuration.

Message 1 Archive Service Error Message

```
% Warning: archive service is not enabled yet.
```

Figure 12-1. Entering Archive Mode

```
FTOS#archive config
% Warning: archive service is not enabled yet.
FTOS#config
FTOS(conf)#archive
FTOS(conf-archive)#exit
FTOS(conf)#exit
FTOS#archive config
configuration archived as archive_1
FTOS#
```

You do not have to enable the archive service again if you save the running configuration after completing task. If you reload the system or upgrade your FTOS version without saving the running configuration you must enable the archive service again.

Archiving a Configuration File

Archive the current running configuration file using the command **archive config** from EXEC Privilege mode.

Figure 12-2. Archiving a Configuration File

```
R1#archive ?
config                  Archive the running configuration
backup                  Backup the archive file
R1#archive config
configuration archived as archive_0
R1#show archive
Archive directory: flash:/CFGARCH_DIR

#    Archive        Date        Time      Size        Comment
0    archive_0      11/19/2007  14:29:26  6040        Most recently archived
1    -
2    -
3    -
4    -
5    -
6    -
7    -
8    -
9    -
10   -
11   -
12   -
13   -
14   -
R1#
```

Viewing the Archive Directory

The archive directory is a hidden directory that FTOS does not display in the output of the command **dir**. View the archive directory using the command **show archive** from EXEC Privilege mode, as shown in Figure 12-2.

Replacing the Current Running Configuration

Replace the current running configuration with an archived configuration using the command **configure replace** from EXEC Privilege mode.

In Figure 12-3:

1. The hostname of the Dell Force10 system is changed from “R1” to “FTOS.”
2. The running configuration is replaced with archive_0, in which the hostname is “R1.”

Figure 12-3. Replacing the Running-configuration with and Archived Configuration

```
R1#config
R1(conf)#hostname FTOS
FTOS#configure replace archive_0

This will apply all nessesary additions and deletions
to replace the current running-config with the contents
of the specified configuration file,
which is assumed to be complete configuration,
not a partial configuration
Please confirm if you want to proceed [yes/no]:yes
2d3h3m: %RPM0-P:CP %CLI-6-RBACKSTART: start rollback to file flash:/CFGARCH_DIR/archive_0
2d3h3m: %RPM0-P:CP %SYS-5-CONFIG_LOAD: Loading configuration file
2d3h3m: %RPM0-P:CP %CLI-6-RBACKCOMPLETE: completed rollback to flash:/CFGARCH_DIR/archive_0
R1#
```

Confirmation Dialog

Use the keyword **force** to bypass the FTOS confirmation dialog, as shown in Figure 12-4.

Figure 12-4. Replacing the Running-configuration without a Confirmation Dialog

```
R1(conf)#hostname FTOS
FTOS#exit
FTOS#configure replace archive_0 force
2d3h8m: %RPM0-P:CP %CLI-6-RBACKSTART: start rollback to file flash:/CFGARCH_DIR/archive_0
2d3h8m: %RPM0-P:CP %SYS-5-CONFIG_LOAD: Loading configuration file
2d3h8m: %RPM0-P:CP %CLI-6-RBACKCOMPLETE: completed rollback to flash:/CFGARCH_DIR/archive_0
R1#
```

Rolling Back to the Previous Configuration

FTOS allows you to implement an archived configuration for a specified amount of time, before reverting to the previous running-configuration using the command **configure replace** from EXEC Privilege mode; FTOS requires you to enter the amount of time in seconds. This feature enables you to test a configuration before committing the system to it.

- If you do not like the configuration, wait for the specified time to expire, as shown in Figure 12-5.
- If you like the configuration, enter the command **configure confirm** from EXEC Privilege mode before the specified time, as shown in Figure 12-6.

Figure 12-5. Configuring FTOS to Rollback to a Previous Configuration

```
FTOS#configure replace archive_0 time ?
<60-1800>               Time value (in seconds)
FTOS#configure replace archive_0 time 60

This will apply all nessesary additions and deletions
to replace the current running-config with the contents
of the specified configuration file,
which is assumed to be complete configuration,
not a partial configuration
Please confirm if you want to proceed [yes/no]:yes
3d4h45m: %RPM0-P:CP %CLI-6-RBACKSTART: start rollback to file flash:/CFGARCH_DIR/archive_0
3d4h45m: %RPM0-P:CP %SYS-5-CONFIG_LOAD: Loading configuration file
3d4h45m: %RPM0-P:CP %CLI-6-RBACKCOMPLETE: completed rollback to flash:/CFGARCH_DIR/archive_0
R1#
R1#Warning: time is expired before confirm. Replace with flash://CFGARCH_DIR/archive_1
FTOS#
```

Figure 12-6. Committing to an Archived Configuration

```
FTOS#config replace archive_0 time 60

This will apply all nessesary additions and deletions
to replace the current running-config with the contents
of the specified configuration file,
which is assumed to be complete configuration,
not a partial configuration
Please confirm if you want to proceed [yes/no]:yes
3d5h26m: %RPM0-P:CP %CLI-6-RBACKSTART: start rollback to file flash:/CFGARCH_DIR/archive_0
3d5h26m: %RPM0-P:CP %SYS-5-CONFIG_LOAD: Loading configuration file
3d5h26m: %RPM0-P:CP %CLI-6-RBACKCOMPLETE: completed rollback to flash:/CFGARCH_DIR/archive_0
R1#configure confirm
```

Configuring an Archive File Maximum

The maximum number of archive files is configurable between 2 and 15. Default maximum is 10. Use the command **maximum** from ARCHIVE mode to configure this parameter, as shown in Figure 12-7.

- If you attempt to archive more configurations than the maximum allowed, the oldest archived configuration is deleted (Figure 12-8) to create space. However, the number in the name of the archived file is still incremented (up to 14, after which the numbering convention restarts at 0; if present, archive_0 is overwritten).
- If you configure a maximum less than the number of archived files you already have, then archived files are deleted to satisfy the maximum.

Figure 12-7. Configuring an Archive Maximum

```
R1(conf-archive)#maximum 2
R1(conf-archive)#show config
!
archive
 maximum 2
R1(conf-archive)#
```

Figure 12-8. Configuring the Maximum Number of Archive Files (continued)

```
R1#archive config
configuration archived as archive_1
R1#show archive
Archive directory: flash:/CFGARCH_DIR

#    Archive        Date        Time      Size        Comment
0    archive_0      11/20/2007  09:45:24  6120        Archived
1    archive_1      11/20/2007  10:54:12  6120        Most recently archived
2    -
3    -
4    -
5    -
6    -
7    -
8    -
9    -
10   -
11   -
12   -
13   -
14   -
R1#archive config
configuration archived as archive_2
R1#show archive
Archive directory: flash:/CFGARCH_DIR

#    Archive        Date        Time      Size        Comment
0    -                                              Deleted
1    archive_1      11/20/2007  10:54:12  6120        Archived
2    archive_2      11/20/2007  10:54:28  6120        Most recently archived
3    -
4    -
5    -
6    -
7    -
8    -
9    -
10   -
11   -
12   -
13   -
14   -
R1#
```

Configuring Auto-archive

You can configure the system to archive the running-configuration periodically so that you do not have to archive manually. Configure auto-archiving using the command **time-period** from ARCHIVE mode. Note that if you do not make any changes to the running-configuration for the configured length of time, then the running-configuration is not archived, and periodic archiving pauses; it resumes when you make a change to the running-configuration.

Figure 12-9. Configuring an Archive Time-period

```
R1(conf-archive)#time-period 5
R1(conf-archive)#show config
!
archive
 maximum 2
 time-period 5
R1(conf-archive)#
```

Copying and Deleting an Archive File

Copy an archive file to another location using the command **archive backup**, as shown in Figure 12-10. Delete an archive file using the command **archive delete** from CONFIG ARCHIVE mode.

Viewing and Editing the Contents of an Archive File

You cannot view or edit the contents of archived files. FTOS disallows these functions to ensure that archived configurations are error-free when they are used in a replace or rollback function. You can, however, copy the file to another location using the command **archive backup**, and then view and edit the copy. If you copy the file to another location on FTOS, then you can view the contents of the file using the command **show file**, as shown in Figure 12-10.

Figure 12-10. Viewing an Archive File

```
R1#archive backup archive_2 flash://archive_2
!
6120 bytes successfully copied
R1#dir
Directory of flash:

  1  drw-      32768   Jan 01 1980 00:00:00  .
  2  drwx        512   Nov 16 2007 13:20:22  ..
  3  drw-       8192   Mar 11 2007 00:23:40  TRACE_LOG_DIR
  4  drw-       8192   Mar 11 2007 00:23:40  CRASH_LOG_DIR
  5  drw-       8192   Mar 11 2007 00:23:40  NVTRACE_LOG_DIR
  6  drw-       8192   Mar 11 2007 00:23:40  CORE_DUMP_DIR
  7  d---       8192   Mar 11 2007 00:23:40  ADMIN_DIR
  8  -rw-       6115   Nov 19 2007 18:35:32  startup-config
  9  -rw-   32999090   Jun 11 2007 20:22:32  FTOS-EF-7.4.1.0.bin
 10  -rw-   33059550   May 31 2007 20:58:56  FTOS-EF-7.4.2.0.bin
 11  -rw-   23234380   May 30 2007 06:38:14  FTOS-EF-6.5.4.0.bin
 12  -rw-       6115   Nov 19 2007 18:15:00  startup-config.bak
 13  -rw-         34   Nov 19 2007 19:23:00  arc_delta.cfg
 14  -rw-       6120   Nov 20 2007 11:17:52  archive_2

flash: 520962048 bytes total (320643072 bytes free)
R1#show file flash://archive_2
! Version E_MAIN4.7.5.353
! Last configuration change at Tue Nov 20 10:54:05 2007 by default
! Startup-config last updated at Mon Nov 19 18:35:30 2007 by default
!
boot system rpm0 primary flash://FTOS-EF-4.7.5.353.bin
boot system rpm0 secondary flash://FTOS-EF-7.4.2.0.bin
boot system rpm0 default flash://FTOS-EF-6.5.4.0.bin
boot system rpm1 primary flash://FTOS-EF-7.5.1.0.bin
boot system rpm1 secondary flash://FTOS-EF-7.4.2.0.bin
boot system rpm1 default flash://FTOS-EF-6.5.4.0.bin
!
redundancy auto-failover-limit count 3 period 60
redundancy auto-synchronize full
redundancy disable-auto-reboot rpm
redundancy primary rpm0
!
--More--
```

Viewing the Difference between Configuration Files

View the difference between the running-configuration and an archived configuration using the command **show run diff**. In Figure 12-11, the running-configuration is archived as archive_3, and then the hostname is changed to “FTOS.” The command **show run diff** lists each difference in the two files; in this case, there is only one, the hostname.

Figure 12-11. Viewing the Difference between Configuration Files

```
R1#archive config
configuration archived as archive_3
R1(conf)#hostname FTOS
FTOS(conf)#do show run diff archive_3
running-config
-------
< hostname FTOS

flash:/CFGARCH_DIR/archive_3
-------
> hostname R1

FTOS(conf)#
```

13

Dynamic Host Configuration Protocol

Dynamic Host Configuration Protocol is available on platforms: C E S

This chapter contains the following sections:

Protocol Overview

Dynamic Host Configuration Protocol (DHCP) is an application layer protocol that dynamically assigns IP addresses and other configuration parameters to network end-stations (hosts) based on configuration policies determined by network administrators. DHCP:

- relieves network administrators of manually configuring hosts, which is a can be a tedious and error-prone process when hosts often join, leave, and change locations on the network.
- reclaims IP addresses that are no longer in use to prevent address exhaustion.

DHCP is based a client-server model. A host discovers the DHCP server and requests an IP address, and the server either leases or permanently assigns one. There are three types of devices that are involved in DHCP negotiation:

- **DHCP Server**—a network device offering configuration parameters to the client.
- **DHCP Client**—a network device requesting configuration parameters from the server.
- **Relay agent**—an intermediary network device that passes DHCP messages between the client and server when the server is not on the same subnet as the host.

DHCP Packet Format and Options

DHCP uses UDP as its transport protocol. The server listens on port 67 and transmits to port 68; the client listens on port 68 and transmits to port 67. The configuration parameters are carried as options in the DHCP packet in Type, Length, Value (TLV) format; many options are specified in RFC 2132. To limit the number parameters that servers must provide, hosts specify the parameters that they require, and the server sends only those; some common options are given in Table 13-1.

Figure 13-1. DHCP Packet Format

Table 13-1. Common DHCP Options

Option	Code	Description
Subnet Mask	1	Specifies the clients subnet mask.
Router	3	Specifies the router IP addresses that may serve as the client's default gateway.
Domain Name Server	6	Specifies the the DNS servers that are available to the client.
Domain Name	15	Specifies the domain name that client should use when resolving hostnames via DNS.
IP Address Lease Time	51	Specifies the amount of time that the client is allowed to use an assigned IP address.
DHCP Message Type	53	1: DHCPDISCOVER 2: DHCPOFFER 3: DHCPREQUEST 4: DHCPDECLINE 5:DHCPACK 6:DHCPNACK 7:DHCPRELEASE 8:DHCPINFORM
Parameter Request List	55	Clients use this option to tell the server which parameters it requires. It is a series of octects where each octet is DHCP option code.
Renewal Time	58	Specifies the amount of time after the IP address is granted that the client attempts to renew its lease with the *original* server.
Rebinding Time	59	Specifies the amount of time after the IP address is granted that the client attempts to renew its lease with *any* server, if the original server does not respond.
End	255	Signals the last option in the DHCP packet.

Assigning an IP Address using DHCP

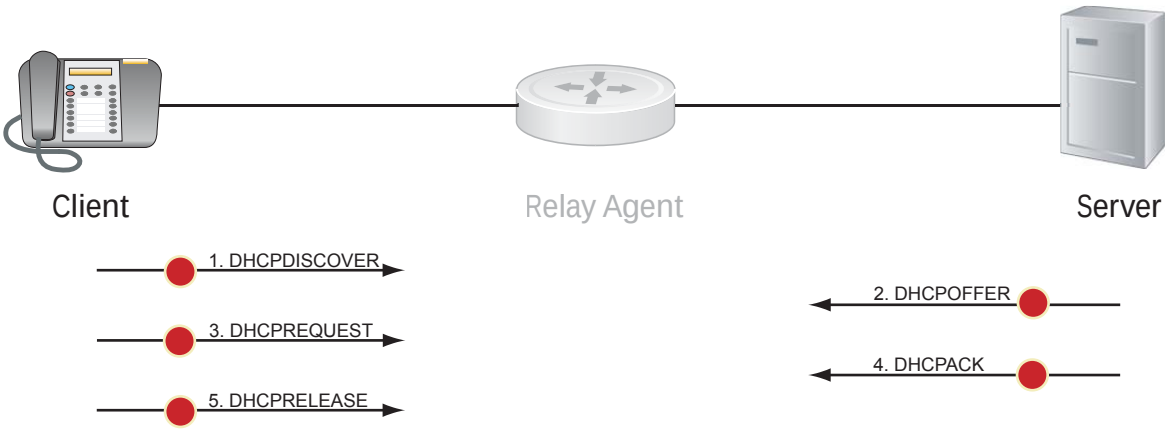
When a client joins a network:

1. The client initially broadcasts a **DHCPDISCOVER** message on the subnet to discover available DHCP servers. This message includes the parameters that the client requires and might include suggested values for those parameters.
2. Servers unicast or broadcast a **DHCPOFFER** message in response to the DHCPDISCOVER that offers to the client values for the requested parameters. Multiple servers might respond to a single DHCPDISCOVER; the client might wait a period of time and then act on the most preferred offer.
3. The client broadcasts a **DHCPREQUEST** message in response to the offer, requesting the offered values.
4. Upon receiving a DHCPREQUEST, the server binds the clients' unique identifier (the hardware address plus IP address) to the accepted configuration parameters and stores the data in a database called a *binding table*. The server then broadcasts a **DHCPACK** message, which signals to the client that it may begin using the assigned parameters.
5. When the client leaves the network, or the lease time expires, returns its IP address to the server in a **DHCPRELEASE** message.

There are additional messages that are used in case the DHCP negotiation deviates from the process previously described and shown in Figure 13-2.

- **DHCPDECLINE—**A client sends this message to the server in response to a DHCPACK if the configuration parameters are unacceptable, for example, if the offered address is already in use. In this case, the client starts the configuration process over by sending a DHCPDISCOVER.
- **DHCPINFORM—**A client uses this message to request configuration parameters when it assigned an IP address manually rather than with DHCP. The server responds by unicast.
- **DHCPNAK—**A server sends this message to the client if it is not able to fulfill a DHCPREQUEST, for example if the requested address is already in use. In this case, the client starts the configuration process over by sending a DHCPDISCOVER.

Figure 13-2. Assigning Network Parameters using DHCP



Implementation Information

- The Dell Force10 implementation of DHCP is based on RFC 2131 and RFC 3046.
- DHCP is available on VLANs and Private VLANs.
- IP Source Address Validation is a sub-feature of DHCP Snooping; FTOS uses ACLs internally to implement this feature and as such, you cannot apply ACLs to an interface which has IP Source Address Validation. If you configure IP Source Address Validation on a member port of a VLAN and then attempt to apply a access list to the VLAN, FTOS displays the first line in Message 1. If you first apply an ACL to a VLAN and then attempt enable IP Source Address Validation an one of its member ports, FTOS displays the second line in Message 1.

Message 1 DHCP Snooping with VLAN ACL Compatibility Error

```
% Error: Vlan member has access-list configured.
% Error: Vlan has an access-list configured.
```

- FTOS provides 40K entries that can be divided between leased addresses and excluded addresses. By extension, the maximum number of pools you can configure depends on the on the subnet mask that you give to each pool. FTOS displays an error message for configurations that exceed the allocated memory.
- E-Series supports 16K DHCP Snooping entries across 500 VLANs.
- C-Series and S-Series support 4K DHCP Snooping entries.
- All platforms support DAI on 16 VLANs per system.

Configuration Tasks

- Configure the System to be a DHCP Server on page 314
- Configure the System to be a Relay Agent on page 320
- Configure Secure DHCP on page 321

Configure the System to be a DHCP Server

Configure the System to be a DHCP Server is supported only on platforms: C S

A DHCP server is a network device that has been programmed to provide network configuration parameters to clients upon request. Servers typically serve many clients, making host management much more organized and efficient.

The key responsibilities of DHCP servers are:

1. **Address Storage and Management**: DHCP servers are the owners of the addresses used by DHCP clients.The server stores the addresses and manages their use, keeping track of which addresses have been allocated and which are still available.

2. **Configuration Parameter Storage and Management**: DHCP servers also store and maintain other parameters that are sent to clients when requested. These parameters specify in detail how a client is to operate.
3. **Lease Management**: DHCP servers use leases to allocate addresses to clients for a limited time. The DHCP server maintains information about each of the leases, including lease length.
4. **Responding To Client Requests**: DHCP servers respond to different types of requests from clients, primarily, granting, renewing, and terminating leases.
5. **Providing Administration Services**: The DHCP server includes functionality that allows an administrator to implement policies that govern how DHCP performs its other tasks.

Configuration Tasks

To configure DHCP, an administrator must first set up a DHCP server and provide it with configuration parameters and policy information including IP address ranges, lease length specifications, and configuration data that DHCP hosts need.

Configuring the Dell Force10 system to be a DHCP server is a 3-step process:

1. Configure the Server for Automatic Address Allocation
2. Specify a Default Gateway
3. Enable DHCP Server

Related Configuration Tasks

- Configure a Method of Hostname Resolution on page 317
- Allocate Addresses to BOOTP Clients on page 318
- Create Manual Binding Entries on page 318
- Check for Address Conflicts on page 319
- DHCP Clear Commands on page 320

Configure the Server for Automatic Address Allocation

Automatic Address Allocation is an address assignment method, by which the DHCP server leases an IP address to a client from a pool of available addresses.

Create an IP Address Pool

An address pool is a range of IP addresses that may be assigned by the DHCP server. Address pools are indexed by subnet number.

To create an address pool:

Step	Task	Command Syntax	Command Mode
1	Access the DHCP server CLI context.	**ip dhcp server**	CONFIGURATION
2	Create an address pool and give it a name.	**pool** *name*	DHCP
3	Specify the range of IP addresses from which the DHCP server may assign addresses. • *network* is the subnet address. • *prefix-length* specifies the number of bits used for the network portion of the address you specify.	**network** *network /prefix-length* Prefix-length Range: 17-31	DHCP <POOL>
4	Display the current pool configuration.	**show config**	DHCP <POOL>

Once an IP address is leased to a client, only that client may release the address. FTOS performs a IP + MAC source address validation to ensure that no client can release another clients address. This is a default behavior, and is separate from IP+MAC Source Address Validation on page 328.

Exclude Addresses from the Address Pool

The DHCP server assumes that all IP addresses in a DHCP address pool are available for assigning to DHCP clients. You must specify the IP address that the DHCP server should not assign to clients.

Task	Command Syntax	Command Mode
Exclude an address range from DHCP assignment. The exclusion applies to all configured pools.	**excluded-address**	DHCP

Specify an Address Lease Time

Task	Command Syntax	Command Mode
Specify an address lease time for the addresses in a pool.	**lease** {**days** [**hours**] [**minutes**] | **infinite**} Default: 24 hours	DHCP <POOL>

Specify a Default Gateway

The IP address of the default router should be on the same subnet as the client.

Task	Command Syntax	Command Mode
Specify default gateway(s) for the clients on the subnet, in order of preference.	**default-router** *address*	DHCP <POOL>

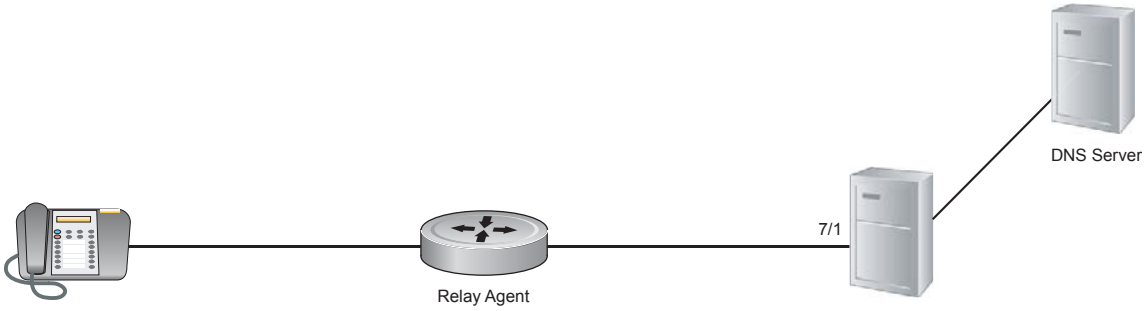
Enable DHCP Server

DHCP server is disabled by default.

Step	Task	Command Syntax	Command Mode
1	Enter the DHCP command-line context.	**ip dhcp server**	CONFIGURATION
2	Enable DHCP server.	**no disable** Default: Disabled	DHCP
3	Display the current DHCP configuration.	**show config**	DHCP

In Figure 13-3, an IP phone is powered by PoE and has acquired an IP address from the Dell Force10 system, which is advertising LLDP-MED. The leased IP address is displayed using **show ip dhcp binding,** and confirmed with **show lldp neighbors**.

Figure 13-3. Configuring DHCP Server



Configure a Method of Hostname Resolution

Dell Force10 systems are capable of providing DHCP clients with parameters for two methods of hostname resolution.

Address Resolution using DNS

A domain is a group of networks. DHCP clients query DNS IP servers when they need to correlate host names to IP addresses.

Step	Task	Command Syntax	Command Mode
1	Create a domain.	**domain-name** *name*	DHCP <POOL>
2	Specify in order of preference the DNS servers that are available to a DHCP client.	**dns-server** *address*	DHCP <POOL>

Address Resolution using NetBIOS WINS

Windows Internet Naming Service (WINS) is a name resolution service that Microsoft DHCP clients use to correlate host names to IP addresses within a group of networks. Microsoft DHCP clients can be one of four types of NetBIOS nodes: broadcast, peer-to-peer, mixed, or hybrid.

Step	Task	Command Syntax	Command Mode
1	Specify the NetBIOS Windows Internet Naming Service (WINS) name servers, in order of preference, that are available to Microsoft Dynamic Host Configuration Protocol (DHCP) clients.	**netbios-name-server** *address*	DHCP <POOL>
2	Specify the NetBIOS node type for a Microsoft DHCP client. Dell Force10 recommends specifying clients as hybrid.	**netbios-node-type** *type*	DHCP <POOL>

Allocate Addresses to BOOTP Clients

Network segments may have both BOOTP and DHCP clients. In this kind of environment, there might be a BOOTP server and a DHCP server to serve the two types of clients separately. However, DHCP servers respond to a BOOTP requests, which in this case would be undesirable because BOOTP clients might receive an address from the DHCP pool. To prevent this, you can configure the DHCP server to ignore BOOTP request packets so that only the BOOTP server serves BOOTP clients.

Task	Command Syntax	Command Mode
Enables address allocation to BOOTP clients. The addresses are from the DHCP address pool for the subnet.	**ip dhcp bootp automatic**	DHCP
Selectively ignore BOOTP request packets.	**ip dhcp bootp ignore**	DHCP

Create Manual Binding Entries

An address binding is a mapping between the IP address and Media Access Control (MAC) address of a client. The DHCP server assign the client an available IP address automatically, and then creates a entry in the binding table. However, the administrator can manually create an entry for a client; manual bindings are useful when you want to guarantee that particular network device receives a particular IP address. Manual bindings can be considered single-host address pools. There is no limit on the number of manual bindings, but you can only configure one manual binding per host.

Note: FTOS does not prevent you from using a network IP as a host IP; be sure to not use a network IP as a host IP.

To create a manual binding:

Step	Task	Command Syntax	Command Mode
1	Create an address pool	**pool** *name*	DHCP
2	Specify the client IP address.	**host** *address*	DHCP <POOL>
3	Specify the client hardware address or client-identifier. • *hardware-address* is the client MAC address. *type* is the protocol of the hardware platform. The default protocol is Ethernet. *client-identifier* is required for Microsoft clients instead of a hardware addresses. The client identifier is formed by concatenating the media type and the MAC address of the client. Refer to the "Address Resolution Protocol Parameters" section of RFC 1700—Assigned Numbers, for a list of media type codes.	**hardware-address** *hardware-address type* **client-identifier** *unique-identifier*	DHCP <POOL>

Check for Address Conflicts

By default, the DHCP server pings an address from the pool twice before assigning the address to a client to attempt to verify that it is not in use. If the ping is unanswered, the DHCP server assumes that the address is not in use and assigns the address. By default, the DHCP server waits 500 milliseconds before timing out a ping packet.

Task	Command Syntax	Command Mode
Specify the number of ping packets the DHCP server sends to the pool address before assigning the address.	**ip dhcp ping packets** *number* Default: 2	CONFIGURATION
Change the amount of time the server waits for a ping reply before considering the ping a failure.	**ip dhcp ping timeout** *milliseconds* Default: 500 milliseconds	CONFIGURATION

An address conflict occurs when two hosts use the same IP address. The server checks for a conflict using ping and the client checks for conflict using gratuitous ARP. If a conflict is detected, the address is removed from the pool. The address will not be assigned until the administrator resolves the conflict.

Task	Command Syntax	Command Mode
Log IP address conflicts.	**ip dhcp conflict logging**	CONFIGURATION

DHCP Clear Commands

Task	Command Syntax	Command Mode
Clear DHCP binding entries for the entire binding table.	**clear ip dhcp binding**	EXEC Privilege
Clear a DHCP binding entry for an individual IP address.	**clear ip dhcp binding** *ip address*	EXEC Privilege
Clear a DHCP address conflict.	**clear ip dhcp conflict**	EXEC Privilege
Clear DHCP server counters.	**clear ip dhcp server statistics**	EXEC Privilege

Configure the System to be a Relay Agent

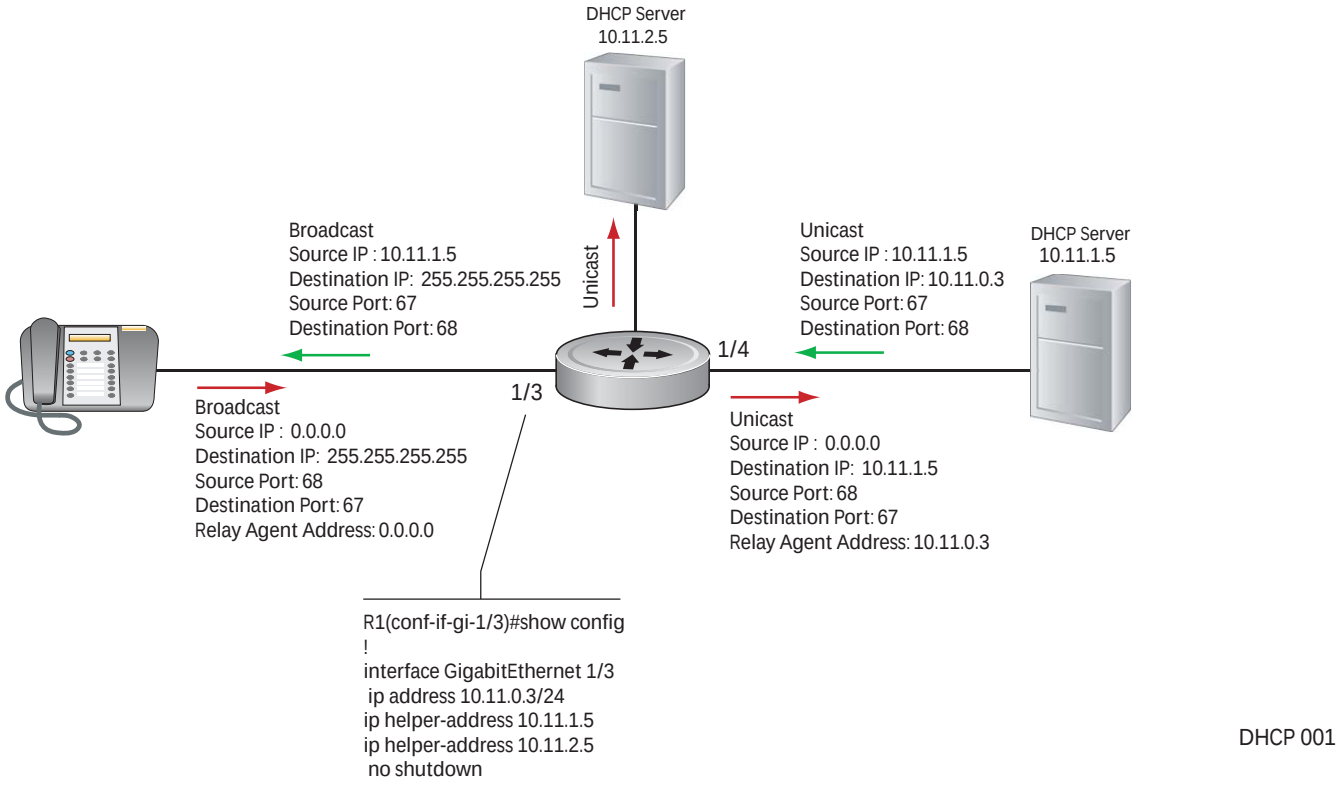
DHCP clients and servers request and offer configuration information via broadcast DHCP messages. Routers do not forward broadcasts, so if there are no DHCP servers on the subnet, the client does not receive a response to its request and therefore cannot access the network.

You can configure an interface on the Dell Force10 system to relay the DHCP messages to a specific DHCP server using the command **ip helper-address** *dhcp-address* from INTERFACE mode, as shown in Figure 13-4. Specify multiple DHCP servers by entering the **ip helper-address** *dhcp-address* command multiple times.

When **ip helper-address** is configured, the system listens for DHCP broadcast messages on port 67. The system rewrites packets received from the client and forwards it via unicast; the system rewrites the destination IP address and writes its own address as the relay device. Responses from the server are unicast back to the relay agent on port 68, and the relay agent rewrites the destination address and forwards the packet to the client subnet via broadcast.

Note: DHCP Relay is not available on Layer 2 interfaces.
Note: In a Private VLAN, **ip helper-address** is configured from Interface VLAN mode of the Primary VLAN. When **ip helper-address** is configured, the system listens for DHCP broadcast messages on port 67. The system rewrites packets received from the clients on primary and secondary (community and isolated) VLANs and forwards it via unicast; the system rewrites the destination IP address and writes primary VLANs IP as the relay device. Responses from the server are unicast back to the relay agent on port 68, and the relay agent rewrites the destination address and forwards the packet to the client subnet via broadcast.

Figure 13-4. Configuring Dell Force10 Systems as a DHCP Relay Device



To view the **ip helper-address** configuration for an interface, use the command **show ip interface** from EXEC privilege mode, Figure 250.

Figure 13-5. Displaying the Helper Address Configuration

```
R1_E600#show ip int gig 1/3
GigabitEthernet 1/3 is up, line protocol is down
Internet address is 10.11.0.1/24
Broadcast address is 10.11.0.255
Address determined by user input
IP MTU is 1500 bytes
Helper address is 192.168.0.1
                  192.168.0.2
Directed broadcast forwarding is disabled
Proxy ARP is enabled
Split Horizon is enabled
Poison Reverse is disabled
ICMP redirects are not sent
ICMP unreachables are not sent
```

Configure Secure DHCP

DHCP as defined by RFC 2131 provides no authentication or security mechanisms. Secure DHCP is a suite of features that protects networks that use dynamic address allocation from spoofing and attacks.

- Option 82 on page 322

- DHCP Snooping on page 322
- Dynamic ARP Inspection on page 325
- Source Address Validation on page 327

Option 82

RFC 3046 (Relay Agent Information option, or Option 82) is used for class-based IP address assignment.

The code for the Relay Agent Information option is 82, and is comprised of two sub-options, Circuit ID and Remote ID.

- **Circuit ID** is the interface on which the client-originated message is received.
- **Remote ID** identifies the host from which the message is received. The value of this sub-option is the MAC address of the relay agent that adds Option 82.

The DHCP relay agent inserts Option 82 before forwarding DHCP packets to the server. The server can use this information to:

- track the number of address requests per relay agent; restricting the number of addresses available per relay agent can harden a server against address exhaustion attacks.
- associate client MAC addresses with a relay agent to prevent offering an IP address to a client spoofing the same MAC address on a different relay agent.
- assign IP addresses according to the relay agent. This prevents generating DHCP offers in response to requests from an unauthorized relay agent.

The server echoes the option back to the relay agent in its response, and the relay agent can use the information in the option to forward a reply out the interface on which the request was received rather than flooding it on the entire VLAN.

The relay agent strips Option 82 from DHCP responses before forwarding them to the client.

Task	Command Syntax	Command Mode
Insert Option 82 into DHCP packets. For routers between the relay agent and the DHCP server, enter the **trust-downstream** option.	**ip dhcp relay information-option** [**trust-downstream**]	CONFIGURATION

DHCP Snooping

DHCP Snooping protects networks from spoofing. In the context of DHCP Snooping, all ports are either trusted or untrusted. By default, all ports are untrusted. Trusted ports are ports through which attackers cannot connect. Manually configure ports connected to legitimate servers and relay agents as trusted.

When DHCP Snooping is enabled, the relay agent builds a binding table—using DHCPACK messages—containing the client MAC address, IP addresses, IP address lease time, port, VLAN ID, and binding type. Every time the relay agent receives a DHCPACK on an trusted port, it adds an entry to the table.

The relay agent then checks all subsequent DHCP client-originated IP traffic (DHCPRELEASE, DHCPNACK, and DHCPDECLINE) against the binding table to ensure that the MAC-IP address pair is legitimate, and that the packet arrived on the correct port; packets that do not pass this check are forwarded to the the server for validation. This check-point prevents an attacker from spoofing a client and declining or releasing the real client's address. Server-originated packets (DHCPOFFER, DHCPACK, DHCPNACK) that arrive on an untrusted port are also dropped. This check-point prevents an attacker from impostering as a DHCP server to facilitate a man-in-the-middle attack.

Binding table entries are deleted when a lease expires, or the relay agent encounters a DHCPRELEASE, DHCPNACK, DHCPDECLINE.

FTOS Behavior: Introduced in FTOS version 7.8.1.0, DHCP Snooping was available for Layer 3 only and dependent on DHCP Relay Agent (**ip helper-address**). FTOS version 8.2.1.0 extends DHCP Snooping to Layer 2, and you do not have to enable relay agent to snoop on Layer 2 interfaces.

FTOS Behavior: Binding table entries are deleted when a lease expires or when the relay agent encounters a DHCPRELEASE. Starting with FTOS Release 8.2.1.2, line cards maintain a list of snooped VLANs. When the binding table is exhausted, DHCP packets are dropped on snooped VLANs, while these packets are forwarded across non-snooped VLANs. Since DHCP packets are dropped, no new IP address assignments are made. However, DHCPRELEASE and DHCPDECLINE packets are allowed so that the DHCP snooping table can decrease in size. Once the table usage falls below the maximum limit of 4000 entries, new IP address assignments are allowed.

FTOS Behavior: In 8.2.1 releases, **ip dhcp snooping trust** was required on the port-channel interface as well as on channel members. In subsequent releases, it is no longer necessary nor permitted to configure port-channel members as trusted; configuring the port-channel interface alone as trusted is sufficient, and ports must have the default configuration to be a channel members. When upgrading from 8.2.1 releases, the channel-member configurations are applied first, so when the port-channel is configured, its membership configuration is rejected, since the member ports no longer have the default configuration. In this case, you must manually remove **ip dhcp snooping trust** on the channel members add the ports to the port-channel.

Note: DHCP server packets will be dropped on all untrusted interfaces of a system configured for DHCP snooping. To prevent these packets from being dropped, configure **ip dhcp snooping trust** on the server-connected port.

Enable DCHP snooping

Step	Task	Command Syntax	Command Mode
1	Enable DHCP Snooping globally.	**ip dhcp snooping**	CONFIGURATION
2	Specify ports connected to DHCP servers as trusted.	**ip dhcp snooping trust**	INTERFACE
3	Enable DHCP Snooping on a VLAN.	**ip dhcp snooping vlan**	CONFIGURATION

Add a static entry in the binding table

Task	Command Syntax	Command Mode
Add a static entry in the binding table.	**ip dhcp snooping binding mac**	EXEC Privilege

Clear the binding table

Task	Command Syntax	Command Mode
Delete all of the entries in the binding table	**clear ip dhcp snooping binding**	EXEC Privilege

Display the contents of the binding table

Task	Command Syntax	Command Mode
Display the contents of the binding table.	**show ip dhcp snooping**	EXEC Privilege

View the DHACP Snooping statistics with the **show ip dhcp snooping** command.

Figure 13-6. Command example: show ip dhcp snooping

```
FTOS#show ip dhcp snooping

IP DHCP Snooping                          : Enabled.
IP DHCP Snooping Mac Verification         : Disabled.
IP DHCP Relay Information-option          : Disabled.
IP DHCP Relay Trust Downstream            : Disabled.

Database write-delay (In minutes)         : 0


DHCP packets information
Relay Information-option packets          : 0
Relay Trust downstream packets            : 0
Snooping packets                          : 0

Packets received on snooping disabled L3 Ports   : 0
Snooping packets processed on L2 vlans    : 142

DHCP Binding File Details
Invalid File                              : 0
Invalid Binding Entry                     : 0
Binding Entry lease expired               : 0
List of Trust Ports                       :Te 0/49
List of DHCP Snooping Enabled Vlans       :Vl 10
List of DAI Trust ports                   :Te 0/49
```

Drop DHCP packets on snooped VLANs only

Binding table entries are deleted when a lease expires, or the relay agent encounters a DHCPRELEASE.

Starting with FTOS Release 8.2.1.1, line cards maintain a list of snooped VLANs. When the binding table fills, DHCP packets are dropped only on snooped-VLANs, while such packets will be forwarded across non-snooped VLANs. Since DHCP packets are dropped, no new IP address assignments are made. However, DHCP Release and Decline packets are allowed so that the DHCP snooping table can decrease in size. Once the table usage falls below the max limit of 4000 entries, new IP address assignments are allowed.

View the number of entries in the table with the **show ip dhcp snooping binding** command. This output displays the snooping binding table created using the ACK packets from the trusted port.

Figure 13-7. Command example: show ip dhcp snooping binding

```
FTOS#show ip dhcp snooping binding

Codes :  S - Static D - Dynamic

IP Address       MAC Address        Expires(Sec)  Type  VLAN   Interface
=====================================================================
10.1.1.251       00:00:4d:57:f2:50    172800        D    Vl 10    Gi 0/2
10.1.1.252       00:00:4d:57:e6:f6    172800        D    Vl 10    Gi 0/1
10.1.1.253       00:00:4d:57:f8:e8    172740        D    Vl 10    Gi 0/3
10.1.1.254       00:00:4d:69:e8:f2    172740        D    Vl 10    Te 0/50

Total number of Entries in the table : 4
```

Dynamic ARP Inspection

Dynamic ARP inspection prevents ARP spoofing by forwarding only ARP frames that have been validated against the DHCP binding table.

ARP is a stateless protocol that provides no authentication mechanism. Network devices accepts ARP request and replies from any device, and ARP replies are accepted even when no request was sent. If a client receives an ARP message for which a relevant entry already exists in its ARP cache, it overwrites the existing entry with the new information.

The lack of authentication in ARP makes it vulnerable to spoofing. ARP spoofing is a technique attackers use to inject false IP to MAC mappings into the ARP cache of a network device. It is used to launch man-in-the-middle (MITM), and denial-of-service (DoS) attacks, among others.

A spoofed ARP message is one in which MAC address in the sender hardware address field and the IP address in the sender protocol field are strategically chosen by the attacker. For example, in an MITM attack, the attacker sends a client an ARP message containing the attacker's MAC address and the gateway's IP address. The client then thinks that the attacker is the gateway, and sends all internet-bound

packets to it. Likewise, the attacker sends the gateway an ARP message containing the attacker's MAC address and the client's IP address. The gateway then thinks that the attacker is the client, and forwards all packets addressed to the client to it. As a result, the attacker is able to sniff all packets to and from the client.

Other attacks using ARP spoofing include:

- broadcast—an attacker can broadcast an ARP reply that specifies FF:FF:FF:FF:FF:FF as the gateway's MAC address, resulting in all clients broadcasting all internet-bound packets.
- MAC flooding—an attacker can send fraudulent ARP messages to the gateway until the ARP cache is exhausted, after which, traffic from the gateway is broadcast.
- denial of service—an attacker can send a fraudulent ARP messages to a client to associate a false MAC address with the gateway address, which would blackhole all internet-bound packets from the client.

Note: DAI uses entries in the L2SysFlow CAM region, a sub-region of SystemFlow. One CAM entry is required for every DAI-enabled VLAN, and you can enable DAI on up to 16 VLANs on a system. However, the ExaScale default CAM profile allocates only 9 entries to the L2SysFlow region for DAI. You can configure 10 to 16 DAI-enabled VLANs by allocating more CAM space to the L2SysFlow region before enabling DAI.

SystemFlow has 102 entries by default. This region is comprised of two sub-regions: L2Protocol and L2SystemFlow. L2Protocol has 87 entries, and L2SystemFlow has 15 entries. Six L2SystemFlow entries are used by Layer 2 protocols, leaving 9 for DAI. L2Protocol can have a maximum of 100 entries, and this region must be expanded to capacity before you can increase the size of L2SystemFlow. This is relevant when you are enabling DAI on VLANs. If, for example, you want to enable DAI on 16 VLANs, you need 7 more entries; in this case, reconfigure the SystemFlow region for 122 entries:

layer-2 eg-acl *value* **fib** *value* **frrp** *value* **ing-acl** *value* **learn** *value* **l2pt** *value* **qos** *value* **system-flow 122**

The logic is as follows:

L2Protocol has 87 entries by default and must be expanded to its maximum capacity, 100 entries, before L2SystemFlow can be increased; therefore 13 more L2Protocol entries are required. L2SystemFlow has 15 entries by default, but only 9 are for DAI; to enable DAI on 16 VLANs, 7 more entries are required. 87 L2Protocol + 13 additional L2Protocol + 15 L2SystemFlow + 7 additional L2SystemFlow equals 122.

Step	Task	Command Syntax	Command Mode
1	Enable DHCP Snooping.		
2	Validate ARP frames against the DHCP Snooping binding table.	**arp inspection**	INTERFACE VLAN

View the number of entries in the ARP database with the **show arp inspection database** command.

Figure 13-8. Command example: show arp inspection database

```
FTOS#show arp inspection database

Protocol    Address         Age(min)  Hardware Address    Interface   VLAN  CPU
--------------------------------------------------------------------------------
Internet    10.1.1.251          -   00:00:4d:57:f2:50    Gi 0/2     Vl 10  CP
Internet    10.1.1.252          -   00:00:4d:57:e6:f6    Gi 0/1     Vl 10  CP
Internet    10.1.1.253          -   00:00:4d:57:f8:e8    Gi 0/3     Vl 10  CP
Internet    10.1.1.254          -   00:00:4d:69:e8:f2    Te 0/50     Vl 10  CP
FTOS#
```

Use **show arp inspection statistics** command to see how many valid and invalid ARP packets have been processed.

Figure 13-9. Command example: show arp inspection database

```
FTOS#show arp inspection statistics

Dynamic ARP Inspection (DAI) Statistics
---------------------------------------
Valid ARP Requests                      : 0
Valid ARP Replies                       : 1000
Invalid ARP Requests                    : 1000
Invalid ARP Replies                     : 0
FTOS#
```

Bypass the ARP Inspection

You can configure a port to skip ARP inspection by defining the interface as trusted, which is useful in multi-switch environments. ARPs received on trusted ports bypass validation against the binding table. All ports are untrusted by default.

Task	Command Syntax	Command Mode
Specify an interface as trusted so that ARPs are not validated against the binding table.	**arp inspection-trust**	INTERFACE

FTOS Behavior: Introduced in FTOS version 8.2.1.0, Dynamic ARP Inspection (DAI) was available for Layer 3 only. FTOS version 8.2.1.1 extends DAI to Layer 2.

Source Address Validation

Using the DHCP binding table, FTOS can perform three types of source address validation (SAV):

- IP Source Address Validation on page 328 prevents IP spoofing by forwarding only IP packets that have been validated against the DHCP binding table.

- DHCP MAC Source Address Validation on page 328 verifies a DHCP packet's source hardware address matches the client hardware address field (CHADDR) in the payload.
- IP+MAC Source Address Validation on page 328 verifies that the IP source address and MAC source address are a legitimate pair.

IP Source Address Validation

IP Source Address Validation (SAV) prevents IP spoofing by forwarding only IP packets that have been validated against the DHCP binding table. A spoofed IP packet is one in which the IP source address is strategically chosen to disguise the attacker. For example, using ARP spoofing an attacker can assume a legitimate client's identity and receive traffic addressed to it. Then the attacker can spoof the client's IP address to interact with other clients.

The DHCP binding table associates addresses assigned by the DHCP servers, with the port on which the requesting client is attached. When IP Source Address Validation is enabled on a port, the system verifies that the source IP address is one that is associated with the incoming port. If an attacker is impostering as a legitimate client the source address appears on the wrong ingress port, and the system drops the packet. Likewise, if the IP address is fake, the address will not be on the list of permissible addresses for the port, and the packet is dropped.

Task	Command Syntax	Command Mode
Enable IP Source Address Validation	**ip dhcp source-address-validation**	INTERFACE

DHCP MAC Source Address Validation

DHCP MAC Source Address Validation (SAV) validates a DHCP packet's source hardware address against the client hardware address field (CHADDR) in the payload.

FTOS Release 8.2.1.1 ensures that the packet's source MAC address is checked against the CHADDR field in the DHCP header only for packets from snooped VLANs.

Task	Command Syntax	Command Mode
Enable DHCP MAC Source Address Validation.	**ip dhcp snooping verify mac-address**	CONFIGURATION

IP+MAC Source Address Validation

IP+MAC Source Address Validation is available on platforms: C S

IP Source Address Validation validates the IP source address of an incoming packet against the DHCP Snooping binding table. IP+MAC Source Address Validation ensures that the IP source address and MAC source address are a legitimate pair, rather validating each attribute individually. IP+MAC Source Address Validation cannot be configured with IP Source Address Validation.

Step	Task	Command Syntax	Command Mode
1	Allocate at least one FP block to the ipmacacl CAM region.	**cam-acl l2acl**	CONFIGURATION
2	Save the running-config to the startup-config.	**copy running-config startup-config**	EXEC Privilege
3	Reload the system.	**reload**	EXEC Privilege
4	Enable IP+MAC Source Address Validation.	**ip dhcp source-address-validation ipmac**	INTERFACE

FTOS creates an ACL entry for each IP+MAC address pair in the binding table and applies it to the interface.

Task	Command Syntax	Command Mode
Display the IP+MAC ACL for an interface for for the entire system.	**show ip dhcp snooping source-address-validation** [**interface**]	EXEC Privilege

14

Equal Cost Multi-Path

This chapter describes how to configure:

- ECMP for Flow-based Affinity (E-Series), including the configurable hash algorithm
- Configurable ECMP Hash Algorithm (C- and S-Series)

ECMP for Flow-based Affinity (E-Series)

ECMP for Flow-based Affinity (E-Series) is available on platform: E

The hashing algorithm on E-Series TeraScale and E-Series ExaScale are different:

- On ExaScale, the hashing algorithm is based on CRC, checksum, or XOR.
- On TeraScale, the hashing algorithm is based on checksum only.

If flow-based affinity is to be maintained by an ExaScale and TeraScale chassis, they must both use the same hashing algorithm and seed value, and ECMP must deterministically choose a next hop. To reconfigure these values, see:

- Configurable Hash Algorithm (E-Series) on page 331
- Configurable Hash Algorithm Seed on page 332
- Deterministic ECMP Next Hop on page 332

Configurable Hash Algorithm (E-Series)

TeraScale has one algorithm that is used for LAGs, ECMP, and NH-ECMP, and ExaScale can use three different algorithms for each of these features. To adjust the ExaScale behavior to match TeraScale, use the following command:

Task	Command Syntax	Command Mode
Change the ExaScale hash-algorithm for LAG, ECMP, and NH-ECMP to match TeraScale.	**hash-algorithm ecmp checksum 0 lag checksum 0 nh-ecmp checksum 0**	CONFIGURATION

FTOS Behavior: In FTOS versions prior to 8.2.1.2, the ExaScale default hash-algorithm is 0. Beginning with version 8.2.1.2, the default hash-algorithm is 24.

For information on the load-balancing criteria used by the hash algorithm to distribute traffic among ECMP paths and LAG members on an E-Series system, see E-Series load-balancing on page 436.

Deterministic ECMP Next Hop

Deterministic ECMP Next Hop arranges all ECMPs in order before writing them into the CAM. For example, suppose the RTM learns 8 ECMPs in the order that the protocols and interfaces came up. In this case, the FIB and CAM sort them so that the ECMPs are always arranged.This implementation ensures that every chassis having the same prefixes orders the ECMPs the same.

With 8 or less ECMPs, the ordering is lexicographic and deterministic. With more than 8 ECMPs, ordering is deterministic, but it is not in lexicographic order.

Task	Command Syntax	Command Mode
Enable IPv4 Deterministic ECMP Next Hop.	**ip ecmp-deterministic**	CONFIGURATION
Enable IPv6 Deterministic ECMP Next Hop.	**ipv6 ecmp-deterministic**	CONFIGURATION
Note: Packet loss might occur when you enable **ip/ipv6 ecmp-deterministic** for the first-time only.		

Configurable Hash Algorithm Seed

Deterministic ECMP sorts ECMPs in order even though RTM provides them in a random order. However, the hash algorithm uses as a seed the lower 12 bits of the chassis MAC, which yields a different hash result for every chassis. This means that for a given flow, even though the prefixes are sorted, two unrelated chassis will select different hops.

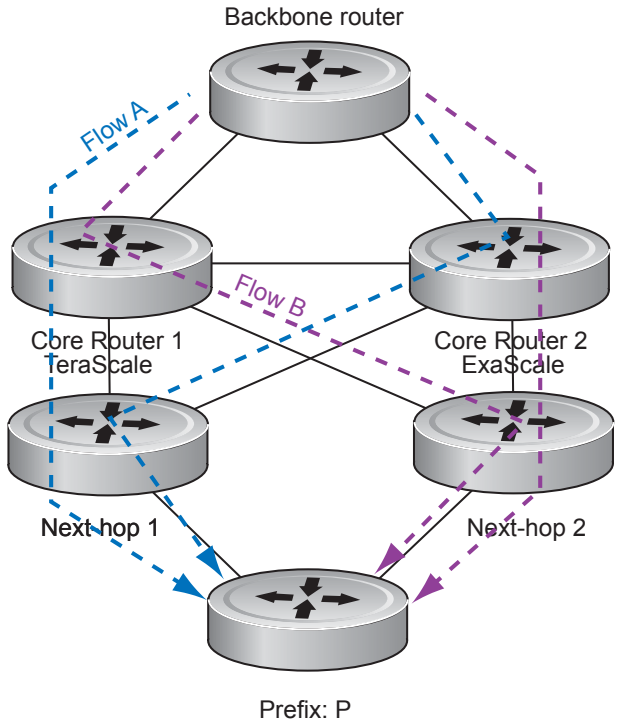
FTOS provides a CLI-based solution for modifying the hash seed to ensure that on each configured system, the ECMP selection is same. When configured, the same seed is set for ECMP, LAG, and NH, and is used for incoming traffic only.

Note: While the seed is stored separately on each port-pipe, the same seed is used across all CAMs.
Note: You cannot separate LAG and ECMP, but you can use different algorithms across chassis with the same seed. If LAG member ports span multiple port-pipes and line cards, set the seed to the same value on each port-pipe to achieve deterministic behavior.
Note: If the hash algorithm configuration is removed. Hash seed will not go to original factory default setting.

Task	Command Syntax	Command Mode
Specify the hash algorithm seed.	**hash-algorithm seed** *value* [**linecard** *number*] [**port-set** *number*] Range: 0 - 4095	CONFIGURATION

In Figure 14-1, Core Router 1 is an E-Series TeraScale and Core Router 2 is an E-Series ExaScale. They have similar configurations and have routes for prefix P with two possible next-hops. When Deterministic ECMP is enabled and the hash algorithm and seed are configured the same, each flow is consistently sent to the same next hop even though they are routed through two different chassis.

Figure 14-1. Deterministic ECMP Next Hop + Configurable Hash Algorithm Seed



Configurable ECMP Hash Algorithm (C- and S-Series)

Configurable ECMP Hash Algorithm (C- and S-Series) is available on platforms: C S

On C-Series and S-Series, the **hash-algorithm** command is specific to ECMP groups and has a different default from the E-Series (see Configurable Hash Algorithm (E-Series)). The default ECMP hash configuration is **crc-lower**, which takes the lower 32 bits of the hash key to compute the egress port. The hash value calculated with the hash algorithm is unique to the entire chassis.

Other options for the ECMP hash-algorithm are:

- **crc-upper** — Uses the upper 32 bits of the hash key to compute the egress port.
- **dest-ip** — Uses destination IP address as part of the hash key.
- **lsb** — Always uses the least significant bit of the hash key to compute the egress port.

To change to a different hash scheme for ECMP, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
hash-algorithm ecmp {**crc-lower** | **crc-upper** | **dest-ip** | **lsb**} Default: **crc-lower**	CONFIGURATION	Change to another algorithm.

The different hash algorithms for ECMP are based on the number of ECMP group members and packet values. The default hash algorithm yields the most balanced results in various test scenarios, but if the default algorithm does not provide satisfactory distribution of traffic, then use this command to designate another algorithm.

When a member leaves or is added to the ECMP group, the hash algorithm is recalculated to balance traffic across the members.

15

Force10 Resilient Ring Protocol

Force10 Resilient Ring Protocol is supported on platforms C E S

The E-Series ExaScale platform is supported with FTOS 8.1.1.0 and later.

Force10 Resilient Ring Protocol (FRRP) provides fast network convergence to Layer 2 switches interconnected in a ring topology, such as a Metropolitan Area Network (MAN) or large campuses. FRRP is similar to what can be achieved with the Spanning Tree Protocol (STP), though even with optimizations, STP can take up to 50 seconds to converge (depending on the size of network and node of failure) may require 4 to 5 seconds to reconverge. FRRP can converge within 150ms to 1500ms when a link in the ring breaks (depending on network configuration).

To operate a deterministic network, a network administrator must run a protocol that converges independently of the network size or node of failure. The Force10 Resilient Ring Protocol (FRRP) is a proprietary protocol that provides this flexibility, while preventing Layer 2 loops. FRRP provides sub-second ring-failure detection and convergence/re-convergence in a Layer 2 network while eliminating the need for running spanning-tree protocol. With its two-way path to destination configuration, FRRP provides protection against any single link/switch failure and thus provides for greater network uptime.

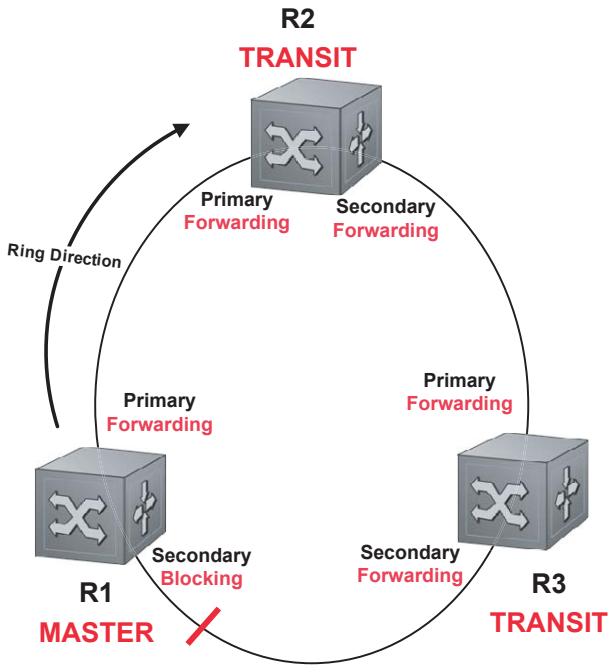
Protocol Overview

FRRP is built on a ring topology. Up to 255 rings can be configured on a system. FRRP uses one Master node and multiple Transit nodes in each ring. There is no limit to the number of nodes on a ring. The Master node is responsible for the intelligence of the Ring and monitors the status of the Ring. The Master node checks the status of the Ring by sending Ring Health Frames (RHF) around the Ring from its Primary port and returning on its Secondary port. If the Master node misses three consecutive RHFs, it determines the ring to be in a failed state. The Master then sends a Topology Change RHF to the Transit Nodes informing them that the ring has changed. This causes the Transit Nodes to flush their forwarding tables, and re-converge to the new network structure.

One port of the Master node is designated the Primary port (P) to the ring; another port is designated as the Secondary port (S) to the ring. In normal operation, the Master node blocks the Secondary port for all non-control traffic belonging to this FRRP group, thereby avoiding a loop in the ring, like STP. Layer 2 switching and learning mechanisms operate per existing standards on this ring.

Each Transit node is also configured with a Primary port and a Secondary port on the ring, but the port distinction is ignored as long as the node is configured as a Transit node. If the ring is complete, the Master node logically blocks all data traffic in the transmit and receive directions on the Secondary port to prevent a loop. If the Master node detects a break in the ring, it unblocks its Secondary port and allows data traffic to be transmitted and received through it. See Figure 15-1 for a simple example of this FRRP topology. Note that ring direction is determined by the Master node's Primary and Secondary ports.

Figure 15-1. Normal Operating FRRP Topology



A Virtual LAN (VLAN) is configured on all node ports in the ring. All ring ports must be members of the Member VLAN and the Control VLAN.

The Member VLAN is the VLAN used to transmit data as described earlier.

The Control VLAN is used to perform the health checks on the ring. The Control VLAN can always pass through all ports in the ring, including the secondary port of the Master node.

Ring Status

The Ring Failure notification and the Ring Status checks provide two ways to ensure the ring remains up and active in the event of a switch or port failure.

Ring Checking

At specified intervals, the Master Node sends a Ring Health Frame (RHF) through the ring. If the ring is complete, the frame is received on its secondary port, and the Master node resets its fail-period timer and continues normal operation.

If the Master node does not receive the Ring Health Frame (RHF) before the fail-period timer expires (a configurable timer), the Master node moves from the Normal state to the Ring-Fault state and unblocks its Secondary port. The Master node also clears its forwarding table and sends a control frame to all other nodes, instructing them to also clear their forwarding tables. Immediately after clearing its forwarding table, each node starts learning the new topology.

Ring Failure

If a Transit node detects a link down on any of its ports on the FRRP ring, it immediately sends a link-down control frame on the Control VLAN to the Master node. When the Master node receives this control frame, the Master node moves from the Normal state to the Ring-Fault state and unblocks its Secondary port. The Master node clears its routing table, and sends a control frame to all other ring nodes, instructing them to clear their routing tables as well. Immediately after clearing its routing table, each node begins learning the new topology.

Ring Restoration

The Master node continues sending Ring Health Frames out its primary port even when operating in the Ring-Fault state. Once the ring is restored, the next status check frame is received on the Master node's Secondary port. This will cause the Master node to transition back to the Normal state. The Master node then logically blocks non-control frames on the Secondary port, clears its own forwarding table, and sends a control frame to the Transit nodes, instructing them to clear their forwarding tables and re-learn the topology.

During the time between the Transit node detecting that its link is restored and the Master node detecting that the ring is restored, the Master node’s Secondary port is still forwarding traffic. This can create a temporary loop in the topology. To prevent this, the Transit node places all the ring ports transiting the newly restored port into a temporary blocked state. The Transit node remembers which port has been temporarily blocked and places it into a pre- forwarding state. When the Transit node in the pre-forwarding state receives the control frame instructing it to clear its routing table, it does so and unblocks the previously blocked ring ports on the newly restored port. Then the Transit node returns to the Normal state.

Multiple FRRP Rings

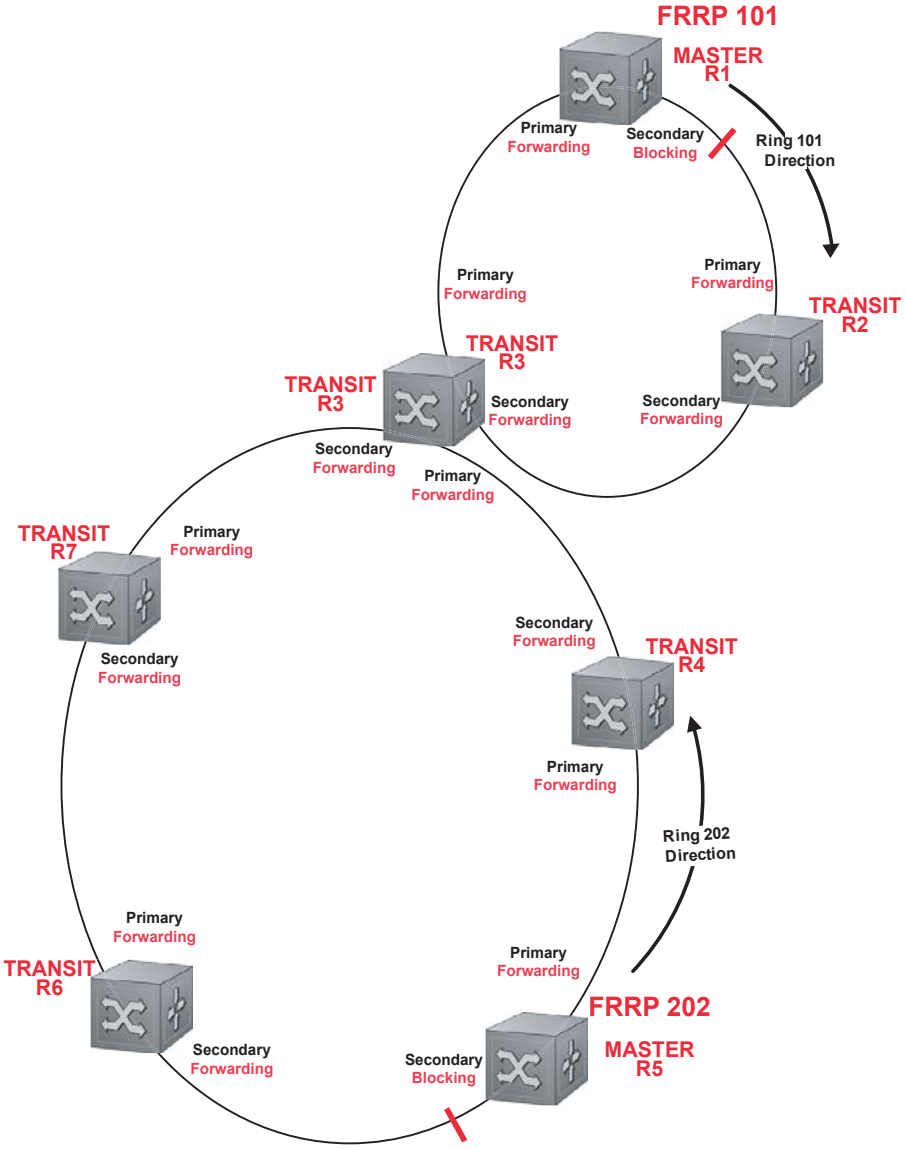
Up to 255 rings allowed per system. However, it is not recommended on the S-Series to have more than 34 rings on the same interface (either a physical interface or a portchannel). More than the recommended number of rings may cause interface instability. Multiple rings can be configured with a single switch connection; a single ring can have multiple FRRP groups; multiple rings can be connected with a common link.

Member VLAN Spanning Two Rings Connected by One Switch

A Member VLAN can span two rings interconnected by a common switch, in a figure-eight style topology. A switch can act as a Master node for one FRRP Group and a Transit for another FRRP group, or it can be a Transit node for both rings.

In the example shown in Figure 15-2, FRRP 101 is a ring with its own Control VLAN, and FRRP 202 has its own Control VLAN running on another ring. A Member VLAN that spans both rings is added as a Member VLAN to both FRRP groups. Switch R3 has two instances of FRRP running on it: one for each ring. The example topology that follows shows R3 assuming the role of a Transit node for both FRRP 101 and FRRP 202.

Figure 15-2. Example of Multiple Rings Connected by Single Switch



Important FRRP Points

FRRP provides a convergence time that can generally range between 150ms and 1500ms. The Master node originates a high-speed frame that circulates around the ring. This frame, appropriately, sets up or breaks down the ring.

- A single FRRP flap will occur wen a line card is reset or a stack unit fails over to the standby.

- Ring Status Check Frames are transmitted by the Master Node at specified intervals.
- Multiple physical rings can be run on the same switch.
- One Master node is supported per ring. All other nodes are Transit nodes.
- Each node has 2 member interfaces: Primary and Secondary.
- There is no limit to the number of nodes on a ring.
- The Master node ring port states are: blocking, pre-forwarding, forwarding, and disabled.
- The Transit node ring port states are: blocking, pre-forwarding, forwarding, and disabled/
- STP is disabled on ring interfaces.
- The Master node secondary port is in blocking state during Normal operation.
- Ring Health Frames (RHF)
  - Hello RHF
    - — Sent at 500ms (hello interval)
    - — Transmitted and processed by Master node only
  - Topology Change RHF
    - — Triggered updates
    - — Processed at all nodes

Important FRRP Concepts

Table 15-1 lists some important FRRP concepts.

Table 15-1. FRRP Components

Concept	Explanation
Ring ID	Each *ring* has a unique 8-bit ring ID through which the ring is identified (e.g. FRRP 101 and FRRP 202 as shown in Figure 15-2.
Control VLAN	Each *ring* has a unique Control VLAN through which tagged Ring Health Frames (RHF) are sent. Control VLANs are used only for sending Ring Health Frames, and cannot be used for any other purpose.
Member VLAN	Each *ring* maintains a list of member VLANs. Member VLANs must be consistent across the entire ring.
Port Role	Each *node* has two ports for each ring: Primary and Secondary. The Master node Primary port generates Ring Health Frames (RHF). The Master node Secondary port receives the RHF frames. On Transit nodes, there is no distinction between a Primary and Secondary interface when operating in the Normal state.

Table 15-1. FRRP Components

Concept	Explanation
Ring Interface State	Each interface (*port*) that is part of the ring maintains one of four states • **Blocking State**: Accepts ring protocol packets but blocks data packets. LLDP, FEFD, or other Layer 2 control packets are accepted. Only the master node Secondary port can enter this state. • **Pre-Forwarding State**: A transition state before moving to the Forward state. Control traffic is forwarded but data traffic is blocked. The Master node Secondary port transitions through this state during ring bring-up. All ports transition through this state when a port comes up. • **Forwarding State**—Both ring control and data traffic is passed. When the ring is in Normal operation, the Primary port on the Master node and both Primary and Secondary ports on the Transit nodes are in forwarding state. When the ring is broken, all ring ports are in this state. • **Disabled State**—When the port is disabled or down, or is not on the VLAN.
Ring Protocol Timers	**Hello Interval**: The interval when ring frames are generated from the Master node's Primary interface (default 500 ms). The Hello interval is configurable in 50 ms increments from 50 ms to 2000 ms. **Dead Interval**: The interval when data traffic is blocked on a port. The default is 3 times the Hello interval rate. The dead interval is configurable in 50 ms increments from 50 ms to 6000 ms.
Ring Status	The state of the FRRP ring. During initialization/configuration, the default ring status is Ring-down (disabled). The Primary and Secondary interfaces, Control VLAN, and Master and Transit node information must be configured for the ring to be up. • **Ring-Up**: Ring is up and operational • **Ring-Down**: Ring is broken or not set up
Ring Health-check Frame (RHF)	Two types of RHFs are generated by the Master node. RHFs never loop the ring because they terminate at the Master node's secondary port. • **Hello RHF** (**HRHF**): These frames are processed only on the Master node's Secondary port. The Transit nodes pass the HRHF through the without processing it. An HRHF is sent at every Hello interval. • **Topology Change RHF** (**TCRHF**): These frames contains ring status, keepalive, and the Control and Member VLAN hash. It is processed at each node of the ring. TCRHFs are sent out the Master Node's Primary and Secondary interface when the ring is declared in a Failed state with the same sequence number, on any topology change to ensure all Transit nodes receive it. There is no periodic transmission of TCRHFs. The TCRHFs are sent on triggered events of ring failure or ring restoration only.

Implementing FRRP

- FRRP is media and speed independent.
- FRRP is a Dell Force10 proprietary protocol that does not interoperate with any other vendor.
- Spanning Tree must be disabled on both Primary and Secondary interfaces before FRRP is enabled.
- All ring ports must be Layer 2 ports. This is required for both Master and Transit nodes.
- A VLAN configured as control VLAN for a ring cannot be configured as a control or member VLAN for any other ring.

- The Control VLAN is used to carry any data traffic; it carries only RHFs.
- The Control VLAN cannot have members that are not ring ports.
- If multiple rings share one or more member VLANs, they cannot share any links between them.
- Member VLANs across multiple rings are not supported in Master nodes.
- Each ring has only one Master node; all others are transit nodes.

FRRP Configuration

These are the tasks to configure FRRP.

- Create the FRRP group
- Configure the Control VLAN
  - Configure Primary and Secondary ports
- Configure and add the Member VLANs
  - Configure Primary and Secondary ports
- Configure the Master node
- Configure a Transit node
- Set FRRP Timers (optional)
- Enable FRRP

Other FRRP related commands are:

- Clear FRRP counters

Create the FRRP group

The FRRP group must be created on each switch in the ring.

Use the commands in the following sequence to create the FRRP group.

Command Syntax	Command Mode	Purpose
protocol frrp *ring-id*	CONFIGURATION	Create the FRRP group with this Ring ID Ring ID: 1-255

Configure the Control VLAN

Control and Member VLANS are configured normally for Layer 2. Their status as Control or Member is determined at the FRRP group commands. For complete information about configuring VLANS in Layer 2 mode, see Chapter 25, Layer 2.

Be sure to follow these guidelines:

- All VLANS must be in Layer 2 mode.
- Only ring nodes can be added to the VLAN.
- A Control VLAN can belong to one FRRP group only.
- Control VLAN ports must be tagged.
- All ports on the ring must use the same VLAN ID for the Control VLAN.
- A VLAN cannot be configured as both a Control VLAN and Member VLAN on the same ring.
- Only two interfaces can be members of a Control VLAN (the Master Primary and Secondary ports).
- Member VLANs across multiple rings are not supported in Master nodes

Use the commands in the following sequence, on the switch that will act as the Master node, to create the Control VLAN for this FRRP group.

Step	Command Syntax	Command Mode	Purpose
1	**interface vlan** *vlan-id*	CONFIGURATION	Create a VLAN with this ID number VLAN ID: 1-4094
2	**tagged** *interface slot/port {range}*	CONFIG-INT-VLAN	Tag the specified interface or range of interfaces to this VLAN. Interface: • For a 10/100/1000 Ethernet interface, enter the keyword keyword **GigabitEthernet** followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. Slot/Port, Range: Slot and Port ID for the interface. Range is entered Slot/Port-Port.
3	**interface primary** *int slot/port* **secondary** *int slot/port* **control-vlan** *vlan id*	CONFIG-FRRP	Assign the Primary and Secondary ports, and the Control VLAN for the ports on the ring. Interface: • For a 10/100/1000 Ethernet interface, enter the keyword keyword **GigabitEthernet** followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information • For a SONET interface, enter the keyword **sonet** followed by slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. Slot/Port: Slot and Port ID for the interface. VLAN ID: The VLAN identification of the Control VLAN.
4	**mode** *master*	CONFIG-FRRP	Configure the Master node

Step	Command Syntax	Command Mode	Purpose
5	**member-vlan** *vlan-id {range}*	CONFIG-FRRP	Identify the Member VLANs for this FRRP group VLAN-ID, Range: VLAN IDs for the ring's Member VLANS.
6	**no disable**	CONFIG-FRRP	Enable FRRP

Configure and add the Member VLANs

Control and Member VLANS are configured normally for Layer 2. Their status as Control or Member is determined at the FRRP group commands. For complete information about configuring VLANS in Layer 2 mode, see Chapter 25, Layer 2.

Be sure to follow these guidelines:

- All VLANS must be in Layer 2 mode.
- Control VLAN ports must be tagged. Member VLAN ports except the Primary/Secondary interface can be tagged or untagged.
- The Control VLAN must be the same for all nodes on the ring.

Use the commands in the following sequence, on all of the Transit switches in the ring, to create the Members VLANs for this FRRP group.

Step	Command Syntax	Command Mode	Purpose
1	**interface vlan** *vlan-id*	CONFIGURATION	Create a VLAN with this ID number VLAN ID: 1-4094
2	**tagged** *interface slot/ port {range}*	CONFIG-INT-VLAN	Tag the specified interface or range of interfaces to this VLAN. Interface: • For a 10/100/1000 Ethernet interface, enter the keyword keyword **GigabitEthernet** followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information • For a SONET interface, enter the keyword **sonet** followed by slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. Slot/Port, Range: Slot and Port ID for the interface. Range is entered Slot/Port-Port.

Step	Command Syntax	Command Mode	Purpose
3	**interface primary** *int slot/port* **secondary** *int slot/port* **control-vlan** *vlan id*	CONFIG-FRRP	Assign the Primary and Secondary ports, and the Control VLAN for the ports on the ring. Interface: • For a 10/100/1000 Ethernet interface, enter the keyword keyword **GigabitEthernet** followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information • For a SONET interface, enter the keyword **sonet** followed by slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. Slot/Port: Slot and Port ID for the interface. VLAN ID: Identification number of the Control VLAN
4	**mode** *transit*	CONFIG-FRRP	Configure a Transit node
5	**member-vlan** *vlan-id {range}*	CONFIG-FRRP	Identify the Member VLANs for this FRRP group VLAN-ID, Range: VLAN IDs for the ring's Member VLANs.
6	**no disable**	CONFIG-FRRP	Enable this FRRP group on this switch.

Set FRRP Timers

Step	Command Syntax	Command Mode	Purpose
1	**timer** *{hello-interval|dead-interval} milliseconds*	CONFIG-FRRP	Enter the desired intervals for Hello-Interval or Dead-Interval times. Hello-Interval: 50-2000, in increments of 50 (default is 500) Dead-Interval: 50-6000, in increments of 50 (default is 1500) The Dead-Interval time should be set at 3x the Hello-Interval.

Clear FRRP counters

Use one of the following commands to clear the FRRP counters.

Command Syntax	Command Mode	Purpose
clear frrp *ring-id*	EXEC PRIVELEGED	Clear the counters associated with this Ring ID Ring ID: 1-255

Command Syntax	Command Mode	Purpose
clear frrp	EXEC PRIVELEGED	Clear the counters associated with all FRRP groups

Show FRRP configuration

Use the following command to view the configuration for the FRRP group.

Command Syntax	Command Mode	Purpose
show configuration	CONFIG-FRRP	Show the configuration for this FRRP group

Show FRRP information

Use one of the following commands show general FRRP information.

Command Syntax	Command Mode	Purpose
show frrp *ring-id*	EXEC *or* EXEC PRIVELEGED	Show the information for the identified FRRP group. Ring ID: 1-255
show frrp summary	EXEC *or* EXEC PRIVELEGED	Show the state of all FRRP groups. Ring ID: 1-255

Troubleshooting FRRP

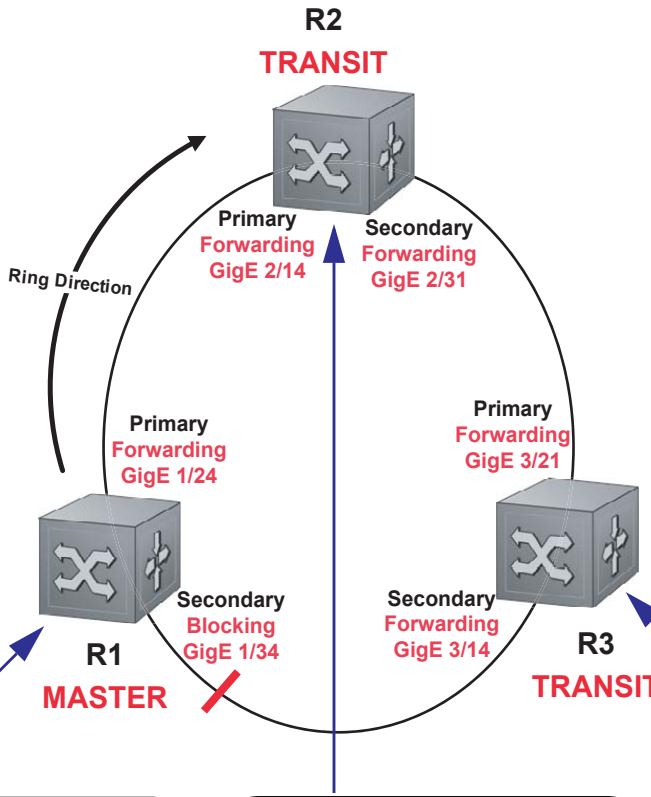
Configuration Checks

- Each Control Ring must use a unique VLAN ID.
- Only two interfaces on a switch can be Members of the same Control VLAN.
- There can be only one Master node for any FRRP Group.
- FRRP can be configured on Layer 2 interfaces only.
- Spanning Tree (if enabled globally) must be disabled on both Primary and Secondary interfaces when FRRP is enabled.
  - When the interface ceases to be a part of any FRRP process, if Spanning Tree is enabled globally, it must be enabled explicitly for the interface.
- The maximum number of rings allowed on a chassis is 255.

Sample Configuration and Topology

Figure 15-3 is an example of a basic FRRP topology. Below the figure are the associated CLI commands.

Figure 15-3. Basic Topology and CLI commands



R1 MASTER

```
interface GigabitEthernet 1/24
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/34
 no ip address
 switchport
 no shutdown
!
interface Vlan 101
 no ip address
 tagged GigabitEthernet 1/24,34
 no shutdown
!
interface Vlan 201
 no ip address
 tagged GigabitEthernet 1/24,34
 no shutdown

!
protocol frrp 101
 interface primary GigabitEthernet 1/24 secondary GigabitEthernet 1/34 control-vlan 101
 member-vlan 201
 mode master
 no disable
```

R2 TRANSIT

```
interface GigabitEthernet 2/14
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 2/31
 no ip address
 switchport
 no shutdown
!
interface Vlan 101
 no ip address
 tagged GigabitEthernet 2/14,31
 no shutdown
!
interface Vlan 201
 no ip address
 tagged GigabitEthernet 2/14,31
 no shutdown
!
protocol frrp 101
 interface primary GigabitEthernet 2/14 secondary GigabitEthernet 2/31 control-vlan 101
 member-vlan 201
 mode transit
 no disable
```

R3 TRANSIT

```
interface GigabitEthernet 3/14
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 3/21
 no ip address
 switchport
 no shutdown
!
interface Vlan 101
 no ip address
 tagged GigabitEthernet 3/14,21
 no shutdown
!
interface Vlan 201
 no ip address
 tagged GigabitEthernet 3/14,21
 no shutdown

!
protocol frrp 101
 interface primary GigabitEthernet 3/21 secondary GigabitEthernet 3/14 control-vlan 101
 member-vlan 201
 mode transit
 no disable
```

16

Force10 Service Agent

Force10 Service Agent is supported on platforms: C E

FTSA is supported on the E-Series ExaScale platform with FTOS 8.2.1.0 and later.

Accurate and timely resolution of problems in your system or network requires gathering relevant data at the time a condition manifests, and getting that information to administrators as soon as possible.

Proper data collection is often impeded because:

- a problem is serious enough that the initial reaction is to reboot the router, which might eliminate the opportunity to gather symptomatic data.
- a data collection plan is complex or misunderstood.
- a problem is intermittent and so collection opportunities are missed.

Force10 Service Agent (FTSA) is designed automate data collection to relieve these issues. It periodically monitors Dell Force10 or user-specified system variables. If a match condition exists, it triggers data collection via the CLI. For example, you can configure FTSA to search for a specific value in the **show** command for output throttles on an interface if CPU usage exceeds 85%. FTSA then automatically E-mails the information in XML format to network administrators, and/or the Dell Force10 Technical Assistance Center.

Implementation Information

It is possible to omit the **admin email** and **smtp server-address** configurations and instead log messages locally (on the Dell Force10 system itself). The **enable all** command can therefore have several outcomes depending on the configuration prior to execution:

1. When no servers are configured and **no enable all** is configured (the default), all policies including log-only policies are inactive.
2. When no servers are configured and **enable all** is configured, log-only policies are active, and messages are logged to the internal flash. All other (no log-only) policies are active.
3. When servers are configured and **enable all** is configured, log-only policies log messages to the internal flash while policies that have no log-only log to the configured email servers.

Configure Force10 Service Agent

The minimal FTSA configuration is four steps:

1. Enable FTSA. See page 348.
2. Specify the SMTP server to which FTSA will send E-mails upon a trigger event. See page 349.
3. Specify the source E-mail address that FTSA should use when generating E-mails. See page 349.
4. Enable the FTSA messaging service for the default recipient (see page 350,) or:
   a Configure a non-default recipient. See page 351.
   a Enable FTSA messaging for the recipient. See page 350.

Related Configuration Tasks

The following configuration tasks are optional, but perform them so that FTSA messaging functionality is fully enabled.

- FTSA Messaging Service on page 350
- FTSA Policies on page 357
- Debugging FTSA on page 371

Enable Force10 Service Agent

FTSA is disabled by default.

Task	Command Syntax	Command Mode
Enable FTSA. • If FTSA is disabled when you execute this command, then FTOS starts the FTSA service and enters the CALLHOME context. • If FTSA is enabled when you enter this command, FTOS only enters the CALLHOME context.	**call-home**	CONFIGURATION
Disable FTSA. Note: This command deletes the FTSA configuration from the running-config. To disable the FTSA messaging function only, use the command **no enable-all** from CALL HOME mode.	**no call-home**	CONFIGURATON

The system displays Message 1 when you enable or disable FTSA. Figure 16-1 shows the default FTSA configuration.

Message 1 FTSA Enabled/Disabled

```
%RPM0-P:CP %CALL-HOME-3-CALLHOME: Call-home service started
%RPM0-P:CP %CALL-HOME-3-CALLHOME: Call-home service ended.
```

Figure 16-1. Displaying the Default FTSA Configuration

```
FTOS(conf-callhome)#show config
!
call-home
 no enable-all
 server Force10
  recipient ftsa@force10networks.com
  keyadd Force10DefaultPublicKey
  encrypt
  no enable
```

Specify an SMTP Server for FTSA

To specify the SMTP server that will receive and forward the E-mail messages generated by FTSA:

Task	Command Syntax	Command Mode
Specify an SMTP server in the form *smtp.domain-name.com*.	**smtp server-address**	CALLHOME

Provide an Administrator E-mail Address

FTSA is designed to send E-mail notification when a test condition is met, Type 4 messaging is enabled, and when Type 3 messaging is enabled.

Step	Task	Command Syntax	Command Mode
1	The administrator E-mail address is the one that FTSA uses to originate E-mails. • Enter the administrator's full E-mail address, in the form: *username@domain.com*, or • Enter the username without the domain name. Dell Force10 recommends using the system name for *username* your company's domain name for *domain.*	**admin-email** *email-address*	CALLHOME
2	If you did not enter the domain name when entering the administrator E-mail address, enter a domain name in the form *domain-name.com* If you specify a domain with both the **admin-email** and **domain-name** command the **domain-name** configuration supersedes.	**domain-name** *domain_name*	CALLHOME

FTSA Messaging Service

The purpose of FTSA is to automatically send information about the switch to the network administrators or Dell Force10 TAC, so that when there is a network problem, the relevant information is collected at the time the problem manifests.

Enable the FTSA Messaging Service

There are five FTSA message types (see FTSA Message Types on page 355 for examples):

- **Type 0**: Call Home Enable
- **Type 1**: Call Home Disable
- **Type 2**: Chassis failover
- **Type 3**: Inventory
- **Type 4**: System Log
- **Type 5**: Action List

The E-mail body of every message always contains the message type, chassis name, transmission time, and, when encrypted, the public encryption key.

FTSA is pre-configured to send PGP5-encrypted E-mails containing basic system inventory information to the Dell Force10 Technical Assistance Center (TAC) at ftsa@force10networks.com, as shown in Figure 16-2. However messaging (for all recipients) is disabled by default.

Note: You may not remove the Dell Force10 server label or default recipient, but you may modify either.

You must still explicitly enable messaging for each recipient, including the default recipient.

Each recipient has a (user-configurable) mnemonic label. FTOS creates a CLI context based on this label from which you can enable messaging and modify the E-mail parameters for the recipient. You can enter the context for a recipient by entering the command **server** *label* from the CALLHOME context. For example, the default label is *Force10*. Enter the context *conf-callhome-Force10* by entering the command **server Force10**, as shown in Figure 16-2.

You may enable messaging for all recipients at once, or enable messaging for each recipient individually.

Task	Command	Command Mode
Enable messaging for all recipients.	**enable-all**	CALLHOME
Enable messaging for a individual recipient.	**enable**	CALLHOME <SERVER-LABEL>

Add Additional Recipients of FTSA E-mails

You can add four more recipients for FTSA E-mails, in addition to Dell Force10 TAC and the administrator, for a total of five recipients.

To add a recipient, you first create a mnemonic label for it. FTOS uses this label to create an FTOS context in which you can configure the E-mail parameters for the recipient. For example, the default recipient is Dell Force10 TAC and the label for this recipient is *Force10*. FTOS creates the context *conf-callhome-Force10* in which you can configure the parameters for E-mails destined for Dell Force10 TAC only, as shown in Figure 16-2.

Figure 16-2. Displaying the Default FTSA E-mail Recipient Configuration

```
FTOS(conf-callhome)#show config
!
call-home
 no enable-all
 server Force10
  recipient ftsa@force10networks.com
  keyadd Force10DefaultPublicKey
  encrypt
  no enable
FTOS(conf-callhome-Force10)#?
enable                  Enable call-home service for this server
encrypt                 Encrypt emails for this server
end                     Exit from configure mode
exit                    Exit from call-home server mode
keyadd                  Add server's public key for encryption
log-messages            Add log information
no                      Reset a command
recipient               Enter server email
show                    Show call-home server configuration
FTOS(conf-callhome-Force10)#show config
 server Force10
  recipient ftsa@force10networks.com
  keyadd Force10DefaultPublicKey
  encrypt
  no enable
```

To add a recipient:

Step	Task	Command	Command Mode
1	Create a mnemonic label for the recipient.	**server** *label*	CONFIGURATION
2	Enter the recipient E-mail address in the form *username@domain-name.com.*	**recipient** *email-address*	CONFIGURATION

Encrypt FTSA Messages

Encrypting FTSA message to a recipient other than the default is supported only on platforms: C E

Per-recipient, you have a choice of sending FTSA E-mails in clear text or with PGP5 encryption. Messages to the default recipient are configured for encryption using a public encryption key, as shown in Figure 16-2.

Step	Task	Command	Command Mode
1	Copy the encryption key file to the internal flash. The key *Force10DefaultPublicKey* for the default recipient is packed with FTOS, so enable encryption for it, proceed to Step 3.	**copy** *source-path*/*file* **flash://** *keyfilename*	EXEC Privilege
2	Specify the key with which E-mails to the recipient will be encrypted.	**keyadd** *keyfilename*	CALLHOME<SERVER-LABEL>
3	Encrypt E-mails to the recipient.	**encrypt**	CALLHOME<SERVER-LABEL>

Create a PGP5 encryption key

Step	Task
1	Use a PGP5-compatible program such as PGP or GnuPG to generate the public or private key. The user name that you choose in the program will be the one that you use in the **server** command.
2	Export the public key to a file.

Provide Administrator Contact Information

Dell Force10 recommends that you provide administrator contact information so that it can be included in Type 3 or greater E-mails.

Task	Command	Command Mode
Provide the postal service mailing address at which the network administrator can be contacted.	**contact-address**	CALLHOME
Provide the E-mail address at which the network administrator can be contacted.	**contact-email**	CALLHOME
Provide the name of the network administrator to be contacted upon an FTSA trigger event.	**contact-name**	CALLHOME
Include a memo in FTSA messages.	**contact-notes**	CALLHOME
Provide the phone number of the network administrator to be contacted upon an FTSA trigger event.	**contact-phone**	CALLHOME
Note: The contact fields may contain any character, however, be aware that FTSA generates all messages are XML format, for example <contact-name> </contact-name>, so non-alphanumeric characters might create XML errors.		

Set the Frequency of FTSA Type 3 Messages

When messaging is enabled, FTSA sends an E-mail every 24 hours containing inventory information to all recipients. There is no facility for setting the frequency for individual recipients.

Task	Command	Command Mode
Set the frequency at which FTSA generates inventory E-mails. Range: 2 to 10080 minutes Default: 1440 minutes (24 hours)	**frequency** *minutes*	CALLHOME

Generate FTSA Type 4 Messages

FTSA can collect and E-mail a user-defined subset of the local system log.These E-mails are Type 4, and Type 4 messages are *enabled* by default:

Step	Task	Command	Command Mode
1	Ensure that system logging is on, and verify the logging severity level.	**logging on** **do show running-config logging**	CONFIGURATION
2	Collect and E-mail system log messages. • If you do not use the **severity** option, FTSA uses 7 by default. • This is the recommended severity level. Lower values will result in partial log data sent to the server because messages with higher values are filtered out.	**log-messages** [**delay** *minutes*] [**severity** *level*] [**filter** *string*]	CONFIGURATION<SERVER>

After the initial message, subsequent messages only include log entries with that were generated after the last message was sent to the server so that log messages are not repeated in multiple E-mails.

Set Parameters FTSA Type 5 Messages

FTSA Type 5 messages have two configurable parameters:

Task	Command	Command Mode
Divert Type 5 messages to the internal flash directory /CALL-HOME-LOGS instead of sending them to FTSA recipients. Filenames include the action list name and a timestamp, and use the file extension *.ftsa*.	**log-only**	CALLHOME ACTIONLIST

Task	Command	Command Mode
All E-mails are generated in XML format by default. For Type 5 messages only, you may generate E-mails in clear text format. The configuration is per action list.	**message-format** {**xml** | **text**}	CALLHOME ACTIONLIST

FTOS Behavior: FTOS versions prior to 8.2.1.0 diverted Type 5 messages to the internal flash root directory when you enter the command log-only. Beginning in version 8.2.1.0, FTOS stores these messages in /CALL-HOME-LOGs on the internal flash.

FTSA Message Types

FTOS displays Message 2 every time FTSA sends a message.

Message 2 FTSA Message Sent

```
%RPM0-P:CP %CALL-HOME-HELPER-3-CALLHOME: Callhome service sent a message to Force10 at
pubslab@training10.com
```

FTSA generates Type 0 messages when you enable a recipient.

Figure 16-3. FTSA Type 0 Message

```
<AgentInfo>
<messagetype>Type - 0</messagetype>
<time>16:25:08.887 UTC Tue Feb 17 2009</time>
<serialnum>0026233      </serialnum>
<hostname>R1_E600</hostname>
<messagenum>0</messagenum>
</AgentInfo>
```

FTSA generates Type 1 messages when messaging is disabled. Messaging is disabled when:

- you **no-enable** a recipient or **no enable-all**
- messaging is enabled and you **no call-home**

Figure 16-4. FTSA Type 1 Message

```
<AgentInfo>
<messagetype>Type - 1</messagetype>
<time>19:49:51.581 UTC Wed Feb 18 2009</time>
<serialnum>0036232      </serialnum>
<hostname>Force10</hostname>
<messagenum>0</messagenum>
</AgentInfo>
```

FTSA generates Type 2 messages for a force or auto failover.

Figure 16-5. FTSA Type 2 Message

```
<AgentInfo>
<messagetype>Type - 2</messagetype>
<time>00:25:36.893 UTC Thu Feb 19 2009</time>
<serialnum>0036232      </serialnum>
<hostname>Force10</hostname>
<messagenum>0</messagenum>
</AgentInfo>
```

FTSA periodically generates Type 3 messages, which contain the output of the command **show inventory**.

Figure 16-6. FTSA Type 3 Message

```
---------------------------------Message Body-----------------------------------------
<AgentInfo>
<messagetype>Type - 3</messagetype>
<time>00:30:46.172 UTC Thu Feb 19 2009</time>
<serialnum>0036232      </serialnum>
<hostname>Force10</hostname>
<messagenum>0</messagenum>
</AgentInfo>

---------------------------------Message Attachment-----------------------------------
Chassis Type     : E300
Chassis Mode     : TeraScale
Software Version : 7.8.1.0

<hardware>
     E300               0036232        7520009603   D
  1  LC-EF3-GE-48T      0029961        7520009704   01
  0* LC-EF3-RPM         FX000017082    7520025400   04
  1  LC-EF3-RPM         0031177        7520013808   05
  0  CC-E-SFM **        0004903        7490007411    A
  1  CC-E-SFM           0010745        7520003706   A
  0  CC-E300-1200-AC    CKTE31420      7520022300   A
  1  CC-E300-1200-AC    CKTE31303      7520022300   A
  2  CC-E300-1200-AC    SJ32670        7520022300   A
  0  CC-E300-FAN        N/A            N/A          N/A

 * - standby

</hardware>
<software>

</software>
```

FTSA periodically generates Type 4 messages, only when Type 4 messaging is enabled, which contains system log messages.

Figure 16-7. FTSA Type 4 Messages

```
---------------------------------Message Body-----------------------------------------
<AgentInfo>
<messagetype>Type - 4</messagetype>
<time>01:57:46.453 UTC Fri Feb 20 2009</time>
<serialnum>0036232      </serialnum>
<hostname>Force10</hostname>
<messagenum>0</messagenum>
</AgentInfo>

---------------------------------Message Attachment-----------------------------------
Chassis Type     : E300
Chassis Mode     : TeraScale
Software Version : 7.8.1.0

<log_messages>
how logging severity 7 session 1 | display xml
<?xml version="1.0" encoding="UTF-8" ?>
<response MajorVersion="1" MinorVersion="0">
<action>
<syslog_properties>
<logging>enabled</logging>
<console_level>debugging</console_level>
<monitor_level>debugging</monitor_level>
<buffer_level>debugging, 36 Messages Logged, Size (40960 bytes)</buffer_level>
<trap_level>informational</trap_level>
<bufferclearedat></bufferclearedat>
</syslog_properties>
<syslog>
<time>1d0h50m</time>
<card>RPM1-P</card>
<slot>CP</slot>
<facility>CALL</facility>
<severity>HOME</severity>
<msgid>HELPER-3-CALLHOME</msgid>
<msg> Callhome service sent a message to Force10 at pubslab@training10.com
</msg>
</syslog>
</action>
</response>
FTOS#
</log_messages>
```

For FTSA Type 5 Messages, see FTSA Policy Sample Configurations on page 364.

FTSA Policies

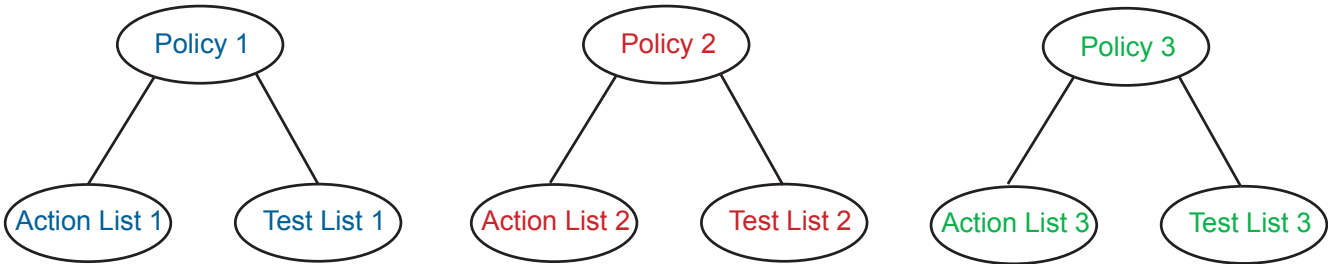
FTSA policies are a list of user-defined problematic conditions for which the FTSA periodically searches. If any of the conditions exist, FTSA executes a user-defined set of actions. FTOS allows up five active FTSA policies and an unlimited number of inactive ones.

To configure an FTSA policy:

1. Create a policy-test-list. See Create an FTSA Policy Test List on page 358.

2. Create the list of actions that FTSA should take if any of the conditions exist. See Create a Policy Action List on page 361.
3. Create a policy and assign a test list and action list. See Create a Policy and Assign a Test and Action List on page 363.
4. Set optional policy parameters. See Additional Policy Configurations on page 364

Figure 16-8. Creating FTSA Policies



Create an FTSA Policy Test List

Create the list of conditions for which FTSA should search. You may include a pre-defined list (Table 16-1) and specify additional test conditions (Table 16-2).

To create a new, empty policy test list:

Task	Command	Command Mode
Create a policy test list and name it.	**policy-test-list** *name*	CALLHOME

Choose test conditions for a policy test list

Once you create a policy test list, FTOS enters the CALLHOME TESTLIST context. The list you created is initially empty. You may choose one of three pre-defined condition lists, or create your own.

The three pre-defined condition lists are shown in Table 16-1.

Table 16-1. Pre-defined Policy Test Lists

Condition	Conditions Tested
Exception	CPU above 85%, crash, task crash, dump, reload due to error, RPM failover due to error

Table 16-1. Pre-defined Policy Test Lists

Condition	Conditions Tested
Hardware	**Hardware Status** SFM status transition from active to another state RPM status transition from active to another state Line Card transition from active to another state Port-pipe error or transition to down PEM transition from up to another state AC PSU transition from up to other state Fan Tray down or individual fan down **Environmental status** Overtemp for each item listed in **show environment** Over/under-voltage for each item listed in **show environment**
Software	SWP Timeout, IPC Timeout, IRC timeout, CPU > 85%, Memory usage > 85%

To add a pre-defined list of conditions to your policy test list:

Task	Command	Command Mode
Add a pre-defined list of conditions to your policy test list.	**default-test [exception | hardware | software]**	CALLHOME TESTLIST

Table 16-2 shows the test conditions that are available to add to a custom policy test list. See the Dell Force10 MIB for further description of the given Object Identifiers (OID). You may only specify one test condition within a policy.

Table 16-2. Custom Policy Test Conditions

Condition	Keyword	Description	OID
CPU Usage	**cpu-1-min**	CPU utilization in percentage for the last 1 minute.	chRpmCpuUtil1Min
	cpu-5-min	CPU utilization in percentage for the last 5 minutes.	chRpmCpuUtil5Min
Interface Rate/ Throttles	**interface-bit-rate**	Interface bandwidth in bits/second.	
	interface-rate	Interface bandwidth in packets/ second.	
	interface-throttles	The total number of valid ingress frames with a length or type field value equal to 0x8808 (Control Frame).	f10IfInThrottles f10IfOutThrottles
Interface Errors	**interface-crc**	Frames received with a length between 64 and 1518 octets, inclusive, that had an incorrect CRC.	f10IfInCRC

Table 16-2. Custom Policy Test Conditions

Condition	Keyword	Description	OID
Memory Usage	**memory-free**	Per-CPU free memory in Megabytes.	chSysProcessorMemSize * (1 - chRpmMemUsageUtil)
	memory-free-percent	Per-CPU total free memory in percent.	1 - chRpmMemUsageUtil
	memory-used	Per-CPU total memory usage in Megabytes.	chSysProcessorMemSize * chRpmMemUsageUtil
	memory-used-percent	Per-CPU total memory usage in percent.	chRpmMemUsageUtil
Match String	**cli-show-text**	Match a string within a command output.	N/A
WRED drops	**wred-drops**	A count of the frames that are dropped using a WRED policy because of excessive traffic.	f10IfOutWredDrops

To add a custom test condition to a policy test list:

Condition	Command	Command Mode
CPU Usage	**test-condition {cpu-1-min** | **cpu-5-min**} *boolean-comparison value* **sample** *number*	CALLHOME TESTLIST
Interface Rate/ Throttles	**test-condition** {**interface-bit-rate | interface-rate** | **interface-throttles**} [**input** | **output**] *slot-number boolean-comparison value* **sample** *number*	CALLHOME TESTLIST
Interface Errors	**test-condition interface-crc** *slot-number boolean-comparison value* **sample** *number*	CALLHOME TESTLIST
Memory Usage	**test-condition** {**memory-free | memory-free-percent | memory-used | memory-used-percent**} {*cpu* | **rpm-any**} *boolean-comparison value* **sample** *number*	CALLHOME TESTLIST
Match String	**cli-show-text "***show command***" contains** *string*	CALLHOME TESTLIST
WRED drops	**test-condition wred-drops** *slot-number boolean-comparison value* **sample** *number*	CALLHOME TESTLIST

The boolean comparison operators behave as follows:

- **decrease**—If the difference between samples, calculated by subtracting the first value from the last, is or less than or equal to the specified value, then the action list is executed.
- **equal-to**—If the value of the probed system variable is the same as the specified value, then the action list is executed.
- **greater-than**—If the value of the probed system variable is greater than the specified value, then the action list is executed.

- **increase**—If the difference between successive samples, calculated by subtracting the first value from the last, is greater than or equal to the previously sampled value, then the action list is executed.
- **less-than**—If the value of the probed system variable is less than the specified value, then the action list is executed.
- **not-equal-to**—If the value of the probed system variable is not the same as the specified value, then the action list is executed.
- **no-change**—If the compared samples are equal, then the action list is executed.

Set the match criterion for test lists

Task	Command	Command Mode
Match any one of the test-conditions, all test conditions, or all conditions during the same sample period. Default: **any**	**match** [**any | all | simultaneous**]	CALLHOME TESTLIST

Create a Policy Action List

Depending on your configuration, if any or all of the conditions in the policy test list exists, FTSA executes the actions contained in the policy action list. Data gathered from the actions is saved in a local file with the same name as the action list and a date and time stamp appended to the filename. FTSA does not overwrite files from previous executions.

While an action list is executing, pending action lists do not execute until the current action list completes. For example, if a test list matches a condition and triggers an action list, and during the execution of the action list another test list matches a condition, execution of the second action list is postponed until the first action list completes.

If a policy action list is unconfigured while executing, data already gathered is stored in a local file, and then data gathering is terminated.

To create a new, empty policy action list:

Task	Command	Command Mode
Create a policy action list and name it.	**policy-action-list** *name*	CALLHOME

Add actions to a policy action list

Once you create a policy action list, FTOS enters the CALLHOME ACTIONLIST context. The list you created is initially empty. You may choose one of three pre-defined action lists and add an unlimited number of custom actions.

Table 16-3. Pre-defined Action Lists

Action List	Keyword	Actions Executed
Exception	**exception**	**show processes cpu** **show processes memory** **show processes communication lp—**all line cards **show tech-support** **show trace** **show trace hardware** **show command history** **debug cpu-traffic-stats** **show cpu-traffic-stats—**x3,10s interval **show ip traffic —**x3,10s interval **no debug cpu-traffic-stats** **show cpu-traffic-stats—**x3,10s interval **show ip traffic—**x3,10s interval
Hardware	**hardware**	**show tech-support** **show trace** **show trace hardware** **show logging driverlog linecard**—all line cards **show logging driverlog cp** **show console lp**—all line cards **show pcdfo** **show command-history** **show cpu-interface-stats cp—**x2, 5s interval **show environment all** **show environment linecard-voltage**
Software	**software**	**show tech-support** **show trace** **show trace hardware** **show processes communication lp—**all line cards **show processes cpu** **show command history** **show processes ipc flow-control** **show processes ipc flow-control lp—**all line cards **show hardware rpm cp party-bus statistics**—x2, 10s interval **show hardware rpm cp data-plane statistics**—x2, 10s interval **show hardware rpm rp1 party-bus statistics**—x2, 10s interval **show hardware rpm rp1 data-plane statistics**—x2, 10s interval **show hardware rpm rp2 party-bus statistics**—x2, 10s interval **show hardware rpm rp2 data-plane statisticsq**— x2, 10s interval **show cpu-interface-stats cp**—x2, 10s interval

To add a pre-defined list of actions to your policy action list:

Task	Command	Command Mode
Add a pre-defined list of actions to your policy action list.	**default-action** [**exception | hardware | software**]	CALLHOME ACTIONLIST

You may add an unlimited number of three types of custom actions:

Task	Command	Command Mode
Execute a recovery action when FTSA reaches the test-limit. You may reload the chassis or reset an RPM or linecard. **Note:** The default test-limit is unlimited and under this condition the recovery action will never execute.	**seq** *number* **cli-action "***command***"**	CALLHOME ACTIONLIST
Execute a show debug when FTSA discovers a test condition. While debug is running, FTSA will execute other pending action list items.	**seq** *number* **cli-debug "***debug-command***" time** *seconds*	CALLHOME ACTIONLIST
Execute a **show** command when FTSA discovers a test condition.	**seq** *number* **cli-show "***show-command***" repeat** *number* **delay** *seconds*	CALLHOME ACTIONLIST
Reset an interface when FTSA discovers a test condition.	**seq** *number* **interface-reset** *interface* **delay** *seconds*	CALLHOME ACTIONLIST

Create a Policy and Assign a Test and Action List

An FTSA minimally must have a policy test list and policy action list assigned to it.

Step	Task	Command	Command Mode
1	Create an FTSA policy and name it.	**policy** *name*	CALLHOME POLICY
2	Assign a test list to a policy.	**test-list** *name*	CALLHOME POLICY
3	Assign a policy action list to a policy.	**action-list** *name*	CALLHOME POLICY

Additional Policy Configurations

Task	Command	Command Mode
Associate a Dell Force10 TAC case number with the policy. Configure a case number only if you already have a case open with Dell Force10 for the policy. This case number is included in action-list messages sent to Dell Force10.	**case-number** *number*	CALLHOME POLICY
Delay the subsequent execution of the test list after a match occurs. This configuration reduces the risk of burdening the CPU with sampling when a failure condition has already been detected. Default: 5 minutes	**dampen** *minutes*	CALLHOME POLICY
Associate a Dell Force10 Problem Report (PR) number with the policy. Configure a PR number only if you already have a case open with Dell Force10 for the policy. This PR number is included in action-list messages sent to Dell Force10.	**pr-number** *number*	CALLHOME POLICY
Execute the action list contingent upon the state of CPU utilization. The CPU utilization is calculated in percentage using the 1 minute rolling average of all RPM CPUs. Default: unconditional	**run-cpu** {*cpu* | {**cp** *slot* **| any}** **| rpm-any**} {{**less-than | greater-than**} *percentage*	CALLHOME POLICY
Note: If a test-list match is found, but the action-list does not execute because the **run-cpu less-than** condition is exists, then FTSA logs the test-list match and indicates that the action-list was prevented. However, if the trigger event was a high or specified CPU usage condition in the test-list, then the FTSA ignores the **run-cpu less-than** configuration.		
Specify the interval at which the system variables specified in the test-list are sampled. Default: 1 minute	**sample-rate**	CALLHOME POLICY
Execute the test list multiple times. If the test limit is greater than 1, the test list is executed immediately after the previous execution is complete. Default: unlimited	**test-limit**	CALLHOME POLICY

FTSA Policy Sample Configurations

Line card state-change policy configuration

The following FTSA policy configuration uses the default test list *hardware*, which contains a line-card-state-change condition, and the default action list *hardware* plus the custom action *show linecard 4 | grep Status*. Linecard 4 is then taken offline to trigger a match against the card-state-change test condition.

Figure 16-9. Configuring an FTSA Policy for a Linecard Down

```
call-home
 admin-email pubsadmin@training10.com
 smtp server-address 192.168.1.1
 no enable-all
 server Force10
  recipient pubslab@training10.com
  keyadd Force10DefaultPublicKey
  no encrypt
  enable
  log-messages delay 60 severity 6
 policy lcdown
  action-list lcdown
  test-list lcdown
 policy-test-list lcdown
  default-test hardware
 policy-action-list lcdown
  default-action hardware
  no log-only
  message-format text
  cli-show "show linecard 4 | grep Status" repeat 1 delay 1
```

Figure 16-10. System Log Messages during an a Linecard Down with FTSA

```
R6_E300(conf-callhome)#do offline linecard 4
2d9h24m: %RPM0-P:CP %CHMGR-2-CARD_DOWN: Line card 4 down - card offline
2d9h24m: %RPM0-P:CP %IFMGR-1-DEL_PORT: Removed port: Gi 4/0-47
R6_E300(conf-callhome)#2d9h24m: %RPM1-S:CP %IFMGR-1-DEL_PORT: Removed port: Gi 4/0-47
2d9h24m: %RPM0-P:CP %CALL-HOME-6-CALLHOME: Call-home executes remote exec command
2d9h24m: %RPM0-P:CP %CLI-0-REMOTE-EXEC: remote-exec CP: dhsTestCp
2d9h25m: %RPM0-P:CP %CALL-HOME-6-CALLHOME: Call-home executes remote exec command
2d9h25m: %RPM0-P:CP %CLI-0-REMOTE-EXEC: remote-exec CP: dhsTestCp
2d9h25m: %RPM0-P:CP %CALL-HOME-HELPER-3-CALLHOME: Callhome service sent a message to
Force10 at pubslab@training10.com
```

Figure 16-11. FTSA Type 5 Message for a Linecard Down Policy

```
---------------------------------Message Body-----------------------------------------
<AgentInfo>
<messagetype>Type - 5</messagetype>
<time>23:19:37.209 UTC Wed Feb 25 2009</time>
<serialnum>0036232      </serialnum>
<hostname>R6_E300</hostname>
<messagenum>0</messagenum>
</AgentInfo>
---------------------------------Message Attachment-----------------------------------
<action_list_message>
 <AgentInfo>
  <messagetype>Type - 5</messagetype>
  <time>23:19:37.230 UTC Wed Feb 25 2009</time>
  <serialnum>0036232      </serialnum>
 </AgentInfo>

 <contact_info>
 </contact_info>

 <F10_info>
  <policy_name>lcdown</policy_name>
 </F10_info>

 <action_list_name>lcdown</action_list_name>
 <test_list_match>
    <match>
     <test_condition>hardware</test_condition>
     <test_value>Line card 4 down</test_value>
    </match>
 </test_list_match>
 <content>
  <item>
   <item_name>show tech-support</item_name>
   <item_time>23:19:37.232 UTC Wed Feb 25 2009</item_time>
   <item_output>
show tech-support
[output omitted]
   </item_output>
  </item>
  <item>
   <item_name>show trace</item_name>
   <item_time>23:19:45.425 UTC Wed Feb 25 2009</item_time>
   <item_output>
show trace
[output omitted]
</item_output>
  </item>
  <item>
   <item_name>show trace hardware</item_name>
   <item_time>23:19:45.988 UTC Wed Feb 25 2009</item_time>
   <item_output>
show trace hardware
[output omitted]
</item_output>
```

Figure 16-12. FTSA Type 5 Message for a Linecard Down Policy (continued)

```
</item>
  <item>
   <item_name>show logging driverlog linecard 1</item_name>
   <item_time>23:19:46.191 UTC Wed Feb 25 2009</item_time>
   <item_output>
show logging driverlog linecard 1
[output omitted]
</item_output>
  </item>
  <item>
   <item_name>show logging driverlog linecard 4</item_name>
   <item_time>23:19:46.577 UTC Wed Feb 25 2009</item_time>
   <item_output>
show logging driverlog linecard 4
[output omitted]
</item_output>
  </item>
  <item>
   <item_name>show logging driverlog cp</item_name>
   <item_time>23:19:46.879 UTC Wed Feb 25 2009</item_time>
   <item_output>
show logging driverlog cp
[output omitted]
</item_output>
  </item>
  <item>
   <item_name>show console lp 1</item_name>
   <item_time>23:19:47.141 UTC Wed Feb 25 2009</item_time>
   <item_output>
show console lp 1
[output omitted]
</item_output>
  </item>
  <item>
   <item_name>show console lp 4</item_name>
   <item_time>23:19:50.686 UTC Wed Feb 25 2009</item_time>
   <item_output>
show console lp 4
[output omitted]
</item_output>
  </item>
  <item>
   <item_name>show pcdfo</item_name>
   <item_time>23:19:54.218 UTC Wed Feb 25 2009</item_time>
   <item_output>
show pcdfo
[output omitted]
</item_output>
  </item>
  <item>
   <item_name>show environment linecard-voltage</item_name>
   <item_time>23:19:54.246 UTC Wed Feb 25 2009</item_time>
   <item_output>
show environment linecard-voltage
[output omitted]
</item_output>
```

Figure 16-13. FTSA Type 5 Message for a Linecard Down Policy (continued)

```
</item>
  <item>
   <item_name>remote-exec cp dhsTestCp</item_name>
   <item_time>23:19:54.597 UTC Wed Feb 25 2009</item_time>
   <item_output>
remote-exec cp dhsTestCp
[output omitted]
</item_output>
  </item>
  <item>
   <item_name>remote-exec cp dhsTestCp</item_name>
   <item_time>23:20:00.663 UTC Wed Feb 25 2009</item_time>
   <item_output>
remote-exec cp dhsTestCp
[output omitted]
</item_output>
  </item>
  <item>
   <item_name>show linecard 4 | grep Status</item_name>
   <item_time>23:20:07.755 UTC Wed Feb 25 2009</item_time>
   <item_output>
show linecard 4 | grep Status
Status        : offline
Power Status  : AC
R6_E300#
   </item_output>
  </item>
 </content>
</action_list_message>
```

Figure 16-14. FTSA Type 5 Message for a BGP Peer Down Policy

```
---------------------------------Message Body-----------------------------------------
<AgentInfo>
<messagetype>Type - 5</messagetype>
<time>17:14:28.394 UTC Thu Feb 26 2009</time>
<serialnum>0036232      </serialnum>
<hostname>R6_E300</hostname>
<messagenum>0</messagenum>
</AgentInfo>
---------------------------------Message Attachment-----------------------------------
<action_list_message>
 <AgentInfo>
  <messagetype>Type - 5</messagetype>
  <time>17:14:28.415 UTC Thu Feb 26 2009</time>
  <serialnum>0036232      </serialnum>
 </AgentInfo>

 <contact_info>
 </contact_info>

 <F10_info>
  <policy_name>bgpdown</policy_name>
 </F10_info>

 <action_list_name>bgpdown</action_list_name>
 <test_list_match>
    <match>
     <test_condition>show logging 10</test_condition>
     <test_value>Search string "BGP-5-ADJCHANGE: Connection with neighbor 172.16.2.1
closed." is matched</test_value>
    </match>
 </test_list_match>
 <content>
  <item>
   <item_name>show ip bgp summary | grep 172.16.2.1</item_name>
   <item_time>17:14:28.417 UTC Thu Feb 26 2009</item_time>
   <item_output>
show ip bgp summary | grep 172.16.2.1
172.16.2.1      200               29       39          0    0     0 00:02:40 (shut)
R6_E300#
   </item_output>
  </item>
 </content>
</action_list_message>
```

Excessive CRC-error policy configuration

The following FTSA policy configuration uses the interface-crc match condition to monitor GigabitEthernet 1/2 for greater than 500 CRC errors. When this condition exists, FTSA triggers the action list, which captures a partial output of the command **show interfaces gigabitethernet 1/2**.

Figure 16-15. Configuring an FTSA Policy for an Excessive CRC-error Condition

```
call-home
 admin-email pubsadmin@training10.com
 smtp server-address 192.168.1.1
 no enable-all
 server Force10
  recipient pubslab@training10.com
  keyadd Force10DefaultPublicKey
  no encrypt
  enable
  no log-messages
 policy crcerror
  action-list crcerror
  test-list crcerror
 policy-test-list crcerror
  test-condition interface-crc 1 greater-than number 500
 policy-action-list crcerror
  no log-only
  message-format text
  cli-show "show int gig 1/2 | grep CRC" repeat 1 delay 1
```

Figure 16-16. System Syslog Messages during an Excessive CRC-error Condition

```
R6_E300(conf)#do show int gig 1/2 | grep CRC
     0 CRC, 0 overrun, 0 discarded
R6_E300(conf)#do show int gig 1/2 | grep CRC
     105 CRC, 0 overrun, 105 discarded
R6_E300(conf)#do show int gig 1/2 | grep CRC
     183 CRC, 0 overrun, 183 discarded
R6_E300(conf)#do show int gig 1/2 | grep CRC
     285 CRC, 0 overrun, 285 discarded
R6_E300(conf)#do show int gig 1/2 | grep CRC
     399 CRC, 0 overrun, 399 discarded
R6_E300(conf)#do show int gig 1/2 | grep CRC
     501 CRC, 0 overrun, 501 discarded
R6_E300(conf)#02:27:01: %RPM0-P:CP %CALL-HOME-HELPER-3-CALLHOME: Callhome service sent
a message to Force10 at pubslab@training10.com
```

Figure 16-17. FTSA Type 5 Message for an Excessive CRC-error Condition

```
---------------------------------Message Body-----------------------------------------
<AgentInfo>
<messagetype>Type - 5</messagetype>
<time>21:10:04.678 UTC Tue Mar 10 2009</time>
<serialnum>0036232      </serialnum>
<hostname>R6_E300</hostname>
<messagenum>0</messagenum>
</AgentInfo>
---------------------------------Message Attachment-----------------------------------
 <action_list_message>
  <AgentInfo>
   <messagetype>Type - 5</messagetype>
   <time>21:10:04.686 UTC Tue Mar 10 2009</time>
   <serialnum>0036232      </serialnum>
  </AgentInfo>
  <contact_info>

  <F10_info>
   <policy_name>crcerror</policy_name>
  </F10_info>

  <action_list_name>crcerror</action_list_name>
  <test_list_match>
     <match>
      <test_condition>interface-crc</test_condition>
      <test_value>The current value 501 is greater than the configured value 500</
test_value>
     </match>
  </test_list_match>
  <content>
   <item>
    <item_name>show int gig 1/2 | grep CRC</item_name>
    <item_time>19:01:07.368 UTC Tue Mar 10 2009</item_time>
    <item_output>
show int gig 1/2 | grep CRC
     501 CRC, 0 overrun, 501 discarded
R6_E300#
    </item_output>
   </item>
  </content>
</action_list_message>
```

Debugging FTSA

Display FTSA messages using the **debug call-home** command from EXEC Privilege mode.

Figure 16-18. Call-home Debug All during Type 5 Message Generation

```
#02:13:49 : CALL-HOME: Sending the following email
02:13:49 : From: pubsadmin@training10.com
        To:
        pubslab@training10.com
        Subject: <messagetype>Type - 5</messagetype>
        Attachment: ramdisk:/crcerror-21_10_04.685.txt
02:13:49 : Message:
<AgentInfo>
<messagetype>Type - 5</messagetype>
<time>21:10:04.678 UTC Tue Mar 10 2009</time>
<serialnum>0036232      </serialnum>
<hostname>R6_E300</hostname>
<messagenum>0</messagenum>
</AgentInfo>

02:13:49: %RPM0-P:CP %CALL-HOME-HELPER-3-CALLHOME: Callhome service sent a message to
Force10 at pubslab@training10.com
02:13:49 : Removing text file ramdisk:/crcerror-21_10_04.685.txt encrypt file
ramdisk:/crcerror-21_10_04.685.txt.asc
02:13:49 : CALL-HOME: Got action list status
```

17

GARP VLAN Registration Protocol

GARP VLAN Registration Protocol is supported on platform C E S

GVRP is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

Protocol Overview

Typical VLAN implementation involves manually configuring each Layer 2 switch that participates in a given VLAN. GARP VLAN Registration Protocol (GVRP), defined by the IEEE 802.1q specification, is a Layer 2 network protocol that provides for automatic VLAN configuration of switches. GVRP-compliant switches use GARP to register and de-register attribute values, such as VLAN IDs, with each other.

GVRP exchanges network VLAN information to allow switches to dynamically forward frames for one or more VLANs. Consequently, GVRP spreads this information and configures the needed VLAN(s) on any additional switches in the network. Data propagates via the exchange of GVRP protocol data units (PDUs).

The purpose of GVRP is to simplify (but not eliminate) static configuration. The idea is to configure switches at the edge and have the information dynamically propagate into the core. As such, the edge ports must still be statically configured with VLAN membership information, and they do not run GVRP. It is this information that is propagated to create dynamic VLAN membership in the core of the network.

Important Points to Remember

- GVRP propagates VLAN membership throughout a network. GVRP allows end stations and switches to issue and revoke declarations relating to VLAN membership.
- VLAN registration is made in the context of the port that receives the GARP PDU and is propagated to the other active ports.
- GVRP is disabled by default; you must enable GVRP for the switch and then for individual ports.
- Dynamic VLANs are aged out after the LeaveAll timer expires three times without receipt of a Join message. Use the **show gvrp statistics** {**interface** *interface* | **summary**} command to display status.
- Per-VLAN Spanning Tree (PVST+) or MSTP and GVRP cannot be enabled at the same time, as shown in Figure 17-1. If Spanning Tree and GVRP are both required, implement RSTP.

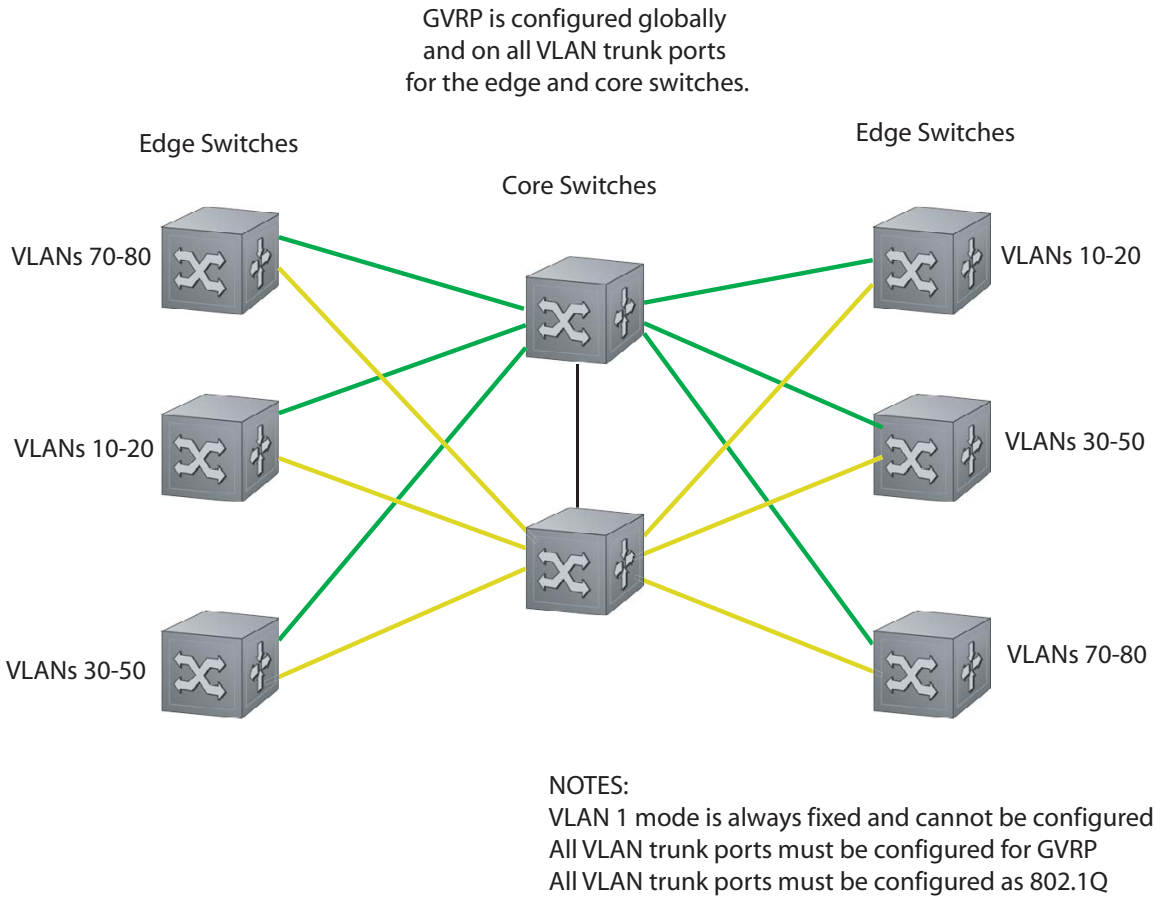
Figure 17-1. GVRP Compatibility Error Message

```
FTOS(conf)#protocol spanning-tree pvst
FTOS(conf-pvst)#no disable
% Error: GVRP running. Cannot enable PVST.

.........
FTOS(conf)#protocol spanning-tree mstp
FTOS(conf-mstp)#no disable
% Error: GVRP running. Cannot enable MSTP.

.........

FTOS(conf)#protocol gvrp
FTOS(conf-gvrp)#no disable
% Error: PVST running. Cannot enable GVRP.
% Error: MSTP running. Cannot enable GVRP.
```

Configuring GVRP

Globally, enable GVRP on each switch to facilitate GVRP communications. Then, GVRP configuration is per interface on a switch-by-switch basis. Enable GVRP on each port that connects to a switch where you want GVRP information exchanged. In Figure 17-2, that kind of port is referred to as a VLAN trunk port, but it is not necessary to specifically identify to FTOS that the port is a trunk port, as described in Chapter 18, VLAN Stacking, on page 367.

Figure 17-2. GVRP Configuration Overview



Basic GVRP configuration is a 2-step process:

1. Enable GVRP globally. See page 376.
2. Enable GVRP on an interface. See page 376.

Related Configuration Tasks

- Configuring GVRP Registration on page 376
- Configuring a GARP Timer on page 377

Enabling GVRP Globally

Enable GVRP for the entire switch using the command **gvrp enable** in CONFIGURATION mode, as shown in Figure 17-3. Use the **show gvrp brief** command to inspect the global configuration.

Figure 17-3. Enabling GVRP Globally

```
FTOS(conf)#protocol gvrp
FTOS(config-gvrp)#no disable
FTOS(config-gvrp)#show config
!
protocol gvrp
 no disable
FTOS(config-gvrp)#
```

Enabling GVRP on a Layer 2 Interface

Enable GVRP on a Layer 2 interface using the command **gvrp enable** in INTERFACE mode, as shown in Figure 17-4. Use **show config** from the INTERFACE mode to inspect the interface configuration, as shown in Figure 17-4, or use the **show gvrp** *interface* command in EXEC or EXEC Privilege mode.

Figure 17-4. Enabling GVRP on a Layer 2 Interface

```
FTOS(conf-if-gi-1/21)#switchport
FTOS(conf-if-gi-1/21)#gvrp enable
FTOS(conf-if-gi-1/21)#no shutdown
FTOS(conf-if-gi-1/21)#show config
!
interface GigabitEthernet 1/21
 no ip address
 switchport
 gvrp enable
 no shutdown
```

Configuring GVRP Registration

- **Fixed Registration Mode**: Configuring a port in fixed registration mode allows for manual creation and registration of VLANs, prevents VLAN de-registration, and registers all VLANs known on other ports on the port. For example, if an interface is statically configured via the CLI to belong to a VLAN, it should not be un-configured when it receives a Leave PDU. So, the registration mode on that interface is FIXED.
- **Forbidden Mode**: Disables the port to dynamically register VLANs, and to propagate VLAN information except information about VLAN 1. A port with forbidden registration type thus allows only VLAN 1 to pass through even though the PDU carries information for more VLANs. So, set the interface to the registration mode of FORBIDDEN if you do not want the interface to advertise or learn about particular VLANS.

Based on the configuration in the example shown in Figure 17-5, the interface 1/21 will not be removed from VLAN 34 or VLAN 35 despite receiving a GVRP Leave message. Additionally, the interface will not be dynamically added to VLAN 45 or VLAN 46, even if a GVRP Join message is received.

Figure 17-5. Configuring GVRP Registration

```
FTOS(conf-if-gi-1/21)#gvrp registration fixed 34,35
FTOS(conf-if-gi-1/21)#gvrp registration forbidden  45,46
FTOS(conf-if-gi-1/21)#show conf
!
interface GigabitEthernet 1/21
 no ip address
 switchport
 gvrp enable
 gvrp registration fixed 34-35
 gvrp registration forbidden 45-46
 no shutdown
FTOS(conf-if-gi-1/21)#
```

Configuring a GARP Timer

GARP timers must be set to the same values on all devices that are exchanging information using GVRP:

- **Join**: A GARP device reliably transmits Join messages to other devices by sending each Join message two times. Use this parameter to define the interval between the two sending operations of each Join message. The FTOS default is 200ms.
- **Leave**: When a GARP device expects to de-register a piece of attribute information, it will send out a Leave message and start this timer. If a Join message does not arrive before the timer expires, the information is de-registered. The Leave timer must be greater than or equal to 3x the Join timer. The FTOS default is 600ms.
- **LeaveAll**: Upon startup, a GARP device globally starts a LeaveAll timer. Upon expiration of this interval, it will send out a LeaveAll message so that other GARP devices can re-register all relevant attribute information. The device then restarts the LeaveAll timer to begin a new cycle. The LeaveAll timer must be greater than or equal to 5x of the Leave timer. The FTOS default is 10000ms.

Figure 17-6. Configuring GVRP Registration

```
FTOS(conf)#garp timer leav 1000
FTOS(conf)#garp timers leave-all 5000
FTOS(conf)#garp timer join 300

Verification:

FTOS(conf)#do show garp timer
GARP Timers     Value (milliseconds)
----------------------------------------
Join Timer        300
Leave Timer       1000
LeaveAll Timer    5000
FTOS(conf)#
```

FTOS displays Message 1 if an attempt is made to configure an invalid GARP timer.

Message 1 GARP Timer Error

```
FTOS(conf)#garp timers join 300
% Error: Leave timer should be >= 3*Join timer.
```

18

High Availability

High Availability is supported on platforms: C E S

High availability is a collection of features that preserves system continuity by maximizing uptime and minimizing packet loss during system disruptions.

To support all the features within the HA collection, you should have the latest boot code. The following table lists the boot code requirements as of this FTOS release.

Component	Boot Code
E-Series TeraScale RPM	2.4.2.1
E-Series TeraScale Line Card	2.3.2.1
E-Series ExaScale RPM	2.5.1.9
E-Series ExaScale Line Card	2.9.1.1
C-Series RPM	2.7.1.1
C-Series Line Card	2.6.0.2
S-Series RPM	2.8.2.0
S-Series Line Card	2.8.2.0

The features in this collection are:

Component Redundancy

Dell Force10 systems eliminates single points of failure by providing dedicated or load-balanced redundancy for each component.

RPM Redundancy

The current version of FTOS supports 1+1 hitless Route Processor Module (RPM) redundancy. The primary RPM performs all routing, switching, and control operations while the standby RPM monitors the primary RPM. In the event that the primary RPM fails, the standby RPM can assume control of the system without requiring a chassis reboot.

This section contains the following sub-sections:

Boot the chassis with a single RPM

You can boot the chassis with one RPM and later add a second RPM, which automatically becomes the standby RPM. FTOS displays Message 1 when the standby RPM is online.

Message 1 Standby RPM is Online

```
%RPM-2-MSG:CP0 %POLLMGR-2-ALT_RPM_STATE: Alternate RPM is present
%IRC-6-IRC_COMMUP: Link to peer RPM is up
%RAM-6-RAM_TASK: RPM1 is in Standby State.
```

On the C-Series, since the RPM also contains the switch fabric, even though the second RPM comes online as the standby, the switch fabric is active and participates in routing. You can achieve line rate on all line cards with a single RPM except for the 8-port 10G line card which requires both RPMs to achieve line rate.

Boot the chassis with dual RPMs

When you boot the system with two RPMs installed, the RPM in slot R0 is the primary RPM by default. Both RPMs should be running the same version of FTOS. You can configure either RPM to be the primary upon the next chassis reboot using the command **redundancy primary** from CONFIGURATION mode.

Version compatibility between RPMs

In general, the two RPMs should have the same FTOS version. However, FTOS tolerates some degree of difference between the two versions, as described in Table 18-1. View the configuration loaded on each RPM using the **show redundancy** command as shown in Figure 18-1.

Table 18-1. System Behavior with RPMs with Mismatched FTOS Versions

Mismatch Condition	Example	Behavior
different FTOS versions with only first two digits matching	Primary: **7.4**.2.0 Standby: **7.4**.1.0	The link to the standby RPM is up, and FTOS block syncs only the startup-config. The failover type is warm upgrade. FTOS displays Message 2.
different FTOS versions with first two digits not matching	Primary: **7.6**.1.0 Standby: **7.5**.1.0	The link to the standby RPM is down, and the standby RPM is in a boot loop. FTOS displays Message 3 and a boot fail prompt.
different FTOS versions with only first three digits matching	Primary: **7.4.2**.0 Standby: **7.4.2**.1	The link to the peer RPM is up, and FTOS performs a complete block sync. The failover type is hot failover. FTOS displays Message 2.

Message 2 FTOS Version Incompatibility Error

```
*********************************************
 *
 *     Warning !!!  Warning !!!  Warning !!!
 *
 * -------------------------------------------
 *
 *       Incompatible SW Version detected !!
 *
 *       This RPM -> 7.4.2.0
 *       Peer RPM -> 7.4.1.0
 *
 *********************************************

00:00:12: %RPM0-U:CP %IRC-4-IRC_VERSION: Current RPM 7.4.2.0 Peer RPM 7.4.1.0 - Different
software version detected
00:00:12: %RPM0-U:CP %IRC-6-IRC_COMMUP: Link to peer RPM is up
00:00:14: %RPM0-U:CP %RAM-6-ELECTION_ROLE: RPM0 is transitioning to Primary RPM.
```

Message 3 Boot Failure on Standby RPM

```
System failed to boot up. Please reboot the chassis !!!
00:12:46: %RPM1-U:CP %TME-0-RPM BRINGUP FAIL: FTOS failed to bring up the system
    Communication between RPMs is not up, check the software version and reboot chassis.

FTOS(standby)(bootfail)#
```

Automatic and manual RPM failover

RPM failover is the process of the standby RPM becoming the primary RPM. FTOS fails over to the standby RPM when:

1. communication is lost between the standby and primary RPMs
2. you request a failover via the CLI
3. you remove the primary RPM

Use the command **show redundancy** from EXEC Privilege mode to display the reason for the last failover.

Figure 18-1. Viewing RPM Redundancy Status

```
FTOS#show redundancy

--  RPM Status  --
-----------------------------------------------
 RPM Slot ID:               0
 RPM Redundancy Role:       Primary
 RPM State:                 Active
 RPM SW Version:            7.6.1.0
 Link to Peer:              Up

--  PEER RPM Status  --
-----------------------------------------------
 RPM State:                 Standby
 RPM SW Version:            7.6.1.0

--  RPM Redundancy Configuration  --
-----------------------------------------------
 Primary RPM:               rpm0
 Auto Data Sync:            Full
 Failover Type:             Hot Failover
 Auto reboot RPM:           Enabled
 Auto failover limit:       3 times in 60 minutes

--  RPM Failover Record  --
-----------------------------------------------
 Failover Count:            0
 Last failover timestamp:   None
 Last failover Reason:      None
 Last failover type:        None

--  Last Data Block Sync Record:  --
-----------------------------------------------
  Line Card Config:         succeeded  May 19 2008 11:34:06
   Start-up Config:         succeeded  May 19 2008 11:34:06
 Runtime Event Log:         succeeded  May 19 2008 11:34:06
    Running Config:         succeeded  May 19 2008 11:34:07
FTOS#
```

Note: A system may not start correctly if it boots from a corrupted FTOS image or an incorrect boot location. To troubleshoot a failed start, you can interrupt the boot process and configure a different boot location with the **boot change** or **boot system** commands. For more information, see Recovering from a Failed Start on page 77.

Communication between RPMs

E-Series RPMs have three CPUs: Control Processor (CP), Routing Processor 1 (RP1), and Routing Processor 2 (RP2). The CPUs use Fast Ethernet connections to communicate to each other and to the line card CPUs (LP) using Inter-Processor Communication (IPC). The CP monitors the health status of the other processors by sending a heartbeat message. If any CPU fails to acknowledge a consecutive number of heartbeat messages, or the CP itself fails to send heartbeat messages (IPC timeout), the primary RPM requests a failover to the standby RPM, and FTOS displays a message similar to Message 4.

C-Series RPMs have one CPU: Control Processor (CP). The CP on the RPM communicates with the LP via IPC. Like the E-Series, the CP monitors the health status of the other processors by sending a heartbeat message. If any CPU fails to acknowledge a consecutive number of heartbeat messages, or the CP itself fails to send heartbeat messages (IPC timeout), the primary RPM requests a failover to the standby RPM, and FTOS displays a message similar to Message 4.

Message 4 RPM Failover due to IPC Timeout

```
%RPM1-P:CP %IPC-2-STATUS: target rp2 not responding
%RPM0-S:CP %RAM-6-FAILOVER_REQ: RPM failover request from active peer: Auto failover on
failure
%RPM0-S:CP %RAM-6-ELECTION_ROLE: RPM0 is transitioning to Primary RPM.
%RPM0-P:CP %TSM-6-SFM_SWITCHFAB_STATE: Switch Fabric: UP
```

In addition to IPC, the CP on the each RPM sends heartbeat messages to the CP on its peer RPM via a process called Inter-RPM Communication (IRC). If the primary RPM fails to acknowledge a consecutive number of heartbeat messages (IRC timeout), the standby RPM responds by assuming the role of primary RPM, and FTOS displays message similar to message Message 5.

Message 5 RPM Failover due to IRC Timeout

```
20:29:07: %RPM1-S:CP %IRC-4-IRC_WARNLINKDN: Keepalive packet 7 to peer RPM is lost
20:29:07: %RPM1-S:CP %IRC-4-IRC_COMMDOWN: Link to peer RPM is down
%RPM1-S:CP %RAM-4-MISSING_HB: Heartbeat lost with peer RPM. Auto failover on heart beat lost.
%RPM1-S:CP %RAM-6-ELECTION_ROLE: RPM1 is transitioning to Primary RPM.
```

IPC and IRC timeouts and failover behavior

IPC or IRC timeouts can occur because heartbeat messages and acknowledgements are lost or arrive out of sequence, or a software or hardware failure occurs that impacts IPC or IRC. Table 18-2 describes the failover behavior for the possible failure scenarios.

Table 18-2. Failover Behaviors

Platform	Failover Trigger	Failover Behavior
C E	CP task crash on the primary RPM	The standby RPM detects the IRC time out and initiates failover, and the failed RPM reboots itself after saving a CP application core dump.
C E	CP IRC timeout for a non-task crash reason on the primary RPM	The standby RPM detects IRC time out and initiates failover. FTOS saves a CP trace log, the CP IPC-related system status, and a CP application core dump. Then the failed RPM reboots itself.

Table 18-2. Failover Behaviors

Platform	Failover Trigger	Failover Behavior
E	RP task or kernel crash on the primary RPM	CP on the primary RPM detects the RP IPC timeout and notifies the standby RPM. The standby RPM initiates a failover. FTOS saves an RP application or kernel core dump, the CP trace log, and the CP IPC-related system status. Then the new primary RPM reboots the failed RPM.
E	RP IPC timeout for a non-task crash reason on the primary RPM	CP on primary RPM detects the RP IPC timeout and notifies standby RPM. Standby RPM initiates a failover. FTOS saves an RP application core dump, RP IPC-related system status, a CP trace log record, and the CP IPC-related system status. Then the new primary RPM reboots the failed RPM.
C E	Hardware error detected on the primary RPM	FTOS detects the hardware error on the primary RPM and notifies the standby RPM. The standby RPM initiates a failover. FTOS saves a CP trace log, and a CP hardware nvtrace log. Then the new primary RPM reboots the failed RPM.
C E	Forced failover via the CLI	CP on primary RPM notifies standby RPM and the standby RPM initiates a failover. FTOS collects no system information. The former primary RPM immediately reboots after failover.
C E	Primary RPM is removed	The standby RPM detects the removal and initiates a failover. FTOS collects no system information.

After a failover, the new primary RPM prompts you for a username and password if authentication methods was configured and that data was synchronized. The standby RPM does not use authentication methods involving client/server protocols, such as RADIUS and TACACS+.

FTOS logs information about IPC timeouts in a log file that you can access. See:

- Chapter 60, C-Series Debugging and Diagnostics, C-Series Debugging and Diagnostics on page 1167
- Chapter 61, E-Series TeraScale Debugging and Diagnostics, Inter-CPU timeouts on page 1206

Support for RPM redundancy by FTOS version

FTOS supports increasing levels of RPM redundancy (warm and hot) as described in Table 18-3.

Table 18-3. Support for RPM Redundancy by FTOS Version

Failover Type	Failover Behavior	Platform
Warm Failover	The new primary RPM remains online, while the failed RPM, all line cards, and all SFMs reboot.	C E
Hot Failover	Only the failed RPM reboots. All line cards and SFMs remain online. All application tasks are spawned on the secondary RPM before failover. The running configuration is synchronized at runtime so it does not need to be reapplied during failover.	C E S

RPM synchronization

Data between the two RPMs is synchronized immediately after bootup. Once the two RPMs have done an initial full synchronization (block sync), thereafter FTOS only updates changed data (incremental sync). The data that is synchronized consists of configuration data, operational data, state and status, and statistics depending on the FTOS version.

Failover Type	Synchronized Data	Platform
Warm Failover	some NVRAM information, startup-configuration, line card configurations, user-access configurations	E C S
Hot Failover	some NVRAM information, startup-config, line card configurations, user-access configurations, running-config, SFM and datapath states, run-time event log and configuration, interface state	E C S

RPM redundancy configuration tasks

Select a Primary RPM

The RPM in slot 0 is the primary RPM by default. Manually select the primary RPM using the command **redundancy primary** from CONFIGURATION mode. View which RPM is the primary using the command show **running-config redundancy** from EXEC Privilege mode, as shown in Figure 18-2.

Figure 18-2. Selecting a Primary RPM

```
FTOS#show running-config redundancy
!
redundancy auto-failover-limit count 3 period 60
redundancy auto-synchronize full
redundancy primary rpm0
FTOS#
```

Force an RPM failover

Trigger an RPM failover between RPMs using the command **redundancy force-failover rpm** from EXEC Privilege mode. Use this feature when:

- you are replacing an RPM, and
- you are performing a warm upgrade

Figure 18-3. Using the redundancy force-failover rpm Command to Copy Software between RPMs

```
FTOS#redundancy force-failover rpm
Peer RPM's SW version is different but HA compatible.
Failover can be done by warm or hitless upgrade.
All linecards will be reset during warm upgrade.

Specify hitless upgrade or warm upgrade [confirm hitless/warm]:hitless
Proceed with warm upgrade [confirm yes/no]:
```

Specify an Auto-failover Limit

When a non-recoverable fatal error is detected, an automatic failover occurs. However, FTOS is configured to auto-failover only three times within any 60 minute period. You may specify a different auto-failover count and period using the command redundancy auto-failover-limit.

To re-enable the auto-failover-limit with its default parameters, in CONFIGURATION mode, use the **redundancy auto-failover-limit** command without parameters.

Disable Auto-reboot

Prevent a failed RPM from rebooting after a failover using the command **redundancy disable-auto-reboot** from CONFIGURATION mode.

Manually Synchronize RPMs

Manually synchronize RPMs at any time using the command **redundancy synchronize full** from EXEC Privilege mode.

Switch Fabric Module redundancy

Switch Fabric Module Redundancy is supported on platform: C

Since the RPM on the C-Series also contains the switch fabric, even though the second RPM comes online as the standby, the switch fabric is active and is automatically available for routing. Change this behavior using the command **redundancy sfm standby** from CONFIGURATION mode. To bring the secondary SFM online, enter **no redundancy sfm standby**. There is sub-second packet-loss anytime an SFM is brought online or taken offline. Use the command **show sfm all** to determine the status of the SFMs on the RPMs.

Online Insertion and Removal

You can add, replace, or remove chassis components while the chassis is operating.

This section contains the following sub-sections:

RPM Online Insertion and Removal

Dell Force10 systems are functional with only one RPM. If a second RPM is inserted, it comes online as the standby RPM, as shown in Figure 18-4.

On the C-Series, when a secondary RPM with a logical SFM is inserted or removed, the system must add or remove the backplane links to the switch fabric trunk. Any time such links are changed, traffic is disrupted. Use the command **redundancy sfm standby** to avoid any traffic disruption when the secondary RPM is inserted. When this command is executed, the logical SFM on the standby RPM is immediately taken offline, and the SFM state set as standby. Use the command **show sfm all** to see SFM status information.

Figure 18-4. Inserting a Second RPM into an Online System

```
FTOS#show rpm all

--  Route Processor Modules --
Slot  Status        NxtBoot    Version
-------------------------------------------------------------------------
  0   active        online     7-5-1-71
  1   not present
%RPM0-P:CP %POLLMGR-2-ALT_RPM_STATE: Alternate RPM is present
%RPM0-P:CP %IRC-6-IRC_COMMUP: Link to peer RPM is up
%RPM1-S:CP %RAM-5-RPM_STATE: RPM1 is in Standby State

FTOS#show rpm all

--  Route Processor Modules --
Slot  Status        NxtBoot    Version
-------------------------------------------------------------------------
  0   active        online     7-5-1-71
  1   standby       online     7-5-1-71
```

Line Card Online Insertion and Removal

FTOS detects the line card type when you insert a line card into a online chassis. FTOS writes the line card type to the running-config and maintains this information as a logical configuration if you remove the card (or the card fails), as shown in Figure 18-5.

Figure 18-5. Inserting and Removing a Line Card

```
FTOS(conf)#do show linecard all

--  Line cards  --
Slot  Status        NxtBoot    ReqTyp   CurTyp   Version     Ports
------------------------------------------------------------------------------
  0   not present
[output omitted]

FTOS(conf)# %RPM0-P:CP %CHMGR-5-CARDDETECTED: Line card 0 present
FTOS(conf)# do show linecard  all

--  Line cards  --
Slot  Status        NxtBoot    ReqTyp   CurTyp   Version     Ports
------------------------------------------------------------------------------
  0   online        online     E48VB    E48VB    7-5-1-71    48
[output omitted]
FTOS(conf)#%RPM0-P:CP %CHMGR-2-CARD_DOWN: Line card 0 down - card removed
FTOS(conf)#do show linecard  all

--  Line cards  --
Slot  Status        NxtBoot    ReqTyp   CurTyp   Version     Ports
------------------------------------------------------------------------------
  0   not present               E48VB
[output omitted]
```

Pre-configure a line card slot

You may also pre-configure an empty line card slot with a logical line card using the command **linecard** from CONFIGURATION mode. After creating the logical line card, you can configure the interfaces on the line card as if it is present, as shown in Figure 18-6.

Figure 18-6. Configuring a Logical Line Card

```
FTOS(conf)#do show linecard 0

--  Line card 0 --
Status        : not present

FTOS(conf)#int gig 0/0
                ^
% Error: No card configured in slot at "^" marker.
FTOS(conf)#linecard 0 E48VB
FTOS(conf)#do show linecard 0

--  Line card 0 --
Status        : not present
Required Type : E48VB - 48-port GE 10/100/1000Base-T line card with RJ45 interfaces and PoE
(CB)

FTOS(conf)#int gig 0/0
FTOS(conf-if-gi-0/0)#
```

Replace a line card

If you are replacing a line card with a line card of the same type, you may replace the card without any additional configuration.

If you are replacing a line card with a line card of a different type, remove the card and then remove the existing line card configuration using the command **no linecard**. If you do not, FTOS reports a card mismatch (Message 6) when you insert the new card, and the installed line card has a card mismatch status. To clear this line card mismatch status and bring the line card online, specify the type of line card you inserted using the command **linecard**, as shown in Figure 18-7.

Message 6 Line card Mismatch Error

```
%RPM0-P:CP %CHMGR-3-CARD_MISMATCH: Mismatch: line card 0 is type E48VB - type E48TB required
```

Figure 18-7. Resolving a Line Card Type Mismatch

```
%RPM0-P:CP %CHMGR-5-CARDDETECTED: Line card 0 present

%RPM0-P:CP %CHMGR-3-CARD_MISMATCH: Mismatch: line card 0 is type E48VB - type E48TB required

FTOS#show linecard  all

--  Line cards  --
Slot  Status        NxtBoot    ReqTyp   CurTyp   Version     Ports
------------------------------------------------------------------------
  0   type mismatch online     E48TB    E48VB    7-5-1-71    48
[output omitted]

FTOS(conf)#linecard 0 E48VB
Aug 6 14:25:22: %RPM0-P:CP %IFMGR-1-DEL_PORT: Removed port: Gi 0/0-47
FTOS(conf)#Aug 6 14:25:24: %RPM0-P:CP %CHMGR-5-LINECARDUP: Line card 0 is up

FTOS#show linecard  all

--  Line cards  --
Slot  Status        NxtBoot    ReqTyp   CurTyp   Version     Ports
------------------------------------------------------------------------
  0   online        online     E48VB    E48VB    7-5-1-71    48
[output omitted]
```

Hitless Behavior

Hitless Behavior is supported only on platform: C E

Hitless behavior is supported on E-Series ExaScale E X with FTOS 8.2.1.0. and later.

Hitless is a protocol-based system behavior that makes an RPM failover on the local system transparent to remote systems. The system synchronizes protocol information on the standby and primary RPMs such that, the event of an RPM failover, there is no need to notify remote systems of a local state change.

Hitless behavior is defined in the context of an RPM failover only and does not include line card, SFM, and power module failures.

- On the E-Series: Failovers triggered by software exception, hardware exception, forced failover via the CLI, and manual removal of the primary RPM are all hitless.
- On the C-Series: Only failovers via the CLI are hitless. The system is not hitless in any other scenario.

Hitless protocols are compatible with other hitless and graceful restart protocols. For example, if hitless OSPF is configured over hitless LACP LAGs, both features work seamlessly to deliver a hitless OSPF-LACP result. However, if hitless behavior involves multiple protocols, all must be hitless in order to achieve a hitless end result. For example, if OSPF is hitless but BFD is not, OSPF operates hitlessly and BFD flaps upon an RPM failover.

The following protocols are hitless:

- Link Aggregation Control Protocol. See Configure LACP as Hitless on page 549.
- Spanning Tree Protocol. See Configuring Spanning Trees as Hitless on page 1064.
- On the E-Series only, Bi-directional Forwarding Detection (line card ports). See Bidirectional Forwarding Detection on page 169.

Graceful Restart

Graceful Restart is supported on platform: E C S

Graceful restart (also called non-stop forwarding) is a protocol-based mechanism that preserves the forwarding table of the restarting router and its neighbors for a specified period to minimize the loss of packets. A graceful-restart router does not immediately assume that a neighbor is permanently down and so does not trigger a topology change. On E-Series, when you configure graceful restart, the system drops no packets during an RPM failover for protocol-relevant destinations in the forwarding table, and is therefore called “hitless”. On the C-Series and S-Series, packet loss is non-zero, but trivial, and so is still called hitless.

FTOS supports graceful restart for the following protocols:

- Border Gateway Protocol. See Enable graceful restart on page 241.
- Open Shortest Path First. Graceful Restart on page 700.
- Protocol Independent Multicast—Sparse Mode. PIM-SM Graceful Restart on page 764.
- Intermediate System to Intermediate System. Chapter 23, “Intermediate System to Intermediate System,” on page 507.

Software Resiliency

During normal operations FTOS monitors the health of both hardware and software components in the background to identify potential failures, even before these failures manifest.

Runtime System Health Check

Runtime System Health Check is supported on platform: E

FTOS runs a system health check to detect data transfer errors within the system. FTOS performs the check during normal operation by interspersing among, test frames among the data frames that carry user and system data. One such check is a data plane loopback test.

There are some differences between the TeraScale and ExaScale line card and RPM testing:

- The TeraScale card test contains a loopback from the RPM to the SFM and a loopback from the line cards to the SFM.
- The ExaScale card test contains a loopback from the RPM to the SFM and a loopback from the line cards to the on-board TSF3.
- For TeraScale, each line card and RPM periodically sends out test frames that loop back through the SFM. The loopback health check determines the overall status of the backplane and can identifies a faulty SFM. If three consecutive RPM loopbacks fail, then the software initiates a fault isolation procedure that sequentially disables one SFM at a time and performs the same loopback test.
- For ExaScale, the RPM alone RPM periodically sends out test frames that loop back through the SFM. The loopback health check determines the overall status of the backplane and can identifies a faulty SFM. If three consecutive RPM loopbacks fail, then the software initiates a fault isolation procedure that sequentially disables one SFM at a time and performs the same loopback test.

Refer to the Chapter 61, E-Series TeraScale Debugging and Diagnostics for details on the different system checks performed.

SFM Channel Monitoring

PCDFO is supported only on platform: E

Another test that is used to check the integrity of the data plane is a Per-channel De-skew FIFO Overflow (PCDFO). Each ingress and egress Buffer and Traffic Manager (BTM/FPTM) maintains nine channel connections to the SFM. The PCDFO test detects a faulty channel on an SFM, RPM, or line card by creating a test frame and striping it across all nine SFM channels between the eBTM/eFPTM and iBTM/iFPTM. The eBTM/eFPTM must receive each segment of striped data within a specified time to be considered to have proper temporal alignment. Small skews less than the specified time are tolerated because of buffering within the BTM/FPTM. If segments are not received within the specified time, the fault is not tolerated, and FTOS initiates additional tests to isolate the fault.

For more information on the PCDFO test, see Chapter 61, E-Series TeraScale Debugging and Diagnostics, Respond to PCDFO events on page 1205.

Note: The BTM applies to E-Series TeraScale, and the FPTM applies to the E-Series ExaScale.

Software Component Health Monitoring

On each of the line cards and the RPM, there are a number of software components. FTOS performs a periodic health check on each of these components by querying the status of a flag, which the corresponding component resets within a specified time.

If any health checks on the RPM fail, then the FTOS fails over to standby RPM. If any health checks on a line card fails, FTOS resets the card to bring it back to the correct state.

System Health Monitoring

FTOS also monitors the overall health of the system. Key parameters like CPU utilization, free memory, error counters (CRC failures, packet loss, etc.) are measured, and upon exceeding a threshold can be used to initiate recovery mechanism.

Failure and Event Logging

Dell Force10 systems provides multiple options for logging failures and events.

Trace Log

Developers interlace messages with software code to track a the execution of a program. These messages are called trace messages; they are primarily used for debugging and provide lower level information than event messages, which are primarily used by system administrators. FTOS retains executed trace messages for hardware and software and stores them in files (logs) on the internal flash.

- NV Trace Log—contains line card bootup trace messages that FTOS never overwrites, and is stored in internal flash under the directory NVTRACE_LOG_DIR.
- Trace Log—contains trace messages related to software and hardware events, state, and errors. Trace Logs are stored in internal flash under the directory TRACE_LOG_DIR.
- Crash Log—contains trace messages related to IPC and IRC timeouts and task crashes on line cards, and is stored under the directory CRASH_LOG_DIR.

For more information on trace logs and configuration options, see:

- Chapter 60, C-Series Debugging and Diagnostics
- Chapter 61, E-Series TeraScale Debugging and Diagnostics

Core Dumps

A core dump is the contents of RAM being used by a program at the time of a software exception and is used to identify the cause of the exception. There are two types of core dumps, application and kernel.

- The kernel is the central component of an operating system that manages system processors and memory allocation and makes these facilities available to applications. A kernel core dump is the contents of the memory in use by the kernel at the time of an exception.
- An application core dump is the contents of the memory allocated to a failed application at the time of an exception.

System Log

Event messages provide system administrators diagnostics and auditing information. FTOS sends event messages to the internal buffer, all terminal lines, the console, and optionally to a syslog server. For more information on event messages and configurable options, see Chapter 4, System Management.

Hot-lock Behavior

FTOS Hot-lock features allow you to append and delete their corresponding CAM entries dynamically without disrupting traffic. Existing entries are simply are shuffled to accommodate new entries.

FTOS offers the following Hot-lock features:

- **Hot-lock IP ACLs** (supported on E-Series, C-Series, and S-Series) allow you to append rules to and delete rules from an Access Control List that is already written to CAM. This behavior is enabled by default and is available for both standard and extended ACLs on ingress and egress. For information on configuring ACLs, see Chapter 8, “IP Access Control Lists (ACL), Prefix Lists, and Route-maps,” on page 133.
- **Hot-lock PBR** (supported on E-Series only) allows you to append rules to and delete rules from a redirect list that is already written to CAM without disrupting traffic. This behavior is enabled by default. For information on configuring Policy-based Routing, see Chapter 37, “Policy-based Routing,” on page 801.

Warm Upgrade

Warm Upgrade is supported on platform E

Warm software upgrades use warm failover, which means that FTOS reboots the secondary RPM and all line cards and SFMs. The chassis remains online during the upgrade, but forwarding is interrupted, as shown in Table 18-4.

FTOS supports warm software upgrades under two conditions:

- between consecutive feature releases where only the second digit differs between the running FTOS version number and the upgrade version number. For example, and upgrade from FTOS version 7.6.1.0 to 7.7.1.0 is warm.
- between two consecutive maintenance releases of the same feature release. For example, upgrading from 7.7.1.0 to 7.7.1.1 is warm.

Table 18-4 show the warm upgrade and downtime impact, if any, which each step.

Table 18-4. Control Plane and Data Plane Status during Warm Upgrade

	Download 6.3.1.1 to RPMs	Reboot RPM1 to Upgrade	Initiate Warm Failover	Reboot RPM0 to Upgrade
RPM 0	7.6.1.0 Primary	7.6.1.0 Primary	7.6.1.0 Secondary	7.7.1.0 Secondary
RPM 1	7.6.1.0 Secondary	7.7.1.0 Secondary	7.7.1.0 Primary	7.7.1.0 Primary
Line Cards	7.6.1.0	7.6.1.0	7.7.1.0	7.7.1.0
Control Plane	Operational	Operational	Interruption	Operational
Forwarding State	Forwarding	Forwarding	Interruption	Forwarding

Configure Cache Boot

Cache Boot is supported on platforms: C E

Cache Boot is supported on E-Series ExaScale E X with FTOS 8.2.1.0. and later.

FTOS Behavior: On E-Series ExaScale, the SFM auto upgrade feature is not supported with cacheboot. If you attempt an SFM auto upgrade, you must reload the chassis to recover.

The Dell Force10 system has the ability to boot the chassis using a cached FTOS image. FTOS stores the system image on the bootflash for each processor so that:

- the processors do not have to download the images during bootup, and
- the processors can boot in parallel rather than serially.

Booting the system by this method significantly reduces the time to bring the system online. Using Cache Boot with Warm Upgrade significantly reduces downtime during an upgrade to bring the system online during routine reloads.

Cache Boot can be configured during runtime. Dell Force10 recommends, however, that it be configured it when the system is offline.

The bootflash is partitioned so that two separate images can be cached, one for each RPM.

Cache Boot Pre-requisites

The system must meet two requirements before you can use the cache boot feature:

1. On the E-Series, the cache boot feature requires RPM hardware revision 2.1 or later. Use the **show rpm** command (Figure 18-8) to determine the version of your RPM. There is no hardware requirement for C-Series.

Figure 18-8. Determining your System Pre-requisites for Cache Boot

```
FTOS#show rpm

-- RPM card 0 --
Status        : active
Next Boot     : online
Card Type     : RPM - Route Processor Module (LC-EF3-RPM)
Hardware Rev  : 2.2i          <-- Hardware Revision 2.1 or later
Num Ports     : 1
Up Time       : 1 day, 4 hr, 25 min
Last Restart  : reset by user
FTOS Version  : 4.7.5.427
Jumbo Capable : yes
CP Boot Flash : A: 2.4.1.1   [booted]   B: 2.4.1.1   <-- Specified boot code version
RP1 Boot Flash: A: 2.4.1.1              B: 2.4.1.1   [booted]
RP2 Boot Flash: A: 2.4.1.1              B: 2.4.1.1   [booted]
CP Mem Size   : 536870912 bytes
RP1 Mem Size  : 1073741824 bytes
RP2 Mem Size  : 1073741824 bytes
MMC Mem Size  : 520962048 bytes
External MMC  : n/a
Temperature   : 32C
Power Status  : AC
Voltage       : ok
Serial Number : FX000017082
--More--
```

2. The cache boot feature requires *at least* the boot code versions in Table 18-5. Use **show rpm** and **show linecard** commands to verify that you have the proper version (Figure 18-8).

Table 18-5. Boot Code Requirements for Cache Boot

Component	Boot Code
E-Series TeraScale RPM	2.4.2.1
E-Series TeraScale Line Card	2.3.2.1
E-Series ExaScale RPM	2.5.0.3
E-Series ExaScale Line Card	2.9.0.5
C-Series RPM	2.7.1.1
C-Series Line Card	2.6.0.1

If you do not have the proper boot code version, the system displays a message similar to Message 7 when you attempt to select a cache boot image (see Select the Cache Boot Image). See *Upgrading the Boot Code* in the Release Notes for instructions on upgrading boot code.

Message 7 Boot Code Upgrade Required for Cache Boot Error

```
% Error: linecard 0 doesn't have cache boot aware bootCode.
```

Select the Cache Boot Image

Select the FTOS image that you want to cache using the command **upgrade system-image**, as shown in Figure 18-9. Dell Force10 recommends using the keyword **all** with this command to avoid any mis-matched configurations.

Note: The cache boot feature is not enabled by default; you must copy the running configuration to the startup configuration (**copy running-config startup-config**) after selecting a cache boot image in order to enable it.

Figure 18-9. Selecting a Cache Boot Image

```
FTOS#upgrade system-image all A flash://FTOS-EF-7.8.1.0.bin

Current cache boot information in the system:
=============================================

Type          A                   B
---------------------------------------------------------
CP            invalid             invalid
RP1           invalid[b][n]       invalid
RP2           invalid             invalid
linecard 0    invalid             invalid
linecard 1 is not present.
linecard 2 is not present.
linecard 3 is not present.
linecard 4    invalid             6.5.1.8
linecard 5 is not present.

        Note: [b] : booted    [n] : next boot
Upgrade cache boot image(4.7.5.427) for all cards [yes/no]: yes

cache boot image downloading in progress...
!!!!!!!!!!!!!!!!!!!!!

cache boot upgrade in progress. Please do NOT power off the card.
Note: Updating Flash Table of Contents...
Erasing TOC area...
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!

Upgrade result :
================

All cache boot image upgraded to 4.7.5.427
FTOS#
```

View your cache boot configuration using the command **show boot system all**, as shown in Figure 18-10.

Figure 18-10. Viewing the Cache Boot Configuration

```
FTOS#show boot system all

Current system image information in the system:
===============================================

Type          Boot Type     A                          B
-----------------------------------------------------------------
CP            DOWNLOAD BOOT 4.7.5.427                  invalid
RP1           DOWNLOAD BOOT 4.7.5.427                  invalid
RP2           DOWNLOAD BOOT 4.7.5.427                  invalid
linecard 0    DOWNLOAD BOOT 4.7.5.427                  invalid
linecard 1 is not present.
linecard 2 is not present.
linecard 3 is not present.
linecard 4    DOWNLOAD BOOT 4.7.5.427                  6.5.1.8
linecard 5 is not present.
FTOS#
```

If you attempt to cache a system image that does not support the cache boot feature, Message 8 appears.

Message 8 System Image does not Support Cache Boot Error

```
%% Error: Given image is not cache boot aware image.
```

Verify that the system is configured to boot with the selected cache boot image using the command **show bootvar** as shown in Figure 18-11.

Figure 18-11. Viewing the Selected Cache Boot Image

```
FTOS#copy running-config startup-config
File with same name already exist.
Proceed to copy the file [confirm yes/no]: yes
!
10496 bytes successfully copied
1d6h32m: %RPM0-P:CP %FILEMGR-5-FILESAVED: Copied running-config to startup-config in flash by default
R4_E300#show bootvar
PRIMARY IMAGE FILE =  system://4.7.5.427
SECONDARY IMAGE FILE =  flash://FTOS-EF-7.7.1.0.bin
DEFAULT IMAGE FILE =  flash://FTOS-EF-7.6.1.0.bin
LOCAL CONFIG FILE =  variable does not exist
PRIMARY HOST CONFIG FILE =  variable does not exist
SECONDARY HOST CONFIG FILE =  variable does not exist
PRIMARY NETWORK CONFIG FILE =  variable does not exist
SECONDARY NETWORK CONFIG FILE =  variable does not exist
CURRENT IMAGE FILE =  flash://FTOS-EF-7.7.1.0.bin
CURRENT CONFIG FILE 1 =  flash://startup-config
CURRENT CONFIG FILE 2 =  variable does not exist
CONFIG LOAD PREFERENCE =  local first
BOOT INTERFACE GATEWAY IP ADDRESS =  variable does not exist
FTOS#
```

In-Service Modular Hot-Fixes

In-Service Modular Hot-Fixes are supported on platforms: E

In-Service Modular Hot-Fixes provides a tool whereby you can install a patch while the system is on-line and running. This feature allows a patch to be added to a running FTOS process to obtain debugging information or to resolve a software issue in a deployed system.

There is no need to reload or reboot the system when the patch is inserted. The In-Service Modular patch replaces the existing process code. Once installation is complete, the system executes the patch code as though it was always there.

Add the patch with the **patch flash://RUNTIME_PATCH_DIR/<***patchname***>** command. Remove a patch and revert to the original code with the **no patch flash://RUNTIME_PATCH_DIR/<***patchname***>** command.

The patch name includes the FTOS version, the platform, the CPU, and the process it affects (FTOS-platform-cpu-process-patchversion.rtp). For example, a patch labeled **7.8.1.0-EH-rp2-l2mgr-1.rtp** identifies that this patch applies to FTOS version 7.8.1.0, on the E-Series platform, for RP2, addressing the layer 2 management process, and this is the first version of this patch.

Use the following to add a patch to the system.

Step	Task	Command Syntax	Command Mode
1	If not already done, copy the patch to the Runtime Patch directory	**copy** *file-origin* **rpm {0|1} flash://** RUNTIME_PATCH_DIR	EXEC Privilege
2	Verify that the patch resides in the system's internal flash. In-Service Modular patches are identified as **.rtp** files	**pwd** flash:/RUNTIME_PATCH_DIR FTOS#**dir** `Directory of flash:/RUNTIME_PATCH_DIR` `1 drwx 8192 Dec 06 2008 22:25:54 +00:00 .` `2 drwx 32768 Jan 01 1980 00:00:00 +00:00 ..` `3 -rwx 1835 Mar 03 2009 16:31:48 +00:00 7.8.0.1-EH-rp1-bgp-1.rtp` `flash: 520962048 bytes total (242778112 bytes free)`	EXEC
3	Insert the patch.	**patch flash://RUNTIME_PATCH_DIR/ <***patchname***>** For example: `patch flash://RUNTIME_PATCH_DIR/7.8.0.1-EH-rp1-bgp-1.rtp`	CONFIGURATION
4	View all In-Service patches on the system.	**show patch** `Patch version Module Cpu Timestamp` `7.8.0.1-EH-rp1-bgp-1.rtp bgp RP1 Tue Mar 3 16:10:08 PST 2009`	EXEC

Note: The **show patch** command can be used on both the primary and secondary RPMs, as shown here:

```
FTOS(standby)#show patch
Patch version            Module     Cpu  Timestamp
E.1.1.bgp.1.0            bgp        RP1  Mon Jun 22 06:51:23 PDT 2009
E.2.1.l2mgr.1.0          l2mgr      RP2  Mon Jun 22 07:11:15 PDT 2009
FTOS(standby)#
```

Process Restartability

Process Restartability is supported on platforms: C E S

Process Restartability is an extension to the FTOS High Availability system component that enables application processes and system protocol tasks to be restarted. This extension increases system reliability and uptime by first attempting to restart the crashed process on the primary RPM and then executing the failover procedure only as a last resort.

Currently, if a software exception occurs, FTOS executes a failover procedure:

- **In a single-RPM system**, the system generates a core dump and reboots. Reloads require the system to read and apply the entire startup-configuration file, which might take long if the startup-configuration is large. Restarting a process saves time because only a portion of the configuration related to the crashed process is read and re-applied.
- **In a dual-RPM system**, the system generates a core dump and fails over to the standby RPM. Restarting a process precludes launching the failover process. Recovery is attempted first locally on the primary RPM, which involves less CPU overhead, increasing system availability for other activities.

For both a single and dual-RPM system, when Process Restartability is enabled and a software exception occurs, FTOS: executes a core dump, frees system resources, restarts the failed process, and then updates the restart counter. By default, a process can be restarted a maximum of 3 times within 1 hour. If this limit is exceeded, the FTOS reloads the system reloads or fails over to the secondary RPM.

The processes that can be restarted are:

- **Management-related processes**—TACACS+, RADIUS, CLI, SSH, Telnet, Console/Aux
  - **TACACS+/RADIUS**—FTOS restarts the process and reapplies the TACACS+ portion of the running configuration. You must enable process restart explicitly.
  - **Console/Aux**—FTOS restarts the process; you must log in again after the restart. The threshold for failover is 3 restarts per hour; the fourth restart triggers a failover. The restart configuration is internal, and not user-configurable.
  - **Telnet/SSH**—Each SSH and Telnet session is an individual process. If a Telnet or SSH software exception occurs, only your session is cleared, and you must log in again; no other sessions are affected. This behavior is an exception among the other restartable processes in that Telnet and SSH are not literally restarted. Because of this, Telnet and SSH are not subject to the default or configured restart limit; that is, the system never reloads due to a software exception in these processes. However, a core dump is generated.

You can select which process may attempt to restart, and the number of consecutive restart attempts before failover, but by default, every process causes a system reload or RPM failover.

Task	Command Syntax	Command Mode
Enable Process Restartability for a process or task.	**process restartable** [*process*] [**count** *number*] [**period** *hours*]	CONFIGURATION

```
FTOS(conf)#process restartable radius ?
count                  Specify max number of auto-restarts (default=3)
<cr>
FTOS(conf)#process restartable radius count ?
<1-3>                  number of times value to auto-restarts
FTOS(conf)#process restartable radius count 3 ?
period                 Specify time span for auto-restart count (default=1)
<cr>
```

Display the processes and tasks configured for restartability.	**show process restartable** [**history**]	EXEC Privilege

```
FTOS#sho processes restartable

 ---------------------------------------------------------------------------------------------
  Process name          State           How many times restarted        Timestamp last restarted
 ---------------------------------------------------------------------------------------------
  radius                enabled                           0           [-]
  tacplus               enabled                           0           [-]
 ---------------------------------------------------------------------------------------------
FTOS#show processes restartable history

 ---------------------------------------------------------------------------------------------
  Process name          Timestamp last crashed
 ---------------------------------------------------------------------------------------------
  radius                 [5/23/2001 10:11:47]
 ---------------------------------------------------------------------------------------------
```

When a process restarts, FTOS displays Message 9.

Message 9 System Message for Process Restarts

```
May 8 06:28:35: %RPM0-P:CP %TME-2-PROC_RESTART: Restarting crashed process tacplus
```

You can specify the timestamp in hours so that if the number of restart attempts exceeds the configured limit within this time frame, no further process restarts are attempted. The next time a software exception occurs within this process, the system reloads for a single-RPM system and fails over to the standby RPM for a dual-RPM system.

When a system exceeds the configured restart threshold, FTOS displays Message 10.

Message 10 System Message for Excessive Process Restarts

```
[for Dual RPM:] May 8 06:35:33: %RPM0-P:CP %TME-2-No More Restarts and do failover: process
tacplus reaches max restart with threshold 3

[for Single RPM:] Mar 1 10:21:47: %RPM0-P:CP %TME-2-No More Restarts: process tacplus reaches
max restart with threshold 3
```

FTOS Behavior: When **debug tacacs** or **debug radius** is enabled, and the respective process restarts, FTOS does not continue to print debug messages after the restart; you must execute **debug tacacs** or **debug radius** again. This is because debugging is not saved to the running configuration, rather, FTOS marks the process for debugging with a flag that is cleared during the restart.

19

Internet Group Management Protocol

Table 19-1. FTOS Support for IGMP and IGMP Snooping

Feature	Platform
IGMP version 1, 2, and 3	C E S
IGMP Snooping version 1, 2, and 3	C E S

Note: When both E-Series TeraScale and ExaScale are supported, only the E symbol is shown. If a feature is supported by one or the other chassis, the specific symbols are shown: E T for E-Series TeraScale or E X for E-Series ExaScale.

Multicast is premised on identifying many hosts by a single destination IP address; hosts represented by the same IP address are a *multicast group*. Internet Group Management Protocol (IGMP) is a Layer 3 multicast protocol that hosts use to join or leave a multicast group. Multicast routing protocols (such as PIM) use the information in IGMP messages to discover which groups are active and to populate the multicast routing table.

IGMP Implementation Information

- FTOS supports IGMP versions 1, 2, and 3 based on RFCs 1112, 2236, and 3376, respectively.
- FTOS does not support IGMP version 3 and versions 1 or 2 on the same subnet.
- IGMP on FTOS supports up to 512 interfaces on E-Series, 31 interfaces on C-Series and S-Series, and an unlimited number of groups on all platforms.
- Dell Force10 systems cannot serve as an IGMP host or an IGMP version 1 IGMP Querier.
- FTOS automatically enables IGMP on interfaces on which you enable a multicast routing protocol.

IGMP Protocol Overview

IGMP has three versions. Version 3 obsoletes and is backwards-compatible with version 2; version 2 obsoletes version 1.

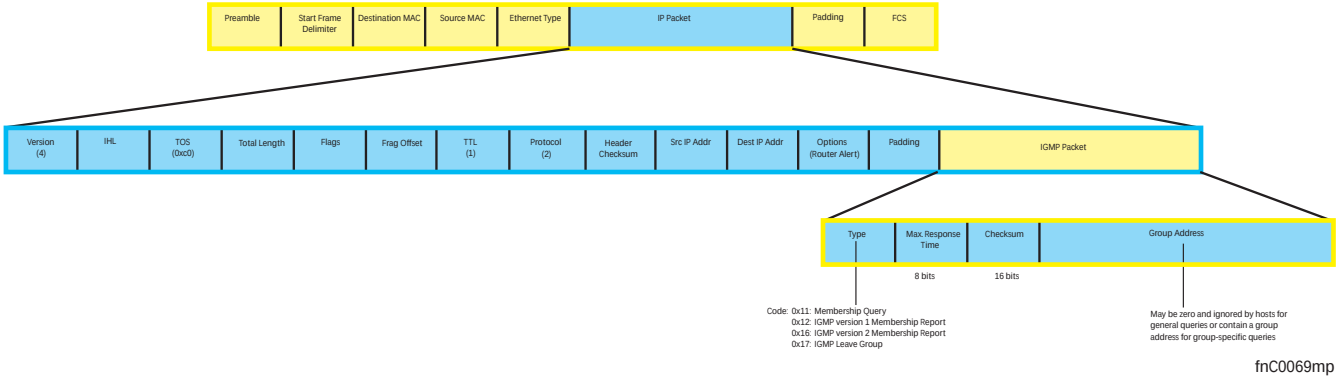
IGMP version 2

IGMP version 2 improves upon version 1 by specifying IGMP Leave messages, which allows hosts to notify routers that they no longer care about traffic for a particular group. Leave messages reduce the amount of time that the router takes to stop forwarding traffic for a group to a subnet (leave latency) after the last host leaves the group. In version 1 hosts quietly leave groups, and the router waits for a query response timer several times the value of the query interval to expire before it stops forwarding traffic.

To receive multicast traffic from a particular source, a host must join the multicast group to which the source is sending traffic. A host that is a member of a group is called a *receiver*. A host may join many groups, and may join or leave any group at any time. A host joins and leaves a multicast group by sending an IGMP message to its IGMP Querier. The querier is the router that surveys a subnet for multicast receivers, and processes survey responses to populate the multicast routing table.

IGMP messages are encapsulated in IP packets, as shown in Figure 19-1.

Figure 19-1. IGMP version 2 Packet Format



Joining a Multicast Group

There are two ways that a host may join a multicast group: it may respond to a general query from its querier, or it may send an unsolicited report to its querier.

Responding to an IGMP Query

1. One router on a subnet is elected as the querier. The querier periodically multicasts (to all-multicast-systems address 224.0.0.1) a general query to all hosts on the subnet.
2. A host that wants to join a multicast group responds with an IGMP Membership Report that contains the multicast address of the group it wants to join (the packet is addressed to the same group). If multiple hosts want to join the same multicast group, only the report from the first host to respond reaches the querier, and the remaining hosts suppress their responses (see Adjusting Query and Response Timers on page 410 for how the delay timer mechanism works).
3. The querier receives the report for a group and adds the group to the list of multicast groups associated with its outgoing port to the subnet. Multicast traffic for the group is then forwarded to that subnet.

Sending an Unsolicited IGMP Report

A host does not have to wait for a general query to join a group. It may send an unsolicited IGMP Membership Report, also called an IGMP Join message, to the querier.

Leaving a Multicast Group

1. A host sends a membership report of type 0x17 (IGMP Leave message) to the all routers multicast address 224.0.0.2 when it no longer cares about multicast traffic for a particular group.
2. The querier sends a Group-Specific Query to determine whether there are any remaining hosts in the group. There must be at least one receiver in a group on a subnet for a router to forward multicast traffic for that group to the subnet.
3. Any remaining hosts respond to the query according to the delay timer mechanism (see Adjusting Query and Response Timers on page 410). If no hosts respond (because there are none remaining in the group) the querier waits a specified period, and sends another query. If it still receives no response, the querier removes the group from the list associated with forwarding port and stops forwarding traffic for that group to the subnet.

IGMP version 3

Conceptually, IGMP version 3 behaves the same as version 2. There are differences:

- Version 3 adds the ability to filter by multicast source, which helps multicast routing protocols avoid forwarding traffic to subnets where there are no interested receivers.
- To enable filtering, routers must keep track of more state information, that is, the list of sources that must be filtered. An additional query type, the Group-and-Source-Specific Query, keeps track of state changes, while the Group-Specific and General queries still refresh existing state.
- Reporting is more efficient and robust: hosts do not suppress query responses (non-suppression helps track state and enables the immediate-leave and IGMP Snooping features), state-change reports are retransmitted to insure delivery, and a single membership report bundles multiple statements from a single host, rather than sending an individual packet for each statement.

The version 3 packet structure is different from version 2 to accommodate these protocol enhancements. Queries (Figure 19-2) are still sent to the all-systems address 224.0.0.1, but reports (Figure 19-3) are sent to the all IGMP version 3-capable multicast routers address 244.0.0.22.

Figure 19-2. IGMP version 3 Membership Query Packet Format

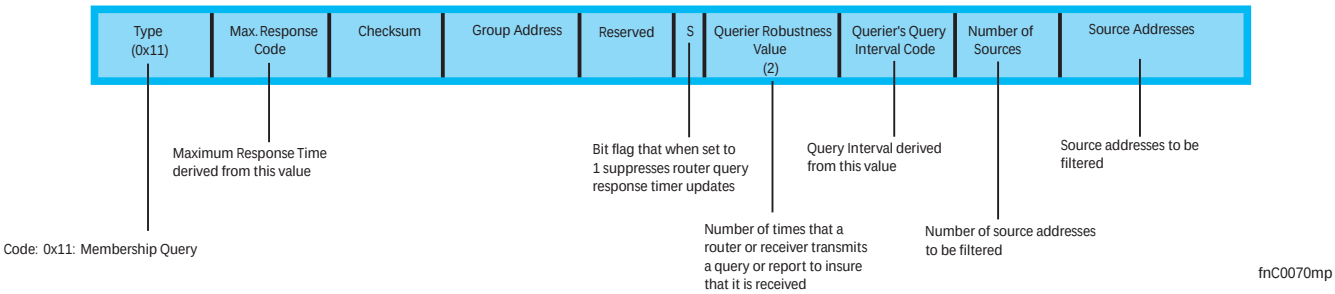
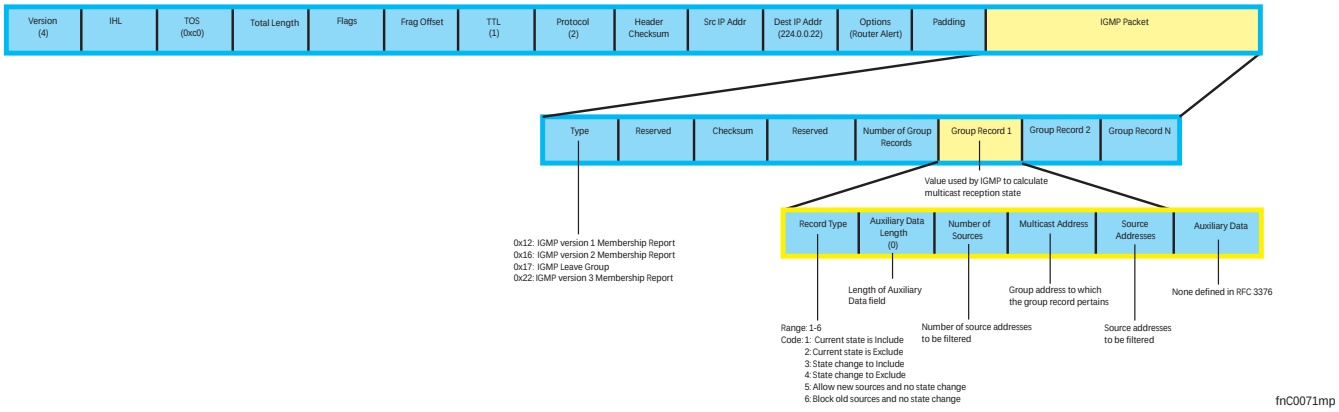


Figure 19-3. IGMP version 3 Membership Report Packet Format

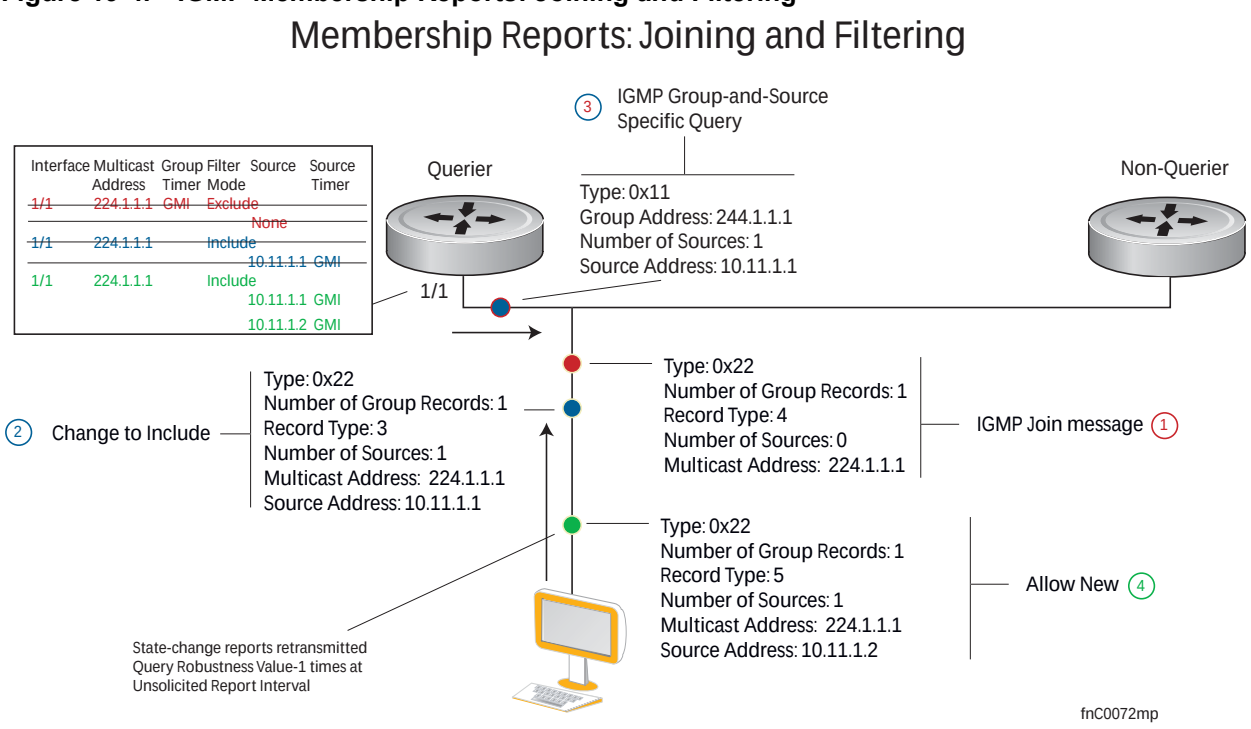


Joining and Filtering Groups and Sources

Figure 19-4 shows how multicast routers maintain the group and source information from unsolicited reports.

1. The first unsolicited report from the host indicates that it wants to receive traffic for group 224.1.1.1.
2. The host's second report indicates that it is only interested in traffic from group 224.1.1.1, source 10.11.1.1. Include messages prevent traffic from all other sources in the group from reaching the subnet, so before recording this request, the querier sends a group-and-source query to verify that there are no hosts interested in any other sources. The multicast router must satisfy all hosts if they have conflicting requests. For example, if another host on the subnet is interested in traffic from 10.11.1.3, then the router cannot record the include request. There are no other interested hosts, so the request is recorded. At this point, the multicast routing protocol prunes the tree to all but the specified sources.
3. The host's third message indicates that it is only interested in traffic from sources 10.11.1.1 and 10.11.1.2. Since this request again prevents all other sources from reaching the subnet, the router sends another group-and-source query so that it can satisfy all other hosts. There are no other interested hosts so the request is recorded.

Figure 19-4. IGMP Membership Reports: Joining and Filtering

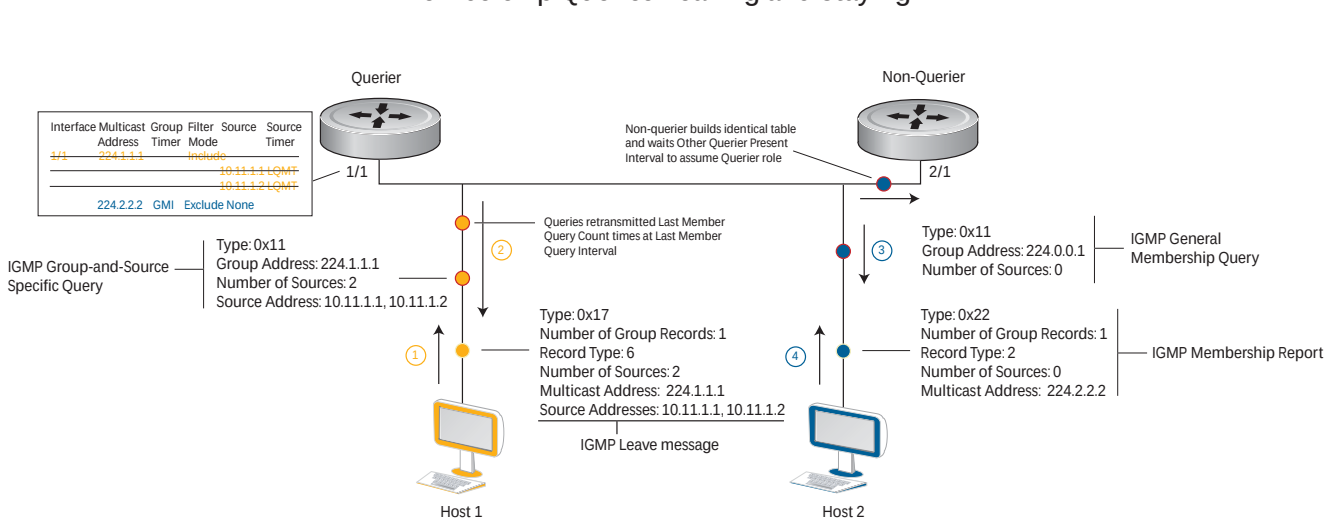


Leaving and Staying in Groups

Figure 19-5 shows how multicast routers track and refresh state changes in response to group-and-specific and general queries.

1. Host 1 sends a message indicating it is leaving group 224.1.1.1 and that the include filter for 10.11.1.1 and 10.11.1.2 are no longer necessary.
2. The querier, before making any state changes, sends a group-and-source query to see if any other host is interested in these two sources; queries for state-changes are retransmitted multiple times. If any are, they respond with their current state information and the querier refreshes the relevant state information.
3. Separately in Figure 19-5, the querier sends a general query to 224.0.0.1.
4. Host 2 responds to the periodic general query so the querier refreshes the state information for that group.

Figure 19-5. IGMP Membership Queries: Leaving and Staying in Groups



Configuring IGMP

Configuring IGMP is a two-step process:

1. Enable multicast routing using the command **ip multicast-routing**.
2. Enable a multicast routing protocol.

Related Configuration Tasks

- Viewing IGMP Enabled Interfaces on page 408
- Selecting an IGMP Version on page 409
- Viewing IGMP Groups on page 409
- Adjusting Timers on page 410
- Configuring a Static IGMP Group on page 411
- Prevent a Host from Joining a Group on page 667
- Enabling IGMP Immediate-leave on page 411
- IGMP Snooping on page 412
- Fast Convergence after MSTP Topology Changes on page 414
- Designating a Multicast Router Interface on page 414

Viewing IGMP Enabled Interfaces

Interfaces that are enabled with PIM-SM are automatically enabled with IGMP. View IGMP-enabled interfaces using the command **show ip igmp interface** command in the EXEC Privilege mode.

Figure 19-6. Viewing IGMP-enabled Interfaces

```
FTOS#show ip igmp interface gig 7/16
GigabitEthernet 7/16 is up, line protocol is up
  Internet address is 10.87.3.2/24
  IGMP is enabled on interface
  IGMP query interval is 60 seconds
  IGMP querier timeout is 300 seconds
  IGMP max query response time is 10 seconds
  Last member query response interval is 199 ms
  IGMP activity: 0 joins, 0 leaves
  IGMP querying router is 10.87.3.2 (this system)
  IGMP version is 2
FTOS#
```

Selecting an IGMP Version

FTOS enables IGMP version 2 by default, which supports version 1 and 2 hosts, but is not compatible with version 3 on the same subnet. If hosts require IGMP version 3, you can switch to IGMP version 3 using the command **ip igmp version** from INTERFACE mode, as shown in Figure 19-7.

Figure 19-7. Selecting an IGMP Version

```
FTOS(conf-if-gi-1/13)#ip igmp version 3
FTOS(conf-if-gi-1/13)#do show ip igmp interface
GigabitEthernet 1/13 is up, line protocol is down
  Inbound IGMP access group is not set
  Interface IGMP group join rate limit is not set
  Internet address is 1.1.1.1/24
  IGMP is enabled on interface
  IGMP query interval is 60 seconds
  IGMP querier timeout is 125 seconds
  IGMP max query response time is 10 seconds
  IGMP last member query response interval is 1000 ms
  IGMP immediate-leave is disabled
  IGMP activity: 0 joins, 0 leaves, 0 channel joins, 0 channel leaves
  IGMP querying router is 1.1.1.1 (this system)
  IGMP version is 3
FTOS(conf-if-gi-1/13)#
```

Viewing IGMP Groups

View both learned and statically configured IGMP groups using the command **show ip igmp groups** from EXEC Privilege mode.

Figure 19-8. Viewing Static and Learned IGMP Groups

```
FTOS(conf-if-gi-1/0)#do sho ip igmp groups
Total Number of Groups: 2
IGMP Connected Group Membership
Group Address    Interface                Uptime    Expires   Last Reporter
224.1.1.1        GigabitEthernet 1/0      00:00:03  Never     CLI
224.1.2.1        GigabitEthernet 1/0      00:56:55  00:01:22  1.1.1.2
```

Adjusting Timers

View the current value of all IGMP timers using the command **show ip igmp interface** from EXEC Privilege mode, as shown in Figure 19-6.

Adjusting Query and Response Timers

The querier periodically sends a general query to discover which multicast groups are active. A group must have at least one host to be active. When a host receives a query, it does not respond immediately, but rather starts a delay timer. The delay time is set to a random value between 0 and the Maximum Response Time. The host sends a response when the timer expires; in version 2, if another host responds before the timer expires, the timer is nullified, and no response is sent.

The Maximum Response Time is the amount of time that the querier waits for a response to a query before taking further action. The querier advertises this value in the query (see Figure 19-1). Lowering this value decreases leave latency but increases response burstiness since all host membership reports must be sent before the Maximum Response Time expires. Inversely, increasing this value decreases burstiness at the expense of leave latency.

- Adjust the period between queries using the command **ip igmp query-interval** from INTERFACE mode.
- Adjust the Maximum Response Time using the command **ip igmp query-max-resp-time** from INTERFACE mode.

When the querier receives a leave message from a host, it sends a group-specific query to the subnet. If no response is received, it sends another. The amount of time that the querier waits to receive a response to the initial query before sending a second one is the Last Member Query Interval (LMQI). The switch waits one LMQI after the second query before removing the group from the state table.

- Adjust the Last Member Query Interval using the command **ip igmp last-member-query-interval** from INTERFACE mode.

Adjusting the IGMP Querier Timeout Value

If there is more than one multicast router on a subnet, only one is elected to be the querier, which is the router that sends queries to the subnet.

1. Routers send queries to the all multicast systems address, 224.0.0.1. Initially, all routers send queries.

2. When a router receives a query it compares the IP address of the interface on which it was received with the source IP address given in the query. If the receiving router IP address is greater than the source address given in the query, the router stops sending queries. By this method, the router with the lowest IP address on the subnet is elected querier and continues to send queries.
3. If a specified amount of time elapses during which other routers on the subnet do not receive a query, those routers assume that the querier is down, and a new querier is elected.

The amount of time that elapses before routers on a subnet assume that the querier is down is the Other Querier Present Interval. Adjust this value using the command **ip igmp querier-timeout** from INTERFACE mode.

Note: The timeout value for an IGMP querier is calculated differently on Dell Force10 FTOS routers than on Cisco IOS routers. As a result, if the IGMP querier in a subnet goes down, a Cisco IOS router may be elected as the new querier before a Dell Force10 FTOS router.

Configuring a Static IGMP Group

Configure a static IGMP group using the command **ip igmp static-group**. Multicast traffic for static groups is always forwarded to the subnet even if there are no members in the group.

View the static groups using the command **show ip igmp groups** from EXEC Privilege mode. Static groups have an expiration value of *Never* and a Last Reporter value of *CLI*, as shown in Figure 19-8.

Enabling IGMP Immediate-leave

If the querier does not receive a response to a group-specific or group-and-source query, it sends another (Querier Robustness Value). Then, after no response, it removes the group from the outgoing interface for the subnet.

IGMP Immediate Leave reduces leave latency by enabling a router to immediately delete the group membership on an interface upon receiving a Leave message (it does not send any group-specific or group-and-source queries before deleting the entry). Configure the system for IGMP Immediate Leave using the command **ip igmp immediate-leave**.

View the enable status of this feature using the command **show ip igmp interface** from EXEC Privilege mode, as shown in Figure 19-7.

IGMP Snooping

Multicast packets are addressed with multicast MAC addresses, which represent a group of devices, rather than one unique device. Switches forward multicast frames out of all ports in a VLAN by default, even though there may be only some interested hosts, which is a waste of bandwidth. IGMP Snooping enables switches to use information in IGMP packets to generate a forwarding table that associates ports with multicast groups so that when they receive multicast frames, they can forward them only to interested receivers.

IGMP Snooping Implementation Information

- IGMP Snooping on FTOS uses IP multicast addresses not MAC addresses.
- IGMP Snooping is not supported on stacked VLANs.
- IGMP Snooping is supported on all S-Series stack members.
- IGMP Snooping reacts to STP and MSTP topology changes by sending a general query on the interface that transitions to the forwarding state.

Configuring IGMP Snooping

Configuring IGMP Snooping is a one-step process. That is, enable it on a switch using the command **ip igmp snooping enable** from CONFIGURATION mode. View the configuration using the command **show running-config** from CONFIGURATION mode, as shown in Figure 19-9. You can disable snooping on for a VLAN using the command **no ip igmp snooping** from INTERFACE VLAN mode.

Figure 19-9. Enabling IGMP Snooping

```
FTOS(conf)#ip igmp snooping enable
FTOS(conf)#do show running-config igmp
ip igmp snooping enable
FTOS(conf)#
```

Related Configuration Tasks

- Enabling IGMP Immediate-leave on page 412
- Disabling Multicast Flooding on page 413
- Specifying a Port as Connected to a Multicast Router on page 413
- Configuring the Switch as Querier on page 413

Enabling IGMP Immediate-leave

Configure the switch to remove a group-port association upon receiving an IGMP Leave message using the command **ip igmp fast-leave** from INTERFACE VLAN mode. View the configuration using the command **show config** from INTERFACE VLAN mode, as shown in Figure 19-10.

Figure 19-10. Enabling IGMP Snooping

```
FTOS(conf-if-vl-100)#show config
!
interface Vlan 100
 no ip address
 ip igmp snooping fast-leave
 shutdown
FTOS(conf-if-vl-100)#
```

Disabling Multicast Flooding

If the switch receives a multicast packet that has an IP address of a group it has not learned (unregistered frame), the switch floods that packet out of all ports on the VLAN.

On the E-Series, you can configure the switch to only forward unregistered packets to ports on a VLAN that are connected to multicast routers (mrouter ports) using the command **no ip igmp snooping flood** from CONFIGURATION mode. When flooding is disabled, if there are no such ports in the VLAN connected to a multicast router, the switch drops the packets.

On the C-Series and S-Series, when you configure **no ip igmp snooping flood**, the system drops the packets immediately. The system does not forward the frames on mrouter ports, even if they are present. On the C-Series and S-Series, Layer 3 multicast must be disabled (**no ip multicast-routing**) in order to disable multicast flooding.

Specifying a Port as Connected to a Multicast Router

You can statically specify a port in a VLAN as connected to a multicast router using the command **ip igmp snooping mrouter** from INTERFACE VLAN mode.

View the ports that are connected to multicast routers using the command **show ip igmp snooping mrouter** from EXEC Privilege mode.

Configuring the Switch as Querier

Hosts that do not support unsolicited reporting wait for a general query before sending a membership report. When the multicast source and receivers are in the same VLAN, multicast traffic is not routed, and so there is no querier. You must configure the switch to be the querier for a VLAN so that hosts send membership reports, and the switch can generate a forwarding table by snooping.

Configure the switch to be the querier for a VLAN by first assigning an IP address to the VLAN interface, and then using the command **ip igmp snooping querier** from INTERFACE VLAN mode.

- IGMP snooping Querier does not start if there is a statically configured multicast router interface in the VLAN.
- The switch may lose the querier election if it does not have the lowest IP address of all potential queriers on the subnet.

- When enabled, IGMP snooping Querier starts after one query interval in case no IGMP general query (with IP SA lower than its VLAN IP address) is received on any of its VLAN members.

Adjusting the Last Member Query Interval

When the querier receives a leave message from a receiver, it sends a group-specific query out of the ports specified in the forwarding table. If no response is received, it sends another. The amount of time that the querier waits to receive a response to the initial query before sending a second one is the Last Member Query Interval (LMQI). The switch waits one LMQI after the second query before removing the group-port entry from the forwarding table.

Adjust the Last Member Query Interval using the command **ip igmp snooping last-member-query-interval** from INTERFACE VLAN mode.

Fast Convergence after MSTP Topology Changes

When a port transitions to the Forwarding state as a result of an STP or MSTP topology change, FTOS sends a general query out of all ports except the multicast router ports. The host sends a response to the general query and the forwarding database is updated without having to wait for the query interval to expire.

When an IGMP snooping switch is not acting as a Querier it sends out the general query, in response to the MSTP triggered link-layer topology change, with the source IP address of 0.0.0.0 to avoid triggering Querier election.

Designating a Multicast Router Interface

You can designate an interface as a multicast router interface with the command **ip igmp snooping mrouter interface**. FTOS also has the capability of listening in on the incoming IGMP General Queries and designate those interfaces as the multicast router interface when the frames have a non-zero IP source address. All IGMP control packets and IP multicast data traffic originating from receivers is forwarded to multicast router interfaces.

20

Interfaces

This chapter describes interface types, both physical and logical, and how to configure them with FTOS.

10/100/1000 Mbps Ethernet, Gigabit Ethernet, and 10 Gigabit Ethernet interfaces are supported on platforms C E S

SONET interfaces are only supported on platform E and are covered in the SONET/SDH chapter.

Basic Interface Configuration:

- Interface Types
- View Basic Interface Information
- Enable a Physical Interface
- Physical Interfaces
- Management Interfaces
- VLAN Interfaces
- Loopback Interfaces
- Null Interfaces on page 427
- Port Channel Interfaces

Advanced Interface Configuration:

- Bulk Configuration
- Interface Range Macros on page 443
- Monitor and Maintain Interfaces
- Link Debounce Timer
- Link Dampening
- Ethernet Pause Frames
- Configure MTU Size on an Interface
- Port-pipes on page 454
- Auto-Negotiation on Ethernet Interfaces
- View Advanced Interface Information

Interface Types

Interface Type	Modes Possible	Default Mode	Requires Creation	Default State
Physical	L2, L3	Unset	No	Shutdown (disabled)
Management	N/A	N/A	No	No Shutdown (enabled)
Loopback	L3	L3	Yes	No Shutdown (enabled)
Null	N/A	N/A	No	Enabled
Port Channel	L2, L3	L3	Yes	Shutdown (disabled)
VLAN	L2, L3	L2	Yes (except default)	L2 - No Shutdown (enabled) L3 - Shutdown (disabled)

View Basic Interface Information

The user has several options for viewing interface status and configuration parameters. The **show interfaces** command in EXEC mode will list all configurable interfaces on the chassis and has options to display the interface status, IP and MAC addresses, and multiple counters for the amount and type of traffic passing through the interface. If a port channel interface is configured, the **show interfaces** command can list the interfaces configured in the port channel.

Note: To end output from the system, such as the output from the **show interfaces** command, enter CTRL+C and FTOS will return to the command prompt.

Note: Unicast counters in the **show interface** output will increment when the interface receives multicast or broadcast packets.

Figure 20-1 displays the configuration and status information for one interface.

Figure 20-1. show interfaces Command Example

```
FTOS#show interfaces tengigabitethernet 1/0
TenGigabitEthernet 1/0 is up, line protocol is up
Hardware is Force10Eth, address is 00:01:e8:05:f3:6a
    Current address is 00:01:e8:05:f3:6a
Pluggable media present, XFP type is 10GBASE-LR.
    Medium is MultiRate, Wavelength is 1310nm
    XFP receive power reading is -3.7685
Interface index is 67436603
Internet address is 65.113.24.238/28
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 10000 Mbit, Mode full duplex, Master
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:09:54
Queueing strategy: fifo
Input Statistics:
     0 packets, 0 bytes
     0 Vlans
     0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     0 Multicasts, 0 Broadcasts
```

Use the **show ip interfaces brief** command in the EXEC Privilege mode to view which interfaces are enabled for Layer 3 data transmission. In Figure 20-2, GigabitEthernet interface 1/5 is in Layer 3 mode since an IP address has been assigned to it and the interface's status is operationally up.

Figure 20-2. show ip interfaces brief Command Example (Partial)

```
FTOS#show ip interface brief
Interface              IP-Address      OK? Method Status                Protocol
GigabitEthernet 1/0    unassigned      NO  Manual administratively down down
GigabitEthernet 1/1    unassigned      NO  Manual administratively down down
GigabitEthernet 1/2    unassigned      YES Manual up                    up
GigabitEthernet 1/3    unassigned      YES Manual up                    up
GigabitEthernet 1/4    unassigned      YES Manual up                    up
GigabitEthernet 1/5    10.10.10.1      YES Manual up                    up
GigabitEthernet 1/6    unassigned      NO  Manual administratively down down
GigabitEthernet 1/7    unassigned      NO  Manual administratively down down
GigabitEthernet 1/8    unassigned      NO  Manual administratively down down
```

Use the **show interfaces configured** command in the EXEC Privilege mode to view only configured interfaces. In Figure 20-2, GigabitEthernet interface 1/5 is in Layer 3 mode since an IP address has been assigned to it and the interface's status is operationally up.

To determine which physical interfaces are available, use the **show running-config** command in EXEC mode. This command displays all physical interfaces available on the line cards. (Figure 158).

Figure 20-3. Interfaces listed in the show running-config Command (Partial)

```
FTOS#show running
Current Configuration ...
!
interface GigabitEthernet 9/6
 no ip address
 shutdown
!
interface GigabitEthernet 9/7
 no ip address
 shutdown
!
interface GigabitEthernet 9/8
 no ip address
 shutdown
!
interface GigabitEthernet 9/9
 no ip address
 shutdown
```

Enable a Physical Interface

After determining the type of physical interfaces available, the user may enter the INTERFACE mode by entering the command **interface** *interface slot/port* to enable and configure the interfaces.

To enter the INTERFACE mode, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**interface** *interface*	CONFIGURATION	Enter the keyword **interface** followed by the type of interface and slot/port information: • For a 10/100/1000 Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For the Management interface on the RPM, enter the keyword **ManagementEthernet** followed by the slot/port information. • For a SONET interface, enter the keyword **sonet** followed by slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information.
2	**no shutdown**	INTERFACE	Enter the **no shutdown** command to enable the interface. If the interface is a SONET interface, enter the **encap ppp** command to enable PPP encapsulation.

To confirm that the interface is enabled, use the **show config** command in the INTERFACE mode.

To leave the INTERFACE mode, use the **exit** command or **end** command.

The user can not delete a physical interface.

Physical Interfaces

The *Management Ethernet interface*, is a single RJ-45 Fast Ethernet port on the Route Processor Module (RPM) of the C-Series and E-Series, and provides dedicated management access to the system. The S-Series systems supported by FTOS do not have this dedicated management interface, but you can use any Ethernet port configured with an IP address and route.

Line card interfaces support Layer 2 and Layer 3 traffic over the 10/100/1000, Gigabit, and 10-Gigabit Ethernet interfaces. SONET interfaces with PPP encapsulation support Layer 3 traffic. These interfaces (except SONET interfaces with PPP encapsulation) can also become part of virtual interfaces such as VLANs or port channels.

Link detection on ExaScale line cards is interrupt-based rather than poll-based, which enables ExaScale cards to bring up and take down links faster.

For more information on VLANs, see Bulk Configuration on page 440 and for more information on port channels, see Port Channel Interfaces on page 428.

FTOS Behavior: S-Series systems use a single MAC address for all physical interfaces while E-Series and C-Series use a unique MAC address for each physical interface, though this results in no functional difference between these platforms.

Configuration Task List for Physical Interfaces

By default, all interfaces are operationally disabled and traffic will not pass through them.

The following section includes information about optional configurations for physical interfaces:

- Overview of Layer Modes on page 420
- Configure Layer 2 (Data Link) Mode on page 420
- Management Interfaces on page 423
- Auto-Negotiation on Ethernet Interfaces on page 455
- Adjust the keepalive timer on page 457
- Clear interface counters on page 461

Overview of Layer Modes

On all systems running FTOS, you can place physical interfaces, port channels, and VLANs in Layer 2 mode or Layer 3 mode.

By default, VLANs are in Layer 2 mode.

Table 20-1. Interfaces Types

Type of Interface	Possible Modes	Requires Creation	Default State
10/100/1000 Ethernet, Gigabit Ethernet, 10 Gigabit Ethernet	Layer 2 Layer 3	No	Shutdown (disabled)
SONET (PPP encapsulation)	Layer 3	No	Shutdown (disabled)
Management	n/a	No	Shutdown (disabled)
Loopback	Layer 3	Yes	No shutdown (enabled)
Null interface	n/a	No	Enabled
Port Channel	Layer 2 Layer 3	Yes	Shutdown (disabled)
VLAN	Layer 2 Layer 3	Yes, except for the default VLAN	No shutdown (active for Layer 2) Shutdown (disabled for Layer 3)

Configure Layer 2 (Data Link) Mode

Use the **switchport** command in INTERFACE mode to enable Layer 2 data transmissions through an individual interface. The user can not configure switching or Layer 2 protocols such as spanning tree protocol on an interface unless the interface has been set to Layer 2 mode.

Figure 20-4 displays the basic configuration found in a Layer 2 interface.

Figure 20-4. show config Command Example of a Layer 2 Interface

```
FTOS(conf-if)#show config
!
interface Port-channel 1
 no ip address
 switchport
 no shutdown
FTOS(conf-if)#
```

To configure an interface in Layer 2 mode, use these commands in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
no shutdown	INTERFACE	Enable the interface.
switchport	INTERFACE	Place the interface in Layer 2 (switching) mode.

For information on enabling and configuring Spanning Tree Protocol, see Chapter 10, Layer 2, on page 47. To view the interfaces in Layer 2 mode, use the command **show interfaces switchport** in the EXEC mode.

Configure Layer 3 (Network) Mode

When you assign an IP address to a physical interface, you place it in Layer 3 mode. Use the **ip address** command and **no shutdown** command in INTERFACE mode to enable Layer 3 mode on an individual interface. In all interface types except VLANs, the **shutdown** command prevents all traffic from passing through the interface. In VLANs, the **shutdown** command prevents Layer 3 traffic from passing through the interface. Layer 2 traffic is unaffected by the **shutdown** command. One of the interfaces in the system must be in Layer 3 mode before you configure or enter a Layer 3 protocol mode (for example, OSPF).

Figure 20-5 shows how the **show config** command displays an example of a Layer 3 interface.

Figure 20-5. show config Command Example of a Layer 3 Interface

```
FTOS(conf-if)#show config
!
interface GigabitEthernet 1/5
 ip address 10.10.10.1 /24
 no shutdown
FTOS(conf-if)#
```

If an interface is in the incorrect layer mode for a given command, an error message is displayed to the user. For example, in Figure 20-6, the command **ip address** triggered an error message because the interface is in Layer 2 mode and the **ip address** command is a Layer 3 command only.

Figure 20-6. Error Message When Trying to Add an IP Address to Layer 2 Interface

```
FTOS(conf-if)#show config
!
interface GigabitEthernet 1/2
 no ip address
 switchport
 no shutdown
FTOS(conf-if)#ip address 10.10.1.1 /24
% Error: Port is in Layer 2 mode Gi 1/2.
FTOS(conf-if)#
```

Error message

To determine the configuration of an interface, you can use the **show config** command in INTERFACE mode or the various **show interface** commands in EXEC mode.

To assign an IP address, use both of the following commands in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
no shutdown	INTERFACE	Enable the interface.

Command Syntax	Command Mode	Purpose
ip address *ip-address mask* [**secondary**]	INTERFACE	Configure a primary IP address and mask on the interface. The *ip-address* must be in dotted-decimal format (A.B.C.D) and the *mask* must be in slash format (/xx). Add the keyword **secondary** if the IP address is the interface's backup IP address.

You can only configure one (1) primary IP address per interface. You can configure up to 255 secondary IP addresses on a single interface.

To view all interfaces to see with an IP address assigned, use the **show ip interfaces brief** command in the EXEC mode (Figure 176).

To view IP information on an interface in Layer 3 mode, use the **show ip interface** command in the EXEC Privilege mode (Figure 159).

Figure 20-7. Command Example: show ip interface

```
FTOS>show ip int vlan 58
Vlan 58 is up, line protocol is up
Internet address is 1.1.49.1/24
Broadcast address is 1.1.49.255
Address determined by config file
MTU is 1554 bytes
Inbound  access list is not set
Proxy ARP is enabled
Split Horizon is enabled
Poison Reverse is disabled
ICMP redirects are not sent
ICMP unreachables are not sent
```

Management Interfaces

Configure Management Interfaces on the E-Series and C-Series

On the E-Series and C-Series, the dedicated Management interface is located on the RPM and provides management access to the system. You can configure this interface with FTOS, but the configuration options on this interface are limited. Gateway addresses and IP addresses cannot be configured if it appears in the main routing table of FTOS. In addition, Proxy ARP is not supported on this interface.

The management interface supports IPv4 and IPv6 addressing, and supports ping and traceroute for IPv4 and IPv6 addresses.

To configure a Management interface, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
interface Managementethernet *interface*	CONFIGURATION	Enter the slot (0-1) and the port (0). In a system with 2 RPMs, therefore, 2 Management interfaces, the slot number differentiates between the two Management interfaces.

To configure IP addresses on a Management interface, use the following command in the MANAGEMENT INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip address *ipv4-address/mask* **ipv6 address** *ipv6-address/mask*	MANAGEMENT INTERFACE	Configure an IPv4 address (A.B.C.D) or IPv6 address (X:X:X:X::X) and mask (/x) on the interface.

If there are two RPMs on the system, each Management interface must be configured with a different IP address. Unless the **management route** command is configured, you can only access the Management interface from the local LAN. To access the Management interface from another LAN, the **management route** command must be configured to point to the Management interface.

Alternatively, you can use **virtual-ip** to manage a system with one or two RPMs. A virtual IP is an IP address assigned to the system (not to any management interfaces) and is a CONFIGURATION mode command. You may enter an IPv4 or IPv6 address. When a virtual IP address is assigned to the system, the active management interface of the RPM is recognized by the virtual IP address—not by the actual interface IP address assigned to it. During an RPM failover, you do not have to remember the IP address of the new RPM's management interface—the system will still recognizes the virtual-IP address.

Important Things to Remember — virtual-ip

- **virtual-ip** is a CONFIGURATION mode command. You may enter an IPv4 or IPv6 address.
- When applied, the management port on the primary RPM assumes the virtual IP address. Entering the **show interfaces** and **show ip interface brief** commands on the primary RPM management interface will display both the virtual IP address and the actual IP address configured on the interface (see Displaying Information on a Management Interface on page 425).
- A duplicate IP address message is printed for management port's virtual IP address on an RPM failover. This is a harmless error that is generated due to a brief transitory moment during failover when both RPMs' management ports own the virtual IP address, but have different MAC addresses.
- The primary management interface will use only the virtual IP address if it is configured. The system can not be accessed through the native IP address of the primary RPM's management interface.
- Once the virtual IP address is removed, the system is accessible through the native IP address of the primary RPM's management interface.
- Primary and secondary management interface IP and virtual IP must be in the same subnet.

Configure Management Interfaces on the S-Series

The user can manage the S-Series from any port. Configure an IP address for the port using the **ip address** command, and enable it using the command **no shutdown**. The user may use the command **description** from INTERFACE mode to note that the interface is the management interface. There is no separate management routing table, so the user must configure all routes in the IP routing table (the **ip route** command).

As shown in Figure 20-8, from EXEC Privilege mode, display the configuration for a given port by entering the command **show interface**, and the routing table with the **show ip route** command.

Figure 20-8. Viewing Management Routes on the S-Series

```
FTOS#show int gig 0/48
GigabitEthernet 0/48 is up, line protocol is up
Description: This is the Managment Interface
Hardware is Force10Eth, address is 00:01:e8:cc:cc:ce
    Current address is 00:01:e8:cc:cc:ce
Pluggable media not present
Interface index is 46449666
Internet address is 10.11.131.240/23
[output omitted]
FTOS#show ip route

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is 10.11.131.254 to network 0.0.0.0

       Destination        Gateway                     Dist/Metric Last Change
       -----------        -------                     ----------- -----------
 *S    0.0.0.0/0          via 10.11.131.254, Gi 0/48          1/0        1d2h
  C    10.11.130.0/23     Direct, Gi 0/48                     0/0        1d2h
```

Displaying Information on a Management Interface

To view information about the primary RPM management port, use the **show interface Managementethernet** command in EXEC or EXEC Privilege mode. If there are two RPMs on the system, you cannot view information on the interface.

Figure 20-9. Command Example: show interface ManagementEthernet

```
FTOS>show interface managementethernet 0/0

ManagementEthernet 0/0 is up, line protocol is up
Hardware is Force10Eth, address is 00:01:e8:5d:b7:4c
    Current address is 00:01:e8:5d:b7:4c
Pluggable media not present
Interface index is 503595208
Internet address is 10.11.197.97/16
Link local IPv6 address: fe80::201:e8ff:fe5d:b74c/64
Global IPv6 address: fdaa:bbbb:cccc:1004::50/64
Virtual-IP address is 10.11.197.99/16
Virtual-IP IPv6 address is fdaa:bbbb:cccc:1004::60/64


FTOS# show running-config interface managementethernet 0/0

interface ManagementEthernet 0/0
 ip address 10.11.197.97/16
 ipv6 address fdaa:bbbb:cccc:1004::50/64
!virtual-ip 10.11.197.99/16
!virtual-ip fdaa:bbbb:cccc:1004::60/64
 no shutdown
```

To view summary information about the primary RPM Management port, use the **show ip interface brief Managementethernet** or **show ipv6 interface brief Managementethernet** commands in EXEC or EXEC Privilege mode.

Figure 20-10. Command Example: show ip interface brief ManagementEthernet

```
FTOS>show ip interface brief managementethernet 0/0
Interface                IP-Address      OK  Method Status                Protocol
ManagementEthernet 0/0   10.11.197.97    YES Manual up                    up
```

Figure 20-11. Command Example: show ipv6 interface brief ManagementEthernet

```
FTOS#show ipv6 interface brief managementethernet 0/0
ManagementEthernet 0/0           [up/up]
    fe80::201:e8ff:fe5d:b74c
    fdaa:bbbb:cccc:1004::50/64
```

VLAN Interfaces

VLANs are logical interfaces and are, by default, in Layer 2 mode. Physical interfaces and port channels can be members of VLANs. For more information on VLANs and Layer 2, refer to Chapter 10, Layer 2, on page 47. See also Chapter 18, VLAN Stacking, on page 367.

Note: To monitor VLAN interfaces, use the Management Information Base for Network Management of TCP/IP-based internets: MIB-II (RFC 1213). Monitoring VLAN interfaces via SNMP is supported only on E-Series.

FTOS supports Inter-VLAN routing (Layer 3 routing in VLANs). You can add IP addresses to VLANs and use them in routing protocols in the same manner that physical interfaces are used. For more information on configuring different routing protocols, refer to the chapters on the specific protocol.

A consideration for including VLANs in routing protocols is that the **no shutdown** command must be configured. (For routing traffic to flow, the VLAN must be enabled.)

Note: An IP address cannot be assigned to the Default VLAN, which, by default, is VLAN 1. To assign another VLAN ID to the Default VLAN, use the **default vlan-id** *vlan-id* command.

Assign an IP address to an interface with the following command the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip address *ip-address mask* [**secondary**]	INTERFACE	Configure an IP address and mask on the interface. • *ip-address mask:* enter an address in dotted-decimal format (A.B.C.D) and the mask must be in slash format (/24). • **secondary:** the IP address is the interface's backup IP address. You can configure up to eight secondary IP addresses.

Figure 20-12 shows a sample configuration of a VLAN participating in an OSPF process.

Figure 20-12. Sample Layer 3 Configuration of a VLAN

```
interface Vlan 10
 ip address 1.1.1.2/24
 tagged GigabitEthernet 2/2-13
 tagged TenGigabitEthernet 5/0
 ip ospf authentication-key force10
 ip ospf cost 1
 ip ospf dead-interval 60
 ip ospf hello-interval 15
 no shutdown
!
```

Loopback Interfaces

A Loopback interface is a virtual interface in which the software emulates an interface. Packets routed to it are processed locally. Since this interface is not a physical interface, you can configure routing protocols on this interface to provide protocol stability. You can place Loopback interfaces in default Layer 3 mode.

To configure a Loopback interface, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
interface loopback *number*	CONFIGURATION	Enter a number as the loopback interface. Range: 0 to 16383.

To view Loopback interface configurations, use the **show interface loopback** *number* command in the EXEC mode.

To delete a Loopback interface, use the **no interface loopback** *number* command syntax in the CONFIGURATION mode.

Many of the same commands found in the physical interface are found in Loopback interfaces.

See also Configuring ACLs to Loopback on page 151.

Null Interfaces

The Null interface is another virtual interface created by the E-Series software. There is only one Null interface. It is always up, but no traffic is transmitted through this interface.

To enter the INTERFACE mode of the Null interface, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
interface null 0	CONFIGURATION	Enter the INTERFACE mode of the Null interface.

The only configurable command in the INTERFACE mode of the Null interface is the **ip unreachable** command.

Port Channel Interfaces

Port channel interfaces support link aggregation, as described in IEEE Standard 802.3ad.

This section covers the following topics:

Port channel definition and standards

Link aggregation is defined by IEEE 802.3ad as a method of grouping multiple physical interfaces into a single logical interface—a Link Aggregation Group (LAG) or port channel. A LAG is “a group of links that appear to a MAC client as if they were a single link” according to IEEE 802.3ad. In FTOS, a LAG is referred to as a port channel interface.

A port channel provides redundancy by aggregating physical interfaces into one logical interface. If one physical interface goes down in the port channel, another physical interface carries the traffic.

Port channel benefits

For the E-Series, a port channel interface provides many benefits, including easy management, link redundancy, and sharing.

Port channels are transparent to network configurations and can be modified and managed as one interface. For example, you configure one IP address for the group and that IP address is used for all routed traffic on the port channel.

With this feature, the user can create larger-capacity interfaces by utilizing a group of lower-speed links. For example, the user can build a 5-Gigabit interface by aggregating five 1-Gigabit Ethernet interfaces together. If one of the five interfaces fails, traffic is redistributed across the four remaining interfaces.

Port channel implementation

FTOS supports two types of port channels:

- **Static**—Port channels that are statically configured

- **Dynamic**—Port channels that are dynamically configured using Link Aggregation Control Protocol (LACP). For details, see Chapter 24, Link Aggregation Control Protocol.

Table 20-2. Number of Port-channels per Platform

Platform	Port-channels	Members/Channel
E-Series	255	16
C-Series	128	8
S-Series: S50 and S25	52	8
S-Series: S55, S60 and S4810	128	8

Table 20-3. Maximum number of configurable Port-channels

Platform	Port-channels	Members/Channel
E-Series ExaScale	512	64

Table 20-4. As soon as a port channel is configured, FTOS treats it like a physical interface. For example, IEEE 802.1Q tagging is maintained while the physical interface is in the port channel.

Member ports of a LAG are added and programmed into hardware in a predictable order based on the port ID, instead of in the order in which the ports come up. With this implementation, load balancing yields predictable results across line card resets and chassis reloads.

A physical interface can belong to only one port channel at a time.

Each port channel must contain interfaces of the same interface type/speed.

Port channels can contain a mix of 10, 100, or 1000 Mbps Ethernet interfaces and Gigabit Ethernet interfaces, and the interface speed (10, 100, or 1000 Mbps) used by the port channel is determined by the first port channel member that is physically up. FTOS disables the interfaces that do match the interface speed set by the first channel member. That first interface may be the first interface that is physically brought up or was physically operating when interfaces were added to the port channel. For example, if the first operational interface in the port channel is a Gigabit Ethernet interface, all interfaces at 1000 Mbps are kept up, and all 10/100/1000 interfaces that are not set to 1000 speed or auto negotiate are disabled.

FTOS brings up 10/100/1000 interfaces that are set to auto negotiate so that their speed is identical to the speed of the first channel member in the port channel.

10/100/1000 Mbps interfaces in port channels

When both 10/100/1000 interfaces and GigE interfaces are added to a port channel, the interfaces must share a common speed. When interfaces have a configured speed different from the port channel speed, the software disables those interfaces.

The common speed is determined when the port channel is first enabled. At that time, the software checks the first interface listed in the port channel configuration. If that interface is enabled, its speed configuration becomes the common speed of the port channel. If the other interfaces configured in that port channel are configured with a different speed, FTOS disables them.

For example, if four interfaces (Gi 0/0, 0/1, 0/2, 0/3) in which Gi 0/0 and Gi 0/3 are set to speed 100 Mb/s and the others are set to 1000 Mb/s, with all interfaces enabled, and you add them to a port channel by entering **channel-member gigabitethernet 0/0-3** while in the port channel interface mode, and FTOS determines if the first interface specified (Gi 0/0) is up. Once it is up, the common speed of the port channel is 100 Mb/s. FTOS disables those interfaces configured with speed 1000 or whose speed is 1000 Mb/s as a result of auto-negotiation.

In this example, you can change the common speed of the port channel by changing its configuration so the first enabled interface referenced in the configuration is a 1000 Mb/s speed interface. You can also change the common speed of the port channel here by setting the speed of the Gi 0/0 interface to 1000 Mb/s.

Configuration task list for port channel interfaces

To configure a port channel (LAG), you use the commands similar to those found in physical interfaces. By default, no port channels are configured in the startup configuration.

- Create a port channel (mandatory)
- Add a physical interface to a port channel on page 431 (mandatory)
- Reassign an interface to a new port channel on page 433 (optional)
- Configure the minimum oper up links in a port channel (LAG) on page 434 (optional)
- Add or remove a port channel from a VLAN on page 434 (optional)
- Assign an IP address to a port channel on page 435 (optional)
- Delete or disable a port channel on page 435 (optional)
- Load balancing through port channels on page 436 (optional)

Create a port channel

You can create up to 255 port channels on an E-Series (255 for TeraScale and ExaScale, 1 to 32 for EtherScale). You can create up to 128 port channels on an C-Series, 52 port channels with 8 port members per group on an S-Series S50 or S25, and 128 port channels with 8 port members per group on an S-Series S55, S60 and S4810.

To configure a port channel, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**interface port-channel** *id-number*	CONFIGURATION	Create a port channel.
2	**no shutdown**	INTERFACE PORT-CHANNEL	Ensure that the port channel is active.

The port channel is now enabled and you can place the port channel in Layer 2 or Layer 3 mode. Use the **switchport** command to place the port channel in Layer 2 mode or configure an IP address to place the port channel in Layer 3 mode.

You can configure a port channel as you would a physical interface by enabling or configuring protocols or assigning access control lists.

Add a physical interface to a port channel

The physical interfaces in a port channel can be on any line card in the chassis, but must be the same physical type.

Note: Port channels can contain a mix of Gigabit Ethernet and 10/100/1000 Ethernet interfaces, but FTOS disables the interfaces that are not the same speed of the first channel member in the port channel (see 10/100/1000 Mbps interfaces in port channels).

You can add any physical interface to a port channel if the interface configuration is minimal. Only the following commands can be configured on an interface if it is a member of a port channel:

- **description**
- **shutdown**/**no shutdown**
- **mtu**
- **ip mtu** (if the interface is on a Jumbo-enabled by default.)

Note: The S-Series supports jumbo frames by default (the default maximum transmission unit (MTU) is 1554 bytes) You can configure the MTU using the **mtu** command from INTERFACE mode.

To view the interface's configuration, enter the INTERFACE mode for that interface and enter the **show config** command or from the EXEC Privilege mode, enter the **show running-config interface** *interface* command.

When an interface is added to a port channel, FTOS recalculates the hash algorithm.

To add a physical interface to a port channel, use these commands in the following sequence in the INTERFACE mode of a port channel:

Step	Command Syntax	Command Mode	Purpose
1	**channel-member** *interface*	INTERFACE PORT-CHANNEL	Add the interface to a port channel. The *interface* variable is the physical interface type and slot/port information.
2	**show config**	INTERFACE PORT-CHANNEL	Double check that the interface was added to the port channel.

To view the port channel's status and channel members in a tabular format, use the **show interfaces port-channel brief** (Figure 177) command in the EXEC Privilege mode.

Figure 20-13. show interfaces port-channel brief Command Example

```
FTOS#show int port brief

LAG Mode  Status       Uptime    Ports
1   L2L3  up           00:06:03  Gi 13/6    (Up) *
                                 Gi 13/12   (Up)
2   L2L3  up           00:06:03  Gi 13/7    (Up) *
                                 Gi 13/8    (Up)
                                 Gi 13/13   (Up)
                                 Gi 13/14   (Up)
FTOS#
```

Figure 20-14 displays the port channel's mode (L2 for Layer 2 and L3 for Layer 3 and L2L3 for a Layer 2 port channel assigned to a routed VLAN), the status, and the number of interfaces belonging to the port channel.

Figure 20-14. show interface port-channel Command Example

```
FTOS>show interface port-channel 20
Port-channel 20 is up, line protocol is up
Hardware address is 00:01:e8:01:46:fa
Internet address is 1.1.120.1/24
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 2000 Mbit
Members in this channel:  Gi 9/10 Gi 9/17
ARP type: ARPA, ARP timeout 04:00:00
Last clearing of "show interface" counters 00:00:00
Queueing strategy: fifo
     1212627 packets input, 1539872850 bytes
     Input 1212448 IP Packets, 0 Vlans 0 MPLS
     4857 64-byte pkts, 17570 over 64-byte pkts, 35209 over 127-byte pkts
     69164 over 255-byte pkts, 143346 over 511-byte pkts, 942523 over 1023-byte pkts
     Received 0 input symbol errors, 0 runts, 0 giants, 0 throttles
     42 CRC, 0 IP Checksum, 0 overrun, 0 discarded
     2456590833 packets output, 203958235255 bytes, 0 underruns
     Output 1640 Multicasts, 56612 Broadcasts, 2456532581 Unicasts
     2456590654 IP Packets, 0 Vlans, 0 MPLS
     0 throttles, 0 discarded
Rate info (interval 5 minutes):
     Input 00.01Mbits/sec,          2 packets/sec
     Output 81.60Mbits/sec,     133658 packets/sec
Time since last interface status change: 04:31:57

FTOS>
```

When more than one interface is added to a Layer 2 port channel, FTOS selects one of the active interfaces in the port channel to be the Primary Port. The primary port replies to flooding and sends protocol PDUs. An asterisk in the **show interfaces port-channel brief** command indicates the primary port.

As soon as a physical interface is added to a port channel, the properties of the port channel determine the properties of the physical interface. The configuration and status of the port channel are also applied to the physical interfaces within the port channel. For example, if the port channel is in Layer 2 mode, you cannot add an IP address or a static MAC address to an interface that is part of that port channel. As Figure 20-15 illustrates, interface GigabitEthernet 1/6 is part of port channel 5, which is in Layer 2 mode, and an error message appeared when an IP address was configured.

Figure 20-15. Error Message

```
FTOS(conf-if-portch)#show config
!
interface Port-channel 5
 no ip address
 switchport
 channel-member GigabitEthernet 1/6
FTOS(conf-if-portch)#int gi 1/6
FTOS(conf-if)#ip address 10.56.4.4 /24
% Error: Port is part of a LAG Gi 1/6.
FTOS(conf-if)#
```

Error message

Reassign an interface to a new port channel

An interface can be a member of only one port channel. If the interface is a member of a port channel, you must remove it from the first port channel and then add it to the second port channel.

Each time you add or remove a channel member from a port channel, FTOS recalculates the hash algorithm for the port channel.

To reassign an interface to a new port channel, use these commands in the following sequence in the INTERFACE mode of a port channel:

Step	Command Syntax	Command Mode	Purpose
1	**no channel-member** *interface*	INTERFACE PORT-CHANNEL	Remove the interface from the first port channel.
2	**interface port-channel** id *number*	INTERFACE PORT-CHANNEL	Change to the second port channel INTERFACE mode.
3	**channel-member** *interface*	INTERFACE PORT-CHANNEL	Add the interface to the second port channel.

Figure 20-16 displays an example of moving the GigabitEthernet 1/8 interface from port channel 4 to port channel 3.

Figure 20-16. Command Example from Reassigning an Interface to a Different Port Channel

```
FTOS(conf-if-portch)#show config
!
interface Port-channel 4
 no ip address
 channel-member GigabitEthernet 1/8
 no shutdown
FTOS(conf-if-portch)#no chann gi 1/8
FTOS(conf-if-portch)#int port 5
FTOS(conf-if-portch)#channel gi 1/8
FTOS(conf-if-portch)#sho conf
!
interface Port-channel 5
 no ip address
 channel-member GigabitEthernet 1/8
 shutdown
FTOS(conf-if-portch)#
```

Configure the minimum oper up links in a port channel (LAG)

You can configure the minimum links in a port channel (LAG) that must be in “oper up” status for the port channel to be considered to be in “oper up” status. Use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
minimum-links *number*	INTERFACE	Enter the number of links in a LAG that must be in “oper up” status. Default: 1

Figure 20-17 displays an example of configuring five minimum “oper up” links in a port channel.

Figure 20-17. Example of using the minimum-links Command

```
FTOS#config t
FTOS(conf)#int po 1
FTOS(conf-if-po-1)#minimum-links 5
FTOS(conf-if-po-1)#
```

Add or remove a port channel from a VLAN

As with other interfaces, you can add Layer 2 port channel interfaces to VLANs. To add a port channel to a VLAN, you must place the port channel in Layer 2 mode (by using the **switchport** command).

To add a port channel to a VLAN, use either of the following commands:

Command Syntax	Command Mode	Purpose
tagged port-channel *id number*	INTERFACE VLAN	Add the port channel to the VLAN as a tagged interface. An interface with tagging enabled can belong to multiple VLANs.
untagged port-channel *id number*	INTERFACE VLAN	Add the port channel to the VLAN as an untagged interface. An interface without tagging enabled can belong to only one VLAN.

To remove a port channel from a VLAN, use either of the following commands:

Command Syntax	Command Mode	Purpose
no tagged port-channel *id number*	INTERFACE VLAN	Remove the port channel with tagging enabled from the VLAN.
no untagged port-channel *id number*	INTERFACE VLAN	Remove the port channel without tagging enabled from the VLAN.

To see which port channels are members of VLANs, enter the **show vlan** command in the EXEC Privilege mode.

Assign an IP address to a port channel

You can assign an IP address to a port channel and use port channels in Layer 3 routing protocols.

To assign an IP address, use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip address *ip-address mask* [**secondary**]	INTERFACE	Configure an IP address and mask on the interface. • *ip-address mask:* enter an address in dotted-decimal format (A.B.C.D) and the mask must be in slash format (/24). • **secondary:** the IP address is the interface's backup IP address. You can configure up to eight secondary IP addresses.

Delete or disable a port channel

To delete a port channel, you must be in the CONFIGURATION mode and use the **no interface portchannel** *channel-number* command.

When you disable a port channel (using the **shutdown** command) all interfaces within the port channel are operationally down also.

Load balancing through port channels

FTOS uses hash algorithms for distributing traffic evenly over channel members in a port channel (LAG). The hash algorithm distributes traffic among ECMP paths and LAG members. The distribution is based on a flow, except for packet-based hashing. A flow is identified by the hash and is assigned to one link. In packet-based hashing, a single flow can be distributed on the LAG and uses one link.

Packet based hashing is used to load balance traffic across a port-channel based on the IP Identifier field within the packet. Load balancing uses source and destination packet information to get the greatest advantage of resources by distributing traffic over multiple paths when transferring data to a destination.

FTOS allows you to modify the hashing algorithms used for flows and for fragments. The **load-balance** and **hash-algorithm** commands are available for modifying the distribution algorithms. Their syntax and implementation are somewhat different between the E-Series and the C-Series and S-Series.

Note: Hash-based load-balancing on MPLS does not work when packet-based hashing (**load-balance ip-selection packet-based**) is enabled.

E-Series load-balancing

On the E-Series, the default **load-balance** criteria are a 5-tuple, as follows:

- IP source address
- IP destination address
- Protocol type
- TCP/UDP source port
- TCP/UDP destination port

Balancing may be applied to IPv4, switched IPv6, and non-IP traffic. For these traffic types, the IP-header-based hash and MAC-based hash may be applied to packets by using the following methods.

Table 20-5. Hash Methods as Applied to Port Channel Types

Hash (Header Based)	Layer 2 Port Channel	Layer 3 Port Channel
5-tuple	X	X
3-tuple	X	X
Packet-based	X	X
MAC source address (SA) and destination address (DA)	X	

On the E-Series, to change the 5-tuple default to 3-tuple, MAC, or packet-based, use the following command in CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
[**no**] **load-balance** [**ip-selection** {**3-tuple	packet-based**}] [**mac**]	CONFIGURATION	To designate a method to balance traffic over a port channel. By default, IP 5-tuple is used to distribute traffic over members port channel. **ip-selection 3-tuple**—Distribute IP traffic based on IP source address, IP destination address, and IP protocol type. **ip-selection packet-based**—Distribute IPV4 traffic based on the IP Identification field in the IPV4 header. **mac**—Distribute traffic based on the MAC source address, and the MAC destination address. See Table 20-7 for more information.

For details on the **load-balance** command, see the IP Routing chapter of the *FTOS Command Reference*.

To distribute IP traffic over an E-Series port channel member, FTOS uses the 5-tuple IP default. The 5-tuple and the 3-tuple hash use the following keys:

Table 20-6. 5-tuple and 3-tuple Keys

Keys	5-tuple	3-tuple
IP source address (lower 32 bits)	X	X
IP destination address (lower 32 bits)	X	X
Protocol type	X	X
TCP/UDP source port	X	
TCP/UDP destination port	X	

Note: For IPV6, only the first 32 bits (LSB) of IP Source Address and IP Destination Address are used for hash generation.

Figure 20-18 shows the configuration and show command for packet-based hashing on the E-Series.

Figure 20-18. Command example: load-balance ip-selection packet-based

```
FTOS(conf)#load-balance ip-selection packet-based


FTOS#show running-config | grep load
load-balance ip-selection packet-based
FTOS#
```

The **load-balance packet based** command can co-exist with **load balance mac** command to achieve the functionality shown in Table 20-7.

IPv4, IPv6, and non-IP traffic handling on the E-Series

The table below presents the combinations of the **load-balance** command and their effect on traffic types.

Table 20-7. The load-balance Commands and Port Channel Types

Configuration Commands	Switched IP Traffic	Routed IP Traffic (IPv4 only)	Switched Non-IP Traffic
Default (IP 5-tuple)	IP 5-tuple (lower 32 bits)	IP 5-tuple	MAC-based
load-balance ip-selection 3-tuple	IP 3-tuple (lower 32 bits)	IP 3-tuple	MAC-based
load-balance ip-selection mac	MAC-based	IP 5-tuple	MAC-based
load-balance ip-selection 3-tuple **load-balance ip-selection mac**	MAC-based	IP 3-tuple	MAC-based
load-balance ip-selection packet-based	Packet based: IPV4 No distribution: IPV6	Packet-based	MAC-based
load-balance ip-selection packet-based **load-balance ip-selection mac**	MAC-based	Packet-based	MAC-based

C-Series and S-Series load-balancing

For LAG hashing on C-Series and S-Series, the source IP, destination IP, source TCP/UDP port, and destination TCP/UDP port are used for hash computation by default. For packets without a Layer 3 header, FTOS automatically uses **load-balance mac source-dest-mac**.

IP hashing or MAC hashing should not be configured at the same time. If you configure an IP and MAC hashing scheme at the same time, the MAC hashing scheme takes precedence over the IP hashing scheme.

To change the IP traffic load balancing default on the C-Series and S-Series, use the following command:

Command Syntax	Command Mode	Purpose
[**no**] **load-balance** {**ip-selection** [**dest-ip** | **source-ip**]} | {**mac** [**dest-mac** | **source-dest-mac** | **source-mac**]} | {**tcp-udp enable**}	CONFIGURATION	Replace the default IP 4-tuple method of balancing traffic over a port channel. You can select one, two, or all three of the following basic hash methods **ip-selection** [**dest-ip | source-ip**]—Distribute IP traffic based on IP destination or source address. **mac** [**dest-mac | source-dest-mac | source-mac**]—Distribute IPV4 traffic based on the destination or source MAC address, or both, along with the VLAN, Ethertype, source module ID and source port ID. **tcp-udp enable**—Distribute traffic based on TCP/UDP source and destination ports.

Hash algorithm

The **load-balance** command discussed above selects the hash criteria applied to port channels.

If even distribution is not obtained with the load-balance command, the **hash-algorithm** command can be used to select the hash scheme for LAG, ECMP and NH-ECMP. The 12 bit Lag Hash can be rotated or shifted till the desired hash is achieved.

The **nh-ecmp** option allows you to change the hash value for recursive ECMP routes independently of non-recursive ECMP routes. This option provides for better traffic distribution over available equal cost links that involve a recursive next hop lookup.

For the E-Series TeraScale and ExaScale, you can select one of 47 possible hash algorithms (16 on EtherScale).

Command Syntax	Command Mode	Purpose
hash-algorithm {*algorithm-number*}| {**ecmp** {**checksum|crc|xor**} [*number*]} **lag** {*checksum|crc|xor*][*number*]}**nh-ecmp** {{*checksum|crc|xor]* [*number*]}}| {**linecard** *number* **ip-sa-mask** *value* **ip-da-mask** *value*}	CONFIGURATION	Change the default (0) to another algorithm and apply it to ECMP, LAG hashing, or a particular line card. **Note:** To achieve the functionality of hash-align on the ExaScale platform, do not use CRC as an hash-algorithm method. For ExaScale systems, set the default hash-algorithm method to ensure CRC is not used for LAG. For example, **hash-algorithm ecmp xor lag checksum nh-ecmp checksum** For details on the algorithm choices, see the command details in the IP Routing chapter of the *FTOS Command Reference*.

Note: E-Series systems require the **lag-hash-align** microcode be configured in the in the CAM profile. E-Series TeraScale E_T includes this microcode as an option with the Default cam profile. E-Series ExaScale E_X systems require that a CAM profile be created and specifically include **lag-hash-align** microcode.

Figure 20-19 shows a sample configuration for the **hash-algorithm** command.

Figure 20-19. Command example: hash-algorithm

```
FTOS(conf)#FTOS(conf)#hash-algorithm ecmp xor 26 lag crc 26 nh-ecmp checksum 26
FTOS(conf)#
```

On C-Series and S-Series, the **hash-algorithm** command is specific to ECMP groups and has different defaults from the E-Series. The default ECMP hash configuration is **crc-lower**. This takes the lower 32 bits of the hash key to compute the egress port. Other options for ECMP hash-algorithms are:

- **crc-upper** — uses the upper 32 bits of the hash key to compute the egress port

- **dest-ip** — uses destination IP address as part of the hash key
- **lsb** — always uses the least significant bit of the hash key to compute the egress port

To change to another method, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
hash-algorithm ecmp {**crc-upper**} | {**dest-ip**} | {**lsb**}	CONFIGURATION	Change to another algorithm.

For more on load-balancing, see "Equal Cost Multipath and Link Aggregation Frequently Asked Questions" in the E-Series FAQ section (login required) of iSupport:

https://www.force10networks.com/CSPortal20/KnowledgeBase/ToolTips.aspx

Bulk Configuration

Bulk configuration enables you to determine if interfaces are present, for physical interfaces, or, configured, for logical interfaces.

Interface Range

An interface range is a set of interfaces to which other commands may be applied, and may be created if there is at least one valid interface within the range. Bulk configuration excludes from configuration any non-existing interfaces from an interface range. A default VLAN may be configured only if the interface range being configured consists of only VLAN ports.

The **interface range** command allows you to create an interface range allowing other commands to be applied to that range of interfaces.

The interface range prompt offers the interface (with slot and port information) for valid interfaces. The maximum size of an interface range prompt is 32. If the prompt size exceeds this maximum, it displays (...) at the end of the output.

Note: Non-existing interfaces are excluded from interface range prompt. In the following example, Tengigabit 3/0 and VLAN 1000 do not exist.

Note: When creating an interface range, interfaces appear in the order they were entered and are not sorted.

The **show range** command is available under interface range mode. This command allows you to display all interfaces that have been validated under the interface range context.

The **show configuration** command is also available under the interface range mode. This command allows you to display the running configuration only for interfaces that are part of interface range.

Bulk Configuration Examples

The following are examples of using the **interface range** command for bulk configuration:

- Create a single-range
- Create a multiple-range
- Exclude duplicate entries
- Exclude a smaller port range
- Overlap port ranges
- Commas
- Add ranges

Create a single-range

Figure 20-20. Creating a Single-Range Bulk Configuration

```
FTOS(config)# interface range gigabitethernet 5/1 - 23
FTOS(config-if-range-gi-5/1-23)# no shutdown
```

Create a multiple-range

Figure 20-21. Creating a Multiple-Range Prompt

```
FTOS(conf)#interface range tengigabitethernet 3/0 , gigabitethernet 2/1 - 47 , vlan 1000
FTOS(conf-if-range-gi-2/1-47,so-5/0)#
```

Exclude duplicate entries

Duplicate single interfaces and port ranges are excluded from the resulting interface range prompt:

Figure 20-22. Interface Range Prompt Excluding Duplicate Entries

```
FTOS(conf)#interface range vlan 1 , vlan 1 , vlan 3 , vlan 3
FTOS(conf-if-range-vl-1,vl-3)#
FTOS(conf)#interface range gigabitethernet 2/0 - 23 , gigabitethernet 2/0 - 23 , gigab 2/0 - 23
FTOS(conf-if-range-gi-2/0-23)#
```

Exclude a smaller port range

If interface range has multiple port ranges, the smaller port range is excluded from prompt:

Figure 20-23. Interface Range Prompt Excluding a Smaller Port Range

```
FTOS(conf)#interface range gigabitethernet 2/0 - 23 , gigab 2/1 - 10
FTOS(conf-if-range-gi-2/0-23)#
```

Overlap port ranges

If overlapping port ranges are specified, the port range is extended to the smallest start port number and largest end port number:

Figure 20-24. Interface Range Prompt Including Overlapping Port Ranges

```
FTOS(conf)#inte ra gi 2/1 - 11 , gi 2/1 - 23
FTOS(conf-if-range-gi-2/1-23)#
```

Commas

The example below shows how to use commas to add different interface types to the range, enabling all Gigabit Ethernet interfaces in the range 5/1 to 5/23 and both Ten Gigabit Ethernet interfaces 1/1 and 1/2.

```
FTOS(config-if)# interface range gigabitethernet 5/1 - 23, tengigabitethernet 1/1 - 2
FTOS(config-if-range-gi-5/1-23)# no shutdown
FTOS(config-if-range-gi-5/1-23)#
```

Figure 20-25. Multiple-Range Bulk Configuration Gigabit Ethernet and Ten-Gigabit Ethernet

Add ranges

The example below shows how to use commas to add VLAN and port-channel interfaces to the range.

Figure 20-26. Multiple-Range Bulk Configuration with VLAN, and Port-channel

```
FTOS(config-ifrange-gi-5/1-23-te-1/1-2)# interface range Vlan 2 – 100 , Port 1 – 25
FTOS(config-if-range-gi-5/1-23-te-1/1-2-so-5/1-vl-2-100-po-1-25)# no shutdown
FTOS(config-if-range)#
```

Interface Range Macros

The user can define an interface-range macro to automatically select a range of interfaces for configuration. Before you can use the macro keyword in the interface-range macro command string, you must define the macro.

To define an interface-range macro, enter this command:

Command Syntax	Command Mode	Purpose
FTOS (config)# **define** *interface-range macro_name* {**vlan** *vlan_ID - vlan_ID*} | {{**gigabitethernet** | **tengigabitethernet**} *slot/interface - interface*} [**,** {**vlan** *vlan_ID - vlan_ID*} {{**gigabitethernet** | **tengigabitethernet**} *slot/interface - interface*}]	CONFIGURATION	Defines the interface-range macro and saves it in the running configuration file.

Define the Interface Range

This example shows how to define an interface-range macro named “test” to select Fast Ethernet interfaces 5/1 through 5/4:

```
FTOS(config)# define interface-range test gigabitethernet 5/1 - 4
```

To show the defined interface-range macro configuration, use the command **show running-config** in the EXEC mode. The example below shows how to display the defined interface-range macro named “test”:

```
FTOS# show running-config | include define

define interface-range test GigabitEthernet5/1 - 4

FTOS#
```

Choose an Interface-range Macro

To use an interface-range macro in the **interface range** command, enter this command:

Command Syntax	Command Mode	Purpose
interface range macro *name*	CONFIGURATION	Selects the interfaces range to be configured using the values saved in a named interface-range macro.

The example below shows how to change to the interface-range configuration mode using the interface-range macro named “test”.

```
FTOS(config)# interface range macro test
FTOS(config-if)#
```

Monitor and Maintain Interfaces

Monitor interface statistics with the **monitor interface** command. This command displays an ongoing list of the interface status (up/down), number of packets, traffic statistics, etc.

Command Syntax	Command Mode	Purpose
monitor interface *interface*	EXEC Privilege	View the interface’s statistics. Enter the type of interface and slot/port information: • For a 10/100/1000 Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For the Management interface on the RPM, enter the keyword **ManagementEthernet** followed by the slot/port information. • For a SONET interface, enter the keyword **sonet** followed by slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information.

The information (Figure 20-27) displays in a continuous run, refreshing every 2 seconds by default. Use the following keys to manage the output.

m - Change mode	c - Clear screen
l - Page up	a - Page down
T - Increase refresh interval (by 1 second)	t - Decrease refresh interval (by 1 second)
q - Quit	

Figure 20-27. Command Example: monitor interface

```
FTOS#monitor interface gi 3/1


 FTOS uptime is 1 day(s), 4 hour(s), 31 minute(s)
   Monitor time: 00:00:00   Refresh Intvl.: 2s

 Interface: Gi 3/1, Disabled, Link is Down, Linespeed is 1000 Mbit

  Traffic statistics:              Current           Rate                 Delta
         Input bytes:                    0          0 Bps                    0
        Output bytes:                    0          0 Bps                    0
       Input packets:                    0          0 pps                    0
      Output packets:                    0          0 pps                    0
          64B packets:                   0          0 pps                    0
    Over 64B packets:                    0          0 pps                    0
   Over 127B packets:                    0          0 pps                    0
   Over 255B packets:                    0          0 pps                    0
   Over 511B packets:                    0          0 pps                    0
  Over 1023B packets:                    0          0 pps                    0
 Error statistics:
      Input underruns:                   0          0 pps                    0
         Input giants:                   0          0 pps                    0
      Input throttles:                   0          0 pps                    0
            Input CRC:                   0          0 pps                    0
    Input IP checksum:                   0          0 pps                    0
        Input overrun:                   0          0 pps                    0
     Output underruns:                   0          0 pps                    0
     Output throttles:                   0          0 pps                    0

         m - Change mode                          c - Clear screen
         l - Page up                              a - Page down
         T - Increase refresh interval            t - Decrease refresh interval
         q - Quit

q
FTOS#
```

Maintenance using TDR

The Time Domain Reflectometer (TDR) is supported on all Dell Force10 switch/routers. TDR is an assistance tool to resolve link issues that helps detect obvious open or short conditions within any of the four copper pairs. TDR sends a signal onto the physical cable and examines the reflection of the signal that returns. By examining the reflection, TDR is able to indicate whether there is a cable fault (when the cable is broken, becomes unterminated, or if a transceiver is unplugged).

TDR is useful for troubleshooting an interface that is not establishing a link, that is, when the link is flapping or not coming up. TDR is not intended to be used on an interface that is passing traffic. When a TDR test is run on a physical cable, it is important to shut down the port on the far end of the cable. Otherwise, it may lead to incorrect test results.

Note: TDR is an intrusive test. Do not run TDR on a link that is up and passing traffic.

To test the condition of cables on 10/100/1000 BASE-T modules, use the **tdr-cable-test** command:

Step	Command Syntax	Command Mode	Usage
1	**tdr-cable-test gigabitethernet** *<slot>/<port>*	EXEC Privilege	To test for cable faults on the GigabitEthernet cable. • Between two ports, the user must not start the test on both ends of the cable. • The user must enable the interface before starting the test. • The port should be enabled to run the test or the test prints an error message.
2	**show tdr gigabitethernet** *<slot>/<port>*	EXEC Privilege	Displays TDR test results.

Link Debounce Timer

Link Debounce Timer is supported on platform E

This feature is supported on E-Series ExaScale EX with FTOS 8.2.0.1 and later

The Link Debounce Timer feature isolates upper layer protocols on Ethernet switches and routers from very short-term, possibly repetitive interface flaps often caused by network jitter on the DWDM equipment connecting the switch and other devices on a SONET ring. The Link Debounce Timer delays link change notifications, thus decreasing traffic loss due to network configuration. All interfaces have a built-in timer to manage traffic. This feature extends the time allowed by the upper layers.

The SONET ring has its own restore time whenever there is a failure. During this time, however, the Ethernet interface connected to the switch will flap. Link Debounce Timer instructs the Ethernet switch to delay the notification of the link change to the upper layers. If the link state changes again within this period, no notification goes to the upper layers, so that the switch remains unaware of the change.

Note: Enabling the link debounce timer causes link up and link down detections to be delayed, resulting in traffic being blackholed during the debouncing period. This situation might affect the convergence and reconvergence of some Layer 2 and Layer 3 protocols.

Important Points to Remember about Link Debounce Timer

- Link Debounce Timer is configurable on physical ports only.
- Only 1G fiber, 10/100/1000 copper, 10G fiber, 10G copper are supported.
- This feature is not supported on management interfaces or SONET interfaces.
- Link Debounce takes effect only when the operational state of the port is up.
- Link Debounce is supported on interfaces that also have link dampening configured.
- Unlike link dampening, link debounce timer does not notify other protocols.

- Changes made do not affect any ongoing debounces. The timer changes take affect from the next debounce onward.

Assign a debounce time to an interface

Command Syntax	Command Mode	Purpose
link debounce time [*milliseconds*]	INTERFACE	Enter the time to delay link status change notification on this interface. Range: 100-5000 ms • Default for Copper is 3100 ms • Default for Fiber is 100 ms

Figure 20-28. Setting Debounce Time

```
FTOS(conf)#int gi 3/1
FTOS(conf-if-gi-3/1)#link debounce time 150
FTOS(conf-if-gi-3/1)#=
```

Show debounce times in an interface

show interface debounce [type**] [**slot/port**]**	EXEC Privilege	Show the debounce time for the specified interface. Enter the interface type keyword followed by the type of interface and slot/port information: • For a 10/100/1000 Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information.

Figure 20-29. Showing Debounce Time

```
FTOS#
FTOS#show interfaces debounce gigabitethernet 3/1
 Interface               Time(ms)
 GigabitEthernet 3/1     200
FTOS#
```

Note: FTOS rounds the entered debounce time up to the nearest hundredth.
Note in Figure 20-28 that the timer was set at 150 ms, but appears as 200 in Figure 20-29.

Disable ports when one only SFM is available (E300 only)

Selected ports can be shut down when a single SFM is available on the E300 system. Each port to be shut down must be configured individually.

When an E300 system boots up and a single SFM is active this configuration, any ports configured with this feature will be shut down. All other ports are booted up.

Similarly, if an SFM fails (or is removed) in an E300 system with two SFM, ports configured with this feature will be shut down. All other ports are treated normally.

When a second SFM is installed or replaced, all ports are booted up and treated as normally. This feature does not take affect until a single SFM is active in the E300 system.

Disable port on one SFM

This feature must be configured for each interface to shut down in the event that an SFM is disabled. Enter the command **disable-on-sfm-failure** from INTERFACE mode to disable the port when only a single SFM is available.

Link Dampening

Interface state changes occur when interfaces are administratively brought up or down or if an interface state changes. Every time an interface changes state or flaps, routing protocols are notified of the status of the routes that are affected by the change in state, and these protocols go through momentous task of re-converging. Flapping therefore puts the status of entire network at risk of transient loops and black holes.

Link dampening minimizes the risk created by flapping by imposing a penalty for each interface flap and decaying the penalty exponentially. Once the penalty exceeds certain threshold, the interface is put in an "error-disabled" state, and for all practical purposes of routing, the interface is deemed to be "down." Once the interface becomes stable and the penalty decays below a certain threshold, the interface comes up again and the routing protocols re-converge.

Link dampening:

- reduces processing on the CPUs by reducing excessive interface flapping.
- improves network stability by penalizing misbehaving interfaces and redirecting traffic
- improves convergence times and stability throughout the network by isolating failures so that disturbances are not propagated.

Important Points to Remember

- Link dampening is not supported on VLAN interfaces
- Link dampening is disabled when the interface is configured for port monitoring
- Link dampening can be applied to Layer 2 and Layer 3 interfaces.
- Link dampening can be configured on individual interfaces in a LAG.

Enable Link Dampening

Enable link dampening using the command **dampening** from INTERFACE mode, as shown in Figure 20-30.

Figure 20-30. Configuring Link Dampening

```
R1(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
 ip address 10.10.19.1/24
 dampening 1 2 3 4
 no shutdown
R1(conf-if-gi-1/1)#exit
```

View the link dampening configuration on an interface using the command **show config**, or view dampening information on all or specific dampened interfaces using the command **show interfaces dampening** from EXEC Privilege mode, as shown in Figure 20-31.

Figure 20-31. Viewing all Dampened Interfaces

```
FTOS# show interfaces dampening
InterfaceState Flaps   Penalty Half-LifeReuse SuppressMax-Sup
Gi 0/0         Up      0       0       5              750     2500           20
Gi 0/1         Up      2       1200    20             500     1500           300
Gi 0/2         Down    4       850     30             600     2000           120
```

View a dampening summary for the entire system using the command show interfaces dampening summary from EXEC Privilege mode, as shown in Figure 20-32.

Figure 20-32. Viewing a System-wide Dampening Summary

```
FTOS# show interfaces dampening summary
20 interfaces are configured with dampening. 3 interfaces are currently suppressed.
Following interfaces are currently suppressed:
Gi 0/2
Gi 3/1
Gi 4/2
FTOS#
```

Clear Dampening Counters

Clear dampening counters and accumulated penalties using the command **clear dampening**, as shown in Figure 20-33.

Figure 20-33. Clearing Dampening Counters

```
FTOS# clear dampening interface Gi 0/1

FTOS# show interfaces dampening GigabitEthernet0/0
InterfaceState Flaps   Penalty Half-LifeReuse SuppressMax-Sup
Gi 0/1  Up     0       0       20             500     1500          300
```

Link Dampening Support for XML

View the output of the following show commands in XML by adding **| display xml** to the end of the command:

- show interfaces dampening
- show interfaces dampening summary
- show interfaces interface x/y

Configure MTU size on an Interface

The E-Series supports a link Maximum Transmission Unit (MTU) of 9252 bytes and maximum IP MTU of 9234 bytes. The link MTU is the frame size of a packet, and the IP MTU size is used for IP fragmentation. If the system determines that the IP packet must be fragmented as it leaves the interface, FTOS divides the packet into fragments no bigger than the size set in the **ip mtu** command.

In FTOS, MTU is defined as the entire Ethernet packet (Ethernet header + FCS + payload)

Since different networking vendors define MTU differently, check their documentation when planing MTU sizes across a network.

Table 20-8 lists the range for each transmission media.

Table 20-8. MTU Range

Transmission Media	MTU Range (in bytes)
Ethernet	594-9252 = link MTU 576-9234 = IP MTU

Ethernet Pause Frames

Ethernet Pause Frames is supported on platforms C E S

Threshold Settings are supported only on platforms: C S

Ethernet Pause Frames allow for a temporary stop in data transmission. A situation may arise where a sending device may transmit data faster than a destination device can accept it. The destination sends a PAUSE frame back to the source, stopping the sender's transmission for a period of time.

The globally assigned 48-bit Multicast address 01-80-C2-00-00-01 is used to send and receive pause frames. To allow full duplex flow control, stations implementing the pause operation instruct the MAC to enable reception of frames with destination address equal to this multicast address.

The PAUSE frame is defined by IEEE 802.3x and uses MAC Control frames to carry the PAUSE commands. Ethernet Pause Frames are supported on full duplex only. The only configuration applicable to half duplex ports is **rx off tx off**.

```
Note that if a port is over-subscribed, Ethernet Pause Frame flow control does not
ensure no loss behavior.
```

```
The following error message appears when trying to enable flow control when half duplex
is already configured: Can’t configure flowcontrol when half duplex is configure,
config ignored.
```

```
The following error message appears when trying to enable half duplex and flow control
configuration is on: Can’t configure half duplex when flowcontrol is on, config
ignored.
```

Threshold Settings

Threshold Settings are supported only on platforms: C S

When the transmission pause is set (**tx on**), 3 thresholds can be set to define the controls more closely. Ethernet Pause Frames flow control can be triggered when either the flow control buffer threshold or flow control packet pointer threshold is reached. The thresholds are:

- Number of flow-control packet pointers: 1-2047 (default = 75)
- Flow-control buffer threshold in KB: 1-2013 (default = 49KB)
- Flow-control discard threshold in KB: 1-2013 (default= 75KB)

The pause is started when *either* the packet pointer or the buffer threshold is met (whichever is met first). When the discard threshold is met, packets are dropped.

The pause ends when *both* the packet pointer and the buffer threshold fall below 50% of the threshold settings.

The discard threshold defines when the interface starts dropping the packet on the interface. This may be necessary when a connected device doesn’t honor the flow control frame sent by S-Series.

The discard threshold should be larger than the buffer threshold so that the buffer holds at least hold at least 3 packets.

Enable Pause Frames

Note: On the C-Series and S-Series platforms, Ethernet Pause Frames TX should be enabled *only after* consulting with the Dell Force10 Technical Assistance Center.

Ethernet Pause Frames flow control must be enabled on all ports on a chassis or a line card. If not, the system may exhibit unpredictable behavior.

On the C-Series and S-Series systems, the flow-control sender and receiver must be on the same port-pipe. Flow control is not supported across different port-pipes on the C-Series or S-Series system.

On 4-port 10G line cards: Changes in the flow-control values are not reflected automatically in the **show interface** output for 10G interfaces. This issue results from the fact that 10G interfaces do not support auto-negotiation per-se. On 1G interfaces, changing the flow control values causes an automatic interface flap, after which PAUSE values are exchanged as part of the auto-negotiation process. As a workaround, apply the new settings, execute **shut** followed by **no shut** on the interface, and then check the running-config of the port.

Command Syntax	Command Mode	Purpose
flowcontrol rx [*off* | *on*] **tx** [*off* | *on*] [*threshold* {<1-2047> <1-2013> <1-2013>}]	INTERFACE	Control how the system responds to and generates 802.3x pause frames on 1 and 10Gig line cards. Defaults: C-Series: **rx off tx off** E-Series: **rx on tx on** S-Series: **rx off tx off**
		Parameters: **rx on**: Enter the keywords **rx on** to process the received flow control frames on this port. **rx off**: Enter the keywords **rx off** to ignore the received flow control frames on this port. **tx on**: Enter the keywords **tx on** to send control frames from this port to the connected device when a higher rate of traffic is received. **tx off**: Enter the keywords **tx off** so that flow control frames are not sent from this port to the connected device when a higher rate of traffic is received. **threshold** (C-Series and S-Series only)**:** When **tx on** is configured, the user can set the threshold values for: Number of flow-control packet pointers: 1-2047 (default = 75) Flow-control buffer threshold in KB: 1-2013 (default = 49KB) Flow-control discard threshold in KB: 1-2013 (default= 75KB) Pause control is triggered when either the flow control buffer threshold or flow control packet pointer threshold is reached.

Configure MTU Size on an Interface

If a packet includes a Layer 2 header, the difference in bytes between the link MTU and IP MTU must be large enough to include the Layer 2 header. For example, for VLAN packets, if the IP MTU is 1400 bytes, the Link MTU must 1422 bytes or greater to accommodate the 22-byte VLAN header:

1400-byte IP MTU + 22-byte VLAN Tag = 1422-byte link MTU

On the E-Series and C-Series, you configure the Link MTU size on an interface by entering the **mtu** command.

On the E-Series, you must manually configure the IP MTU size on an interface by entering the **ip mtu** command. On the C-Series and S-Series, the IP MTU size is automatically configured on an interface by taking into account the Layer 2 overheads and difference in the number of bytes as shown in Table 20-9.

Table 20-9. Difference between Link MTU and IP MTU

Layer 2 Overhead	Difference between link MTU and IP MTU
Ethernet (untagged)	18 bytes
VLAN Tag	22 bytes
Untagged Packet with VLAN-Stack Header	22 bytes
Tagged Packet with VLAN-Stack Header	26 bytes

Link MTU and IP MTU considerations for port channels and VLANs are as follows.

Port Channels:

- All members must have the same link MTU value and the same IP MTU value.
- The port channel link MTU and IP MTU must be less than or equal to the link MTU and IP MTU values configured on the channel members.

Example: If the members have a link MTU of 2100 and an IP MTU 2000, the port channel's MTU values cannot be higher than 2100 for link MTU or 2000 bytes for IP MTU.

VLANs:

- All members of a VLAN must have the same IP MTU value.
- Members can have different Link MTU values. Tagged members must have a link MTU 4 bytes higher than untagged members to account for the packet tag.
- The VLAN link MTU and IP MTU must be less than or equal to the link MTU and IP MTU values configured on the VLAN members.

Example: The VLAN contains tagged members with Link MTU of 1522 and IP MTU of 1500 and untagged members with Link MTU of 1518 and IP MTU of 1500. The VLAN's Link MTU cannot be higher than 1518 bytes and its IP MTU cannot be higher than 1500 bytes.

Port-pipes

A port pipe is a Dell Force10 specific term for the hardware path that packets follow through a system. Port pipes travel through a collection of circuits (ASICs) built into line cards and RPMs on which various processing events for the packets occur. One or two port pipes process traffic for a given set of physical interfaces or a port-set. The E300 only supports one port pipe per slot. On the E1200 and E600 each slot has two port pipes with following specifications:

- 48 port line rate cards have two port pipes on the line card
- 48 port high density cards have only one port pipe on the line card

Note: All references to the E1200 in this section include the E1200i-AC and E1200i-DC. References to E600 include the E600i.

For the purposes of diagnostics, the major difference between the E-Series platforms is the number of port pipes per slot.

- E1200 and E600—Each slot has two port-pipes. Each portpipe has nine 3.125Gbps channels to the backplane, one to each SFM.
- E300—Each slot has one portpipe. Each port-pipe has eight 3.125Gbps channels to the backplane, with four channels to each SFM.

Table 20-10 presents these platform differences again.

Table 20-10. Platform Differences Concerning Port-pipes

Chassis Type	Port-pipes / Slot	Channels / Port-pipe	Capacity of Each Channel (Gbps)	Raw Slot Capacity (Gbps)
E1200/E1200i-AC/DC	2	9	3.125	56.25
E600/E600i	2	9	3.125	56.25
E300	1	8	3.125	25

Auto-Negotiation on Ethernet Interfaces

Setting speed and duplex mode of Ethernet Interfaces

By default, auto-negotiation of speed and duplex mode is enabled on 10/100/1000 Base-T Ethernet interfaces. Only 10GE interfaces do not support auto-negotiation. When using 10GE interfaces, verify that the settings on the connecting devices are set to no auto-negotiation.

Note: Starting with FTOS 7.8.1.0, when a copper SFP2 module with catalog number GP-SFP2-1T is used in the S25P model of the S-Series, its speed can be manually set with the **speed** command. When the speed is set to 10 or 100 Mbps, the **duplex** command can also be executed.

The local interface and the directly connected remote interface must have the same setting, and auto-negotiation is the easiest way to accomplish that, as long as the remote interface is capable of auto-negotiation.

Note: As a best practice, Dell Force10 recommends keeping auto-negotiation enabled. Auto-negotiation should only be disabled on switch ports that attach to devices not capable of supporting negotiation or where connectivity issues arise from interoperability issues.

For 10/100/1000 Ethernet interfaces, the **negotiation auto** command is tied to the **speed** command. Auto-negotiation is always enabled when the **speed** command is set to **1000** or **auto**.

To discover whether the remote and local interface require manual speed synchronization, and to manually synchronize them if necessary, use the following command sequence (see Figure 20-35 on page 456):

Step	Task	Command Syntax	Command Mode
1	Determine the local interface status. See Figure 20-34.	**show interfaces** [*interface* | **linecard** *slot-number*] **status**	EXEC Privilege
2	Determine the remote interface status.	[Use the command on the remote system that is equivalent to the above command.]	EXEC EXEC Privilege
3	Access CONFIGURATION mode.	**config**	EXEC Privilege
4	Access the port.	**interface** *interface slot/port*	CONFIGURATION
5	Set the local port speed.	**speed** {**10** | **100** | **1000** | **auto**}	INTERFACE
6	Optionally, set full- or half-duplex.	**duplex** {**half** | **full**}	INTERFACE
7	Disable auto-negotiation on the port. If the speed was set to 1000, auto-negotiation does not need to be disabled.	**no negotiation auto**	INTERFACE
8	Verify configuration changes.	**show config**	INTERFACE

Note: The **show interfaces status** command displays link status, but not administrative status. For link and administrative status, use **show ip interface** [**interface | brief | linecard slot-number**] [**configuration**].

Figure 20-34. show interfaces status Command Example

```
FTOS#show interfaces status
Port    Description Status Speed      Duplex Vlan
Gi 0/0              Up     1000 Mbit  Auto    --
Gi 0/1              Down   Auto       Auto    1
Gi 0/2              Down   Auto       Auto    1
Gi 0/3              Down   Auto       Auto    --
Gi 0/4 Force10Port  Up     1000 Mbit  Auto   30-130
Gi 0/5              Down   Auto       Auto    --
Gi 0/6              Down   Auto       Auto    --
Gi 0/7              Up     1000 Mbit  Auto   1502,1504,1506-1508,1602
Gi 0/8              Down   Auto       Auto    --
Gi 0/9              Down   Auto       Auto    --
Gi 0/10             Down   Auto       Auto    --
Gi 0/11             Down   Auto       Auto    --
Gi 0/12             Down   Auto       Auto    --
[output omitted]
```

In the example, above, several ports display “Auto” in the Speed field, including port 0/1. In Figure 20-35, the speed of port 0/1 is set to 100Mb and then its auto-negotiation is disabled.

Figure 20-35. Setting Port Speed Example

```
FTOS#configure
FTOS(config)#interface gig 0/1
FTOS(Interface 0/1)#speed 100
FTOS(Interface 0/1)#duplex full
FTOS(Interface 0/1)#no negotiation auto
FTOS(Interface 0/1)#show config
!
interface GigabitEthernet 0/1
no ip address
speed 100
duplex full
no shutdown
```

Setting Auto-Negotiation Options

The **negotiation auto** command provides a **mode** option for configuring an individual port to forced master/forced slave once auto-negotiation is enabled.

Caution: Ensure that only one end of the node is configured as forced-master and the other is configured as forced-slave. If both are configured the same (that is both as forced-master or both as forced-slave), the **show interface** command will flap between an auto-neg-error and forced-master/slave states.

Figure 20-36. Setting Auto-Negotiation Options

```
FTOS(conf)# int gi 0/0
FTOS(conf-if)#neg auto
FTOS(conf-if-autoneg)# ?

end                    Exit from configuration mode
exit                   Exit from autoneg configuration mode
mode                   Specify autoneg mode
no                     Negate a command or set its defaults
```

For details on the **speed**, **duplex**, and **negotiation auto** commands, see the Interfaces chapter of the *FTOS Command Reference*.

Adjust the keepalive timer

Use the **keepalive** command to change the time interval between keepalive messages on the interfaces. The interface sends keepalive messages to itself to test network connectivity on the interface.

To change the default time interval between keepalive messages, use the following command:

Command Syntax	Command Mode	Purpose
keepalive [*seconds*]	INTERFACE	Change the default interval between keepalive messages.

To view the new setting, use the **show config** command in the INTERFACE mode.

View Advanced Interface Information

Display Only Configured Interfaces

The following options have been implemented for **show [ip | running-config] interfaces** commands for (only) linecard interfaces. When the **configured** keyword is used, only interfaces that have non-default configurations are displayed. Dummy linecard interfaces (created with the **linecard** command) are treated like any other physical interface.

Figure 20-37 lists the possible show commands that have the configured keyword available:

Figure 20-37. show Commands with configured Keyword Examples

```
FTOS#show interfaces configured
FTOS#show interfaces linecard 0 configured
FTOS#show interfaces gigabitEthernet 0 configured
FTOS#show ip interface configured
FTOS#show ip interface linecard 1 configured
FTOS#show ip interface gigabitEthernet 1 configured
FTOS#show ip interface br configured
FTOS#show ip interface br linecard 1 configured
FTOS#show ip interface br gigabitEthernet 1 configured
FTOS#show running-config interfaces configured
FTOS#show running-config interface gigabitEthernet 1 configured
```

In EXEC mode, the **show interfaces switchport** command displays only interfaces in Layer 2 mode and their relevant configuration information. The **show interfaces switchport** command (Figure 20-38) displays the interface, whether the interface supports IEEE 802.1Q tagging or not, and the VLANs to which the interface belongs.

Figure 20-38. show interfaces switchport Command Example

```
FTOS#show interfaces switchport
Name: GigabitEthernet 13/0
802.1QTagged: True
Vlan membership:
Vlan    2


Name: GigabitEthernet 13/1
802.1QTagged: True
Vlan membership:
Vlan    2


Name: GigabitEthernet 13/2
802.1QTagged: True
Vlan membership:
```

Configure Interface Sampling Size

Use the **rate-interval** command, in INTERFACE mode, to configure the number of seconds of traffic statistics to display in the **show interfaces** output.

Although any value between 30 and 299 seconds (the default) can be entered, software polling is done once every 15 seconds. So, for example, if you enter “19”, you will actually get a sample of the past 15 seconds.

All LAG members inherit the rate interval configuration from the LAG.

Figure 20-39 shows how to configure rate interval when changing the default value:

Figure 20-39. Configuring Rate Interval Example

```
FTOS#show interfaces
TenGigabitEthernet 10/0 is down, line protocol is down
Hardware is Force10Eth, address is 00:01:e8:01:9e:d9
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 10000 Mbit
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 1d23h44m
Queueing strategy: fifo
     0 packets input, 0 bytes
     Input 0 IP Packets, 0 Vlans 0 MPLS
     0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     Received 0 input symbol errors, 0 runts, 0 giants, 0 throttles
     0 CRC, 0 IP Checksum, 0 overrun, 0 discarded
     0 packets output, 0 bytes, 0 underruns
     Output 0 Multicasts, 0 Broadcasts, 0 Unicasts
     0 IP Packets, 0 Vlans, 0 MPLS
     0 throttles, 0 discarded
Rate info (interval 299 seconds):
     Input 00.00 Mbits/sec,          0 packets/sec, 0.00% of line-rate
     Output 00.00 Mbits/sec,          0 packets/sec, 0.00% of line-rate
Time since last interface status change: 1d23h40m

FTOS(conf)#interface tengigabitethernet 10/0
FTOS(conf-if-te-10/0)#rate-interval 100

FTOS#show interfaces
TenGigabitEthernet 10/0 is down, line protocol is down
Hardware is Force10Eth, address is 00:01:e8:01:9e:d9
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 10000 Mbit
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 1d23h45m
Queueing strategy: fifo
     0 packets input, 0 bytes
     Input 0 IP Packets, 0 Vlans 0 MPLS
     0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     Received 0 input symbol errors, 0 runts, 0 giants, 0 throttles
     0 CRC, 0 IP Checksum, 0 overrun, 0 discarded
     0 packets output, 0 bytes, 0 underruns
     Output 0 Multicasts, 0 Broadcasts, 0 Unicasts
     0 IP Packets, 0 Vlans, 0 MPLS
     0 throttles, 0 discarded
Rate info (interval 100 seconds):
     Input 00.00 Mbits/sec,          0 packets/sec, 0.00% of line-rate
     Output 00.00 Mbits/sec,          0 packets/sec, 0.00% of line-rate
Time since last interface status change: 1d23h42m
```

Default value of 299 seconds

Change rate interval to 100

New rate interval set to 100

Dynamic Counters

By default, counting for the following four applications is enabled:

- IPFLOW
- IPACL
- L2ACL
- L2FIB

For remaining applications, FTOS automatically turns on counting when the application is enabled, and is turned off when the application is disabled. Please note that if more than four counter-dependent applications are enabled on a port pipe, there is an impact on line rate performance.

The following counter-dependent applications are supported by FTOS:

- Egress VLAN
- Ingress VLAN
- Next Hop 2
- Next Hop 1
- Egress ACLs
- ILM
- IP FLOW
- IP ACL
- IP FIB
- L2 ACL
- L2 FIB

Clear interface counters

The counters in the **show interfaces** command are reset by the **clear counters** command. This command does not clear the counters captured by any SNMP program.

To clear the counters, use the following command in the EXEC Privilege mode:

Command Syntax	Command Mode	Purpose
clear counters [*interface*] [**vrrp** [{[**ipv6**] *vrid* | **vrf** *instance*}] | **learning-limit**]	EXEC Privilege	Clear the counters used in the **show interface** commands for all VRRP groups, VLANs, and physical interfaces or selected ones. Without an interface specified, the command clears all interface counters. **(**OPTIONAL) To clear counters on a specified interface, enter one of the following keywords and slot/port or number: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a Loopback interface, enter the keyword **loopback** followed by a number from 0 to 16383. • For a Port Channel interface, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale. • For the management interface on the RPM, enter the keyword **ManagementEthernet** followed by slot/port information. The slot range is 0-1, and the port range is 0. • For a SONET interface, enter the keyword **sonet** followed by the slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094 E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS. **(**OPTIONAL) Enter the keyword **vrrp** to clear the counters of all VRRP groups. To clear the counters of VRRP groups on all IPv6 interfaces, enter **vrrp ipv6**. To clear the counters of a specified VRRP group, enter a *vrid* number from 1 to 255. **E-Series only:** To clear the counters of VRRP groups in a specified VRF instance, enter **vrrp vrf** *instance* (32 characters maximum). **(**OPTIONAL) Enter the keyword **learning-limit** to clear unknown source address (SA) drop counters when MAC learning limit is configured on the interface.

When you enter this command, you must confirm that you want FTOS to clear the interface counters for that interface (Figure 20-40).

Figure 20-40. Clearing an Interface

```
FTOS#clear counters gi 0/0
Clear counters on GigabitEthernet 0/0 [confirm]
FTOS#
```

21

IPv4 Addressing

IPv4 Addressing is supported on platforms C E S

IPv4 addressing is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

FTOS supports various IP addressing features. This chapter explains the basics of Domain Name Service (DNS), Address Resolution Protocol (ARP), and routing principles and their implementation in FTOS.

Table 21-1 lists the defaults for the IP addressing features described in this chapter.

Table 21-1. IP Defaults

IP Feature	Default
DNS	Disabled
Directed Broadcast	Disabled
Proxy ARP	Enabled
ICMP Unreachable	Disabled
ICMP Redirect	Disabled

IP Addresses

FTOS supports IP version 4, as described in RFC 791. It also supports classful routing and Variable Length Subnet Masks (VLSM). With VLSM one network can be can configured with different masks. Supernetting, which increases the number of subnets, is also supported. Subnetting is when a mask is added to the IP address to separate the network and host portions of the IP address.

At its most basic level, an IP address is 32-bits composed of network and host portions and represented in dotted decimal format. For example,

00001010110101100101011110000011

is represented as 10.214.87.131

For more information on IP addressing, refer to RFC 791, *Internet Protocol*.

Implementation Information

In FTOS, you can configure any IP address as a static route except IP addresses already assigned to interfaces.

Note: FTOS versions 7.7.1.0 and later support 31-bit subnet masks (/31, or 255.255.255.254) as defined by RFC 3021. This feature allows you to save two more IP addresses on point-to-point links than 30-bit masks. FTOS supports RFC 3021 with ARP.

Configuration Task List for IP Addresses

The following list includes the configuration tasks for IP addresses:

- Assign IP addresses to an interface on page 464 (mandatory)
- Configure static routes on page 466 (optional)
- Configure static routes for the management interface on page 467 (optional)

For a complete listing of all commands related to IP addressing, refer to *FTOS Command Line Interface Reference*.

Assign IP addresses to an interface

Assign primary and secondary IP addresses to physical or logical (for example, VLAN or port channel) interfaces to enable IP communication between the E-Series and hosts connected to that interface. In FTOS, you can assign one primary address and up to 255 secondary IP addresses to each interface.

To assign an IP address to an interface, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**interface** *interface*	CONFIGURATION	Enter the keyword **interface** followed by the type of interface and slot/port information: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a Loopback interface, enter the keyword **loopback** followed by a number from 0 to 16383. • For the Management interface on the RPM, enter the keyword **ManagementEthernet** followed by the slot/port information. The slot range is 0-1 and the port range is 0. • For a port channel interface, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale. • For a SONET interface, enter the keyword **sonet** followed by the slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN interface, enter the keyword **vlan** followed by a number from 1 to 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.
2	**no shutdown**	INTERFACE	Enable the interface.
3	**ip address** *ip-address mask* [**secondary**]	INTERFACE	Configure a primary IP address and mask on the interface. • *ip-address mask:* IP address must be in dotted decimal format (A.B.C.D) and the mask must be in slash prefix-length format (/24). Add the keyword **secondary** if the IP address is the interface's backup IP address. You can configure up to eight secondary IP addresses.

To view the configuration, use the **show config** command (Figure 246) in the INTERFACE mode or **show ip interface** in the EXEC privilege mode (Figure 247).

Figure 21-1. show config Command Example in the INTERFACE Mode

```
FTOS(conf-if)#show conf
!
interface GigabitEthernet 0/0
 ip address 10.11.1.1/24
 no shutdown
!
FTOS(conf-if)#
```

Figure 21-2. show ip interface Command Example

```
FTOS#show ip int gi 0/8
GigabitEthernet 0/8 is up, line protocol is up
Internet address is 10.69.8.1/24
Broadcast address is 10.69.8.255
Address determined by config file
MTU is 1554 bytes
Inbound  access list is not set
Proxy ARP is enabled
Split Horizon is enabled
Poison Reverse is disabled
ICMP redirects are not sent
ICMP unreachables are not sent


FTOS#
```

Configure static routes

A static route is an IP address that is manually configured and not learned by a routing protocol, such as OSPF. Often static routes are used as backup routes in case other dynamically learned routes are unreachable.

To configure a static route, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip route *ip-address mask* {*ip-address* | *interface* [*ip-address*]} [*distance*] [**permanent**] [**tag** *tag-value*]	CONFIGURATION	Configure a static IP address. Use the following required and optional parameters: • *ip-address*: Enter an address in dotted decimal format (A.B.C.D). • *mask*: Enter a mask in slash prefix-length format (/X). • *interface*: Enter an interface type followed by slot/port information. • *distance* range: 1 to 255 (optional). • **permanent:** Keep the static route in the routing table (if *interface* option is used) even if the interface with the route is disabled. (optional) • **tag** *tag-value* range: 1 to 4294967295. (optional)

You can enter as many static IP addresses as necessary.

To view the configured routes, use the **show ip route static** command.

Figure 21-3. show ip route static Command Example (partial)

```
FTOS#show ip route static
       Destination        Gateway                     Dist/Metric Last Change
       -----------        -------                     ----------- -----------
  S    2.1.2.0/24         Direct, Nu 0                        0/0    00:02:30
  S    6.1.2.0/24         via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.2/32         via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.3/32         via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.4/32         via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.5/32         via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.6/32         via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.7/32         via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.8/32         via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.9/32         via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.10/32        via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.11/32        via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.12/32        via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.13/32        via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.14/32        via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.15/32        via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.16/32        via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    6.1.2.17/32        via 6.1.20.2, Te 5/0                1/0    00:02:30
  S    11.1.1.0/24        Direct, Nu 0                        0/0    00:02:30
                          Direct, Lo 0
--More--
```

FTOS installs a next hop that is on the directly connected subnet of current IP address on the interface (for example, if interface gig 0/0 is on 172.31.5.0 subnet, FTOS installs the static route).

FTOS also installs a next hop that is not on the directly connected subnet but which recursively resolves to a next hop on the interface's configured subnet. For example, if gig 0/0 has ip address on subnet 2.2.2.0 and if 172.31.5.43 recursively resolves to 2.2.2.0, FTOS installs the static route.

- When interface goes down, FTOS withdraws the route.
- When interface comes up, FTOS re-installs the route.
- When recursive resolution is “broken,” FTOS withdraws the route.
- When recursive resolution is satisfied, FTOS re-installs the route.

Configure static routes for the management interface

When an IP address used by a protocol and a static management route exists for the same prefix, the protocol route takes precedence over the static management route.

To configure a static route for the management port, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
management route *ip-address mask* {*forwarding-router-address* | **ManagementEthernet** *slot/port*}	CONFIGURATION	Assign a static route to point to the management interface or forwarding router.

To view the configured static routes for the management port, use the **show ip management-route** command in the EXEC privilege mode.

Figure 21-4. show ip management-route Command Example

```
FTOS>show ip management-route

Destination         Gateway                      State
-----------         -------                      -----
1.1.1.0/24          172.31.1.250                 Active
172.16.1.0/24       172.31.1.250                 Active
172.31.1.0/24       ManagementEthernet 1/0       Connected

FTOS>
```

Directed Broadcast

By default, FTOS drops directed broadcast packets destined for an interface. This default setting provides some protection against Denial of Service (DOS) attacks.

To enable FTOS to receive directed broadcasts, use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip directed-broadcast	INTERFACE	Enable directed broadcast.

To view the configuration, use the **show config** command in the INTERFACE mode.

Resolution of Host Names

Domain Name Service (DNS) maps host names to IP addresses. This feature simplifies such commands as Telnet and FTP by allowing you to enter a name instead of an IP address.

Dynamic resolution of host names is disabled by default. Unless the feature is enabled, the system resolves only host names entered into the host table with the **ip host** or **ipv6 host** command.

Enable dynamic resolution of host names

By default, dynamic resolution of host names (DNS) is disabled.

To enable DNS, use the following commands in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip domain-lookup	CONFIGURATION	Enable dynamic resolution of host names.
ip name-server *ipv4-address* [*ipv4-address2 ... ipv4-address6*] **ipv6 name-server** *ipv6-address* [*ipv6-address2 ... ipv6-address6*]	CONFIGURATION	Specify up to 6 IPv4 or IPv6 name servers. The order you entered the servers determines the order of their use. You may have IPv4 and IPv6 name servers configured at the same time.

To view current bindings, use the **show hosts** command.

Figure 21-5. show hosts Command Example

```
FTOS>show host
Default domain is force10networks.com
Name/address lookup uses domain service
Name servers are not set
Host                      Flags      TTL    Type   Address
--------                  -----      ----   ----   -------
ks                        (perm, OK) -       IP    2.2.2.2
patch1                    (perm, OK) -       IP    192.68.69.2
tomm-3                    (perm, OK) -       IP    192.68.99.2
gxr                       (perm, OK) -       IP    192.71.18.2
f00-3                     (perm, OK) -       IP    192.71.23.1
FTOS>
```

To view the current configuration, use the **show running-config resolve** command.

Specify local system domain and a list of domains

If you enter a partial domain, FTOS can search different domains to finish or fully qualify that partial domain. A fully qualified domain name (FQDN) is any name that is terminated with a period/dot. FTOS searches the host table first to resolve the partial domain. The host table contains both statically configured and dynamically learnt host and IP addresses. If FTOS cannot resolve the domain, it tries the domain name assigned to the local system. If that does not resolve the partial domain, FTOS searches the list of domains configured

To configure a domain name, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip domain-name *name*	CONFIGURATION	Configure one domain name for the E-Series

To configure a list of domain names, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip domain-list *name*	CONFIGURATION	Configure names to complete unqualified host names. Configure this command up to 6 times to specify a list of possible domain names. FTOS searches the domain names in the order they were configured until a match is found or the list is exhausted.

DNS with traceroute

To configure your switch to perform DNS with traceroute, follow the steps below in the CONFIGURATION mode.

Command Syntax	Command Mode	Purpose
ip domain-lookup	CONFIGURATION	Enable dynamic resolution of host names.
ip name-server *ipv4-address* [*ipv4-address2 ... ipv4-address6*] **ipv6 name-server** *ipv6-address* [*ipv6-address2 ... ipv6-address6*]	CONFIGURATION	Specify up to 6 IPv4 or IPv6 name servers. The order you entered the servers determines the order of their use. You may have IPv4 and IPv6 name servers configured at the same time.
traceroute [*host* | *ipv4-address* | *ipv6-address*]	CONFIGURATION	When you enter the traceroute command without specifying an IP address (Extended Traceroute), you are prompted for a target and source IP address, timeout in seconds (default is 5), a probe count (default is 3), minimum TTL (default is 1), maximum TTL (default is 30), and port number (default is 33434). To keep the default setting for those parameters, press the ENTER key.

Figure 21-6 is an example output of DNS using the traceroute command.

Figure 21-6. Traceroute command example

```
FTOS#traceroute www.force10networks.com

Translating "www.force10networks.com"...domain server (10.11.0.1) [OK]
Type Ctrl-C to abort.

------------------------------------------------------------------------------
Tracing the route to www.force10networks.com (10.11.84.18), 30 hops max, 40 byte packets
------------------------------------------------------------------------------
```

ARP

FTOS uses two forms of address resolution: ARP and Proxy ARP.

Address Resolution Protocol (ARP) runs over Ethernet and enables endstations to learn the MAC addresses of neighbors on an IP network. Over time, FTOS creates a forwarding table mapping the MAC addresses to their corresponding IP address. This table is called the ARP Cache and dynamically learned addresses are removed after a defined period of time.

For more information on ARP, see RFC 826, *An Ethernet Address Resolution Protocol*.

In FTOS, Proxy ARP enables hosts with knowledge of the network to accept and forward packets from hosts that contain no knowledge of the network. Proxy ARP makes it possible for hosts to be ignorant of the network, including subnetting.

For more information on Proxy ARP, refer to RFC 925, *Multi-LAN Address Resolution,* and RFC 1027, *Using ARP to Implement Transparent Subnet Gateways.*

Configuration Task List for ARP

The following list includes configuration tasks for ARP:

- Configure static ARP entries on page 471 (optional)
- Enable Proxy ARP on page 472 (optional)
- Clear ARP cache on page 472 (optional)
- ARP Learning via Gratuitous ARP on page 473
- ARP Learning via ARP Request on page 474
- Configurable ARP Retries on page 475

For a complete listing of all ARP-related commands, refer to the *FTOS Command Line Reference Guide*.

Configure static ARP entries

ARP dynamically maps the MAC and IP addresses, and while most network host support dynamic mapping, you can configure an ARP entry (called a static ARP) for the ARP cache.

To configure a static ARP entry, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
arp *ip-address mac-address interface*	CONFIGURATION	Configure an IP address and MAC address mapping for an interface. • *ip-address:* IP address in dotted decimal format (A.B.C.D). • *mac-address:* MAC address in nnnn.nnnn.nnnn format • *interface:* enter the interface type slot/port information.

These entries do not age and can only be removed manually. To remove a static ARP entry, use the **no arp** *ip-address* command syntax.

To view the static entries in the ARP cache, use the **show arp static** command (Figure 253) in the EXEC privilege mode.

Figure 21-7. show arp static Command Example

```
FTOS#show arp

Protocol    Address         Age(min)  Hardware Address    Interface  VLAN    CPU
--------------------------------------------------------------------------------
Internet    10.1.2.4             17   08:00:20:b7:bd:32   Ma 1/0      -      CP
FTOS#
```

Enable Proxy ARP

By default, Proxy ARP is enabled. To disable Proxy ARP, use **no proxy-arp** command in the interface mode.

To re-enable Proxy ARP, use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip proxy-arp	INTERFACE	Re-enable Proxy ARP.

To view if Proxy ARP is enabled on the interface, use the **show config** command in the INTERFACE mode. If it is not listed in the show config command output, it is enabled. Only nondefault information is displayed in the **show config** command output.

Clear ARP cache

To clear the ARP cache of dynamically learnt ARP information, use the following command in the EXEC privilege mode:

Command Syntax	Command Mode	Purpose
clear arp-cache [*interface* | *ip ip-address*] [**no-refresh**]	EXEC privilege	Clear the ARP caches for all interfaces or for a specific interface by entering the following information: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a port channel interface, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale. • For a SONET interface, enter the keyword **sonet** followed by the slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN interface, enter the keyword **vlan** followed by a number between 1 and 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS. **ip** *ip-address* (OPTIONAL) Enter the keyword ip followed by the IP address of the ARP entry you wish to clear. **no-refresh** (OPTIONAL) Enter the keyword **no-refresh** to delete the ARP entry from CAM. Or use this option with *interface* or **ip** *ip-address* to specify which dynamic ARP entries you want to delete. **Note:** Transit traffic may not be forwarded during the period when deleted ARP entries are resolved again and re-installed in CAM. Use this option with extreme caution.

ARP Learning via Gratuitous ARP

Gratuitous ARP can mean an ARP request or reply. In the context of ARP Learning via Gratuitous ARP on FTOS, the gratuitous ARP is a request. A Gratuitous ARP Request is an ARP request that is not needed according to the ARP specification, but one that hosts may send to:

- detect IP address conflicts
- inform switches of their presence on a port so that packets can be forwarded
- update the ARP table of other nodes on the network in case of an address change

In the request, the host uses its own IP address in the Sender Protocol Address and Target Protocol Address fields.

In FTOS versions prior to 8.3.1.0, if a gratuitous ARP is received some time after an ARP request is sent, only RP2 installs the ARP information. For example:

1. At time t=0 FTOS sends an ARP request for IP *A.B.C.D*
2. At time t=1 FTOS receives an ARP request for IP *A.B.C.D*
3. At time t=2 FTOS installs an ARP entry for *A.B.C.D* only on RP2.

Beginning with version 8.3.1.0, when a Gratuitous ARP is received, FTOS installs an ARP entry on all 3 CPUs.

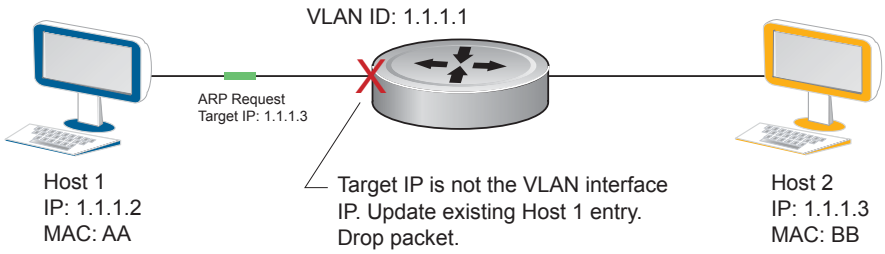
Task	Command Syntax	Command Mode
Enable ARP learning via gratuitous ARP.	**arp learn-enable**	CONFIGURATION

ARP Learning via ARP Request

In FTOS versions prior to 8.3.1.0, FTOS learns via ARP Requests only if the Target IP specified in the packet matches the IP address of the receiving router interface. This is the case when a host is attempting to resolve the gateway address.

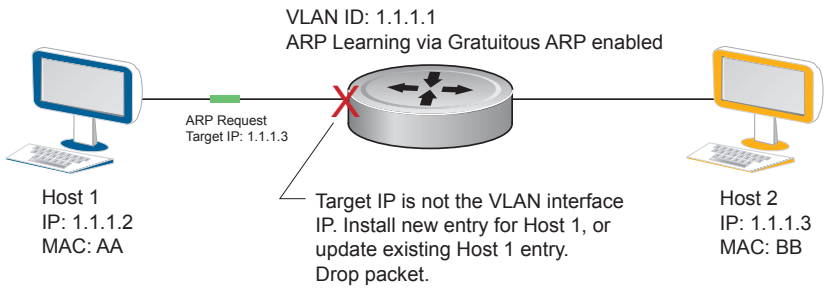
If the Target IP does not match the incoming interface, then the packet is dropped. If there is an existing entry for the requesting host, it is updated.

Figure 21-8. Learning via Gratuitous ARP



Beginning with FTOS version 8.3.1.0, when ARP Learning via Gratuitous ARP is enabled, the system installs a new ARP entry, or updates an existing entry for all received ARP requests.

Figure 21-9. Learning via Gratuitous ARP



Whether ARP Learning via Gratuitous ARP is is enabled or disabled, the system does not look up the Target IP. It only updates the ARP entry for the Layer 3 interface with the source IP of the request.

Configurable ARP Retries

In FTOS versions prior to 8.3.1.0 the number of ARP retries is set to 5 and is not configurable. After 5 retries, FTOS backs off for 20 seconds before it sends a new request. Beginning with FTOS version 8.3.1.0, the number of ARP retries is configurable. The backoff interval remains at 20 seconds.

Task	Command Syntax	Command Mode
Set the number of ARP retries.	**arp retries** *number* Default: 5 Range: 5-20	CONFIGURATION
Display all ARP entries learned via gratuitous ARP.	**show arp retries**	EXEC Privilege

FTOS Behavior: Due to ARP Pruning, the total number of ARP requests sent might exceed, but is never less than, the configured number of ARP retries. This occurs when the ARP Pruning timer expires while ARP retry is in progress.

ARP Pruning is a mechanism that clears stale entries every 1 minute. A stale entry is an IP address for which either the ARP expiry timer—which by default is set to 4 hours—expires, or ARP cannot resolve an IP address.

When there is a coincidence between an ARP Pruning timer expiry and the ARP retry mechanism FTOS sends more than the configured number of ARP requests. To illustrate why this occurs, take a time T1=0 seconds, at which point the pruning timer starts. At time T2=55 seconds, when the pruning timer is 55 seconds, suppose the ARP retry for an unresolved address begins, with ARP retry configured for 20 retries. By time T3=60 seconds, the total number of ARP requests sent is 5. However, at T3, the pruning timer expires and clears all stale entries, including the entry for which ARP retry is in progress. In this case, ARP retry starts over and sends another 20 ARP request over 20 seconds. As a result, the total number of ARP requests sent is 25, not the configured 20.

ICMP

For diagnostics, Internet Control Message Protocol (ICMP) provide routing information to end stations by choosing the best route (ICMP redirect messages) or determining if a router is reachable (ICMP Echo or Echo Reply). ICMP Error messages inform the router of problems in a particular packet. These messages are sent only on unicast traffic

Configuration Task List for ICMP

Use the following steps to configure ICMP:

- Enable ICMP unreachable messages on page 476
- Enable ICMP redirects on page 476

See the *FTOS Command Line Reference Guide* for a complete listing of all commands related to ICMP.

Enable ICMP unreachable messages

By default, ICMP unreachable messages are disabled. When enabled ICMP unreachable messages are created and sent out all interfaces. To disable ICMP unreachable messages, use the **no ip unreachable** command.

To reenable the creation of ICMP unreachable messages on the interface, use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip unreachable	INTERFACE	Set FTOS to create and send ICMP unreachable messages on the interface.

To view if ICMP unreachable messages are sent on the interface, use the **show config** command in the INTERFACE mode. If it is not listed in the **show config** command output, it is enabled. Only nondefault information is displayed in the **show config** command output.

Enable ICMP redirects

Enable ICMP redirects is supported on E platform

By default, ICMP redirect messages are disabled. When enabled, ICMP redirect messages are created and sent out all interfaces. To disable ICMP redirect messages, use the **no ip redirect** command.

To reenable the creation of ICMP redirect messages on the interface, use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip redirect	INTERFACE	Set FTOS to create and send ICMP redirect messages on the interface.

To view if ICMP redirect messages are sent on the interface, use the **show config** command in the INTERFACE mode. If it is not listed in the **show config** command output, it is enabled. Only nondefault information is displayed in the **show config** command output.

UDP Helper

UDP helper allows you to direct the forwarding IP/UDP broadcast traffic by creating special broadcast addresses and rewriting the destination IP address of packets to match those addresses. Configurations using this feature are described in the section Configurations Using UDP Helper on page 478.

Configuring UDP Helper

Configuring FTOS to direct UDP broadcast is a two-step process:

1. Enable UDP helper and specify the UDP ports for which traffic is forwarded. See Enabling UDP Helper on page 477.
2. Configure a broadcast address on interfaces that will receive UDP broadcast traffic. See Configuring a Broadcast Address on page 478.

Important Points to Remember about UDP Helper

- The existing command **ip directed broadcast** is rendered meaningless if UDP helper is enabled on the same interface.
- The broadcast traffic rate should not exceed 200 packets per second when UDP helper is enabled.
- You may specify a maximum of 16 UDP ports.
- UDP helper is compatible with IP helper (**ip helper-address**):
  - UDP broadcast traffic with port number 67 or 68 are unicast to the DHCP server per the **ip helper-address** configuration whether or not the UDP port list contains those ports.
  - If the UDP port list contains ports 67 or 68, UDP broadcast traffic forwarded on those ports.

Enabling UDP Helper

Enable UPD helper using the command **ip udp-helper udp-ports**, as shown in Figure 21-10.

Figure 21-10. Enabling UDP Helper

```
FTOS(conf-if-gi-1/1)#ip udp-helper udp-port 1000
FTOS(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
 ip address 2.1.1.1/24
 ip udp-helper udp-port 1000
 no shutdown
```

View the interfaces and ports on which UDP helper is enabled using the command **show ip udp-helper** from EXEC Privilege mode, as shown in Figure 21-11.

Figure 21-11. Viewing the UDP Broadcast Configuration

```
FTOS#show ip udp-helper
-------------------------------------------------
Port           UDP port list
-------------------------------------------------
Gi 1/1         1000
```

Configuring a Broadcast Address

Configure a broadcast address on an interface using the command **ip udp-broadcast-address**, as shown in Figure 21-12.

Figure 21-12. Configuring a Broadcast Address

```
FTOS(conf-if-vl-100)#ip udp-broadcast-address 1.1.255.255
FTOS(conf-if-vl-100)#show config
!
interface Vlan 100
 ip address 1.1.0.1/24
 ip udp-broadcast-address 1.1.255.255
 untagged GigabitEthernet 1/2
 no shutdown
```

View the configured broadcast address for an interface using the command **show interfaces**, as shown in Figure 21-13.

Figure 21-13. Configuring a Broadcast Address

```
R1_E600(conf)#do show interfaces vlan 100
Vlan 100 is up, line protocol is down
Address is 00:01:e8:0d:b9:7a, Current address is 00:01:e8:0d:b9:7a
Interface index is 1107787876
Internet address is 1.1.0.1/24
IP UDP-Broadcast address is 1.1.255.255
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed auto
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:07:44
Queueing strategy: fifo
Input Statistics:
     0 packets, 0 bytes
Time since last interface status change: 00:07:44
```

Configurations Using UDP Helper

When UDP helper is enabled and the destination IP address of an incoming packet is a broadcast address, FTOS suppresses the destination address of the packet. The following sections describe various configurations that employ UDP helper to direct broadcasts.

- UDP Helper with Broadcast-all Addresses on page 479
- UDP Helper with Subnet Broadcast Addresses on page 479
- UDP Helper with Configured Broadcast Addresses on page 480
- UDP Helper with No Configured Broadcast Addresses on page 481

UDP Helper with Broadcast-all Addresses

When the destination IP address of an incoming packet is the IP broadcast address, FTOS rewrites the address to match the configured broadcast address.

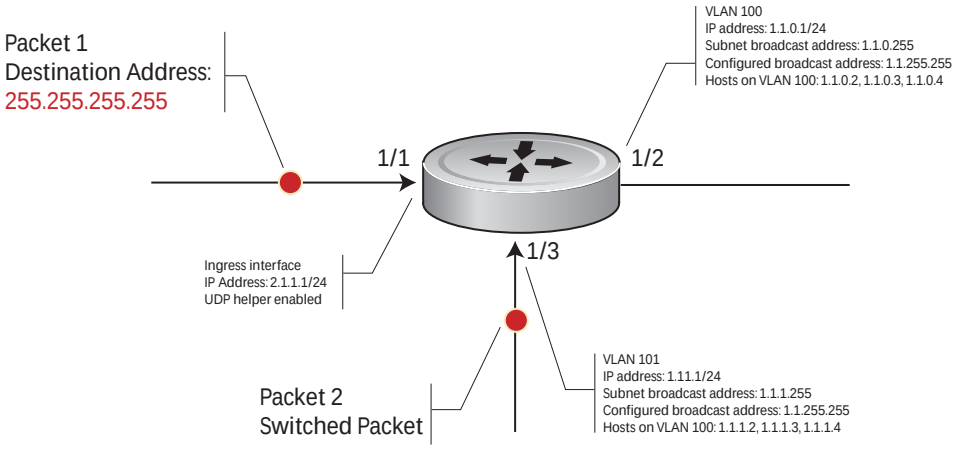
In Figure 21-14:

1. Packet 1 is dropped at ingress if no UDP helper address is configured.
2. If UDP helper (using the command **ip udp-helper udp-port**) is enabled, and the UDP destination port of the packet matches the UDP port configured, the system changes the destination address to the configured broadcast 1.1.255.255 and routes the packet to VLANs 100 and 101. If an IP broadcast address is not configured (using the command **ip udp-broadcast-address**) on VLANs 100 or 101, the packet is forwarded using the original destination destination IP address 255.255.255.255.

Packet 2, sent from a host on VLAN 101 has a broadcast MAC address and IP address. In this case:

1. It is flooded on VLAN 101 without changing the destination address because the forwarding process is Layer 2.
2. If UDP helper is enabled, the system changes the destination IP address to the configured broadcast address 1.1.255.255 and forwards the packet to VLAN 100.
3. Packet 2 is also forwarded to the ingress interface with an unchanged destination address because it does not have broadcast address configured.

Figure 21-14. UDP helper with All Broadcast Addresses



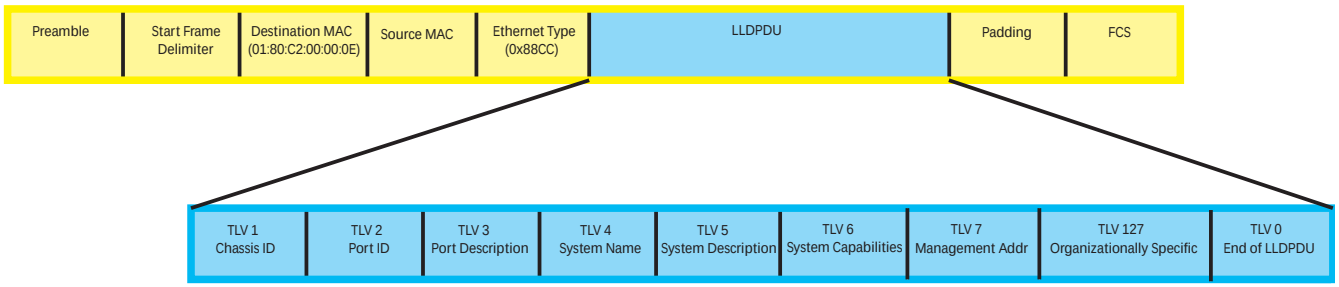
UDP Helper with Subnet Broadcast Addresses

When the destination IP address of an incoming packet matches the subnet broadcast address of any interface, the system changes the address to the configured broadcast address and sends it to matching interface.

In Figure 21-15, Packet 1 has the destination IP address 1.1.1.255, which matches the subnet broadcast address of VLAN 101. If UDP helper is configured and the packet matches the specified UDP port, then the system changes the address to the configured IP broadcast address and floods the packet on VLAN 101.

Packet 2 is sent from host on VLAN 101. It has a broadcast MAC address and a destination IP address of 1.1.1.255. In this case, it is flooded on VLAN 101 in its original condition as the forwarding process is Layer 2.

Figure 21-15. UDP helper with Subnet Broadcast Addresses



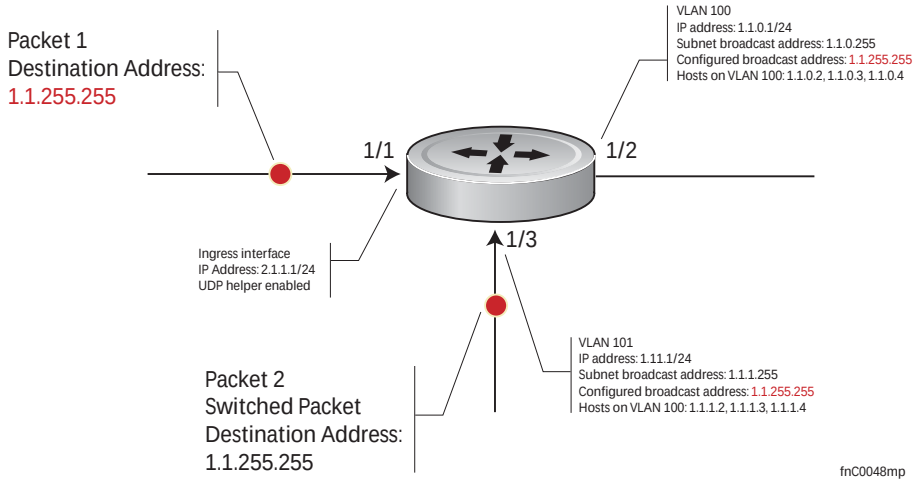
UDP Helper with Configured Broadcast Addresses

Incoming packets with a destination IP address matching the configured broadcast address of any interface are forwarded to the matching interfaces.

In Figure 21-16, Packet 1 has a destination IP address that matches the configured broadcast address of VLAN 100 and 101. If UDP helper is enabled and the UDP port number matches, the packet is flooded on both VLANs with an unchanged destination address.

Packet 2 is sent from a host on VLAN 101. It has broadcast MAC address and a destination IP address that matches the configured broadcast address on VLAN 101. In this case, Packet 2 is flooded on VLAN 101 with the destination address unchanged because the forwarding process is Layer 2. If UDP helper is enabled, the packet is flooded on VLAN 100 as well.

Figure 21-16. UDP Helper with Configured Broadcast Addresses



UDP Helper with No Configured Broadcast Addresses

- If the incoming packet has a broadcast destination IP address, then the unaltered packet is routed to all Layer 3 interfaces.
- If the Incoming packet has a destination IP address that matches the subnet broadcast address of any interface, then the unaltered packet is routed to the matching interfaces.

Troubleshooting UDP Helper

Display debugging information using the command **debug ip udp-helper**, as shown in Figure 21-17.

Figure 21-17. Debugging UDP Broadcast

```
FTOS(conf)# debug ip udp-helper
01:20:22: Pkt rcvd on Gi 5/0 with IP DA (0xffffffff) will be sent on Gi 5/1 Gi 5/2 Vlan 3
01:44:54: Pkt rcvd on Gi 7/0 is handed over for DHCP processing.
```

Use the command **debug ip dhcp** when using the IP helper and UDP helper on the same interface, as shown in Figure 21-18.

Figure 21-18. Debugging IP Helper with UDP Helper

```
Packet 0.0.0.0:68 -> 255.255.255.255:67 TTL 128

2005-11-05 11:59:35 %RELAY-I-PACKET, BOOTP REQUEST (Unicast) received at interface
172.21.50.193 BOOTP Request, XID = 0x9265f901, secs = 0 hwaddr = 00:02:2D:8D:46:DC, giaddr =
0.0.0.0, hops = 2

2005-11-05 11:59:35 %RELAY-I-BOOTREQUEST, Forwarded BOOTREQUEST for 00:02:2D:8D:46:DC to
137.138.17.6

2005-11-05 11:59:36 %RELAY-I-PACKET, BOOTP REPLY (Unicast) received at interface
194.12.129.98 BOOTP Reply, XID = 0x9265f901, secs = 0 hwaddr = 00:02:2D:8D:46:DC, giaddr =
172.21.50.193, hops = 2

2005-07-05 11:59:36 %RELAY-I-BOOTREPLY, Forwarded BOOTREPLY for 00:02:2D:8D:46:DC to
128.141.128.90 Packet 0.0.0.0:68 -> 255.255.255.255:67 TTL 128
```

22

IPv6 Addressing

IPv6 Addressing is supported on platforms: C E S

Note: The basic IPv6 commands are supported on all platforms. However, not all IPv6-based features are supported on all platforms and on all releases. Refer to Table 22-2 to see which FTOS version supports an IPv6 feature on each platform.

IPv6 (Internet Protocol Version 6) is the successor to IPv4. Due to the extremely rapid growth in internet users, and IP addresses, IPv4 is reaching its maximum usage. IPv6 will eventually replace IPv4 usage to allow for the constant expansion.

This chapter provides a brief discussion of the differences between IPv4 and IPv6, and the Dell Force10 support of IPv6. This chapter discusses the following, but is not intended to be a comprehensive discussion of IPv6.

Protocol Overview

IPv6 is an evolution of IPv4. IPv6 is generally installed as an upgrade in devices and operating systems. Most new devices and operating systems support both IPv4 and IPv6.

Some key changes in IPv6 are:

- Extended Address Space
- Stateless Autoconfiguration
- Header Format Simplification
- Improved Support for Options and Extensions

Extended Address Space

The address format is extended from 32 bits to 128 bits. This not only provides room for all anticipated needs, it allows for the use of a hierarchical address space structure to optimize global addressing.

Stateless Autoconfiguration

When a booting device comes up in IPv6 and asks for its network prefix, the device can get the prefix (or prefixes) from an IPv6 router on its link. It can then autoconfigure one or more global IP addresses by using either the MAC address or a private random number to build its unique IP address.

Stateless auto-configuration uses three mechanisms for IPv6 address configuration:

- Prefix Advertisement - Routers use "Router Advertisement" messages to announce the Network Prefix. Hosts then use their interface-identifier MAC address to generate their own valid IPv6 address.
- Duplicate Address Detection (DAD) - Before configuring its IPv6 address, an IPv6 host node device checks whether that address is used anywhere on the network using this mechanism.
- Prefix Renumbering - Useful in transparent renumbering of hosts in the network when an organization changes its service provider.

Note: As an alternative to stateless auto-configuration, network hosts can obtain their IPv6 addresses using Dynamic Host Control Protocol (DHCP) servers via stateful auto-configuration.

Note: FTOS provides the flexibility to add prefixes to advertise responses to RS messages. By default the RA response messages are not sent when an RS message is received. Enable the RA response messages with the **ipv6 nd prefix default** command in INTERFACE mode.

FTOS manipulation of IPv6 stateless auto-configuration supports the router side only. Neighbor Discovery (ND) messages are advertised so the neighbor can use this information to auto-configure its address. However, received Neighbor Discovery (ND) messages are not used to create an IPv6 address.

The router redistribution functionality in Neighbor Discovery Protocol (NDP) is similar to IPv4 router redirect messages. Neighbor Discovery Protocol (NDP) uses ICMPv6 redirect messages (Type 137) to inform nodes that a better router exists on the link.

IPv6 Headers

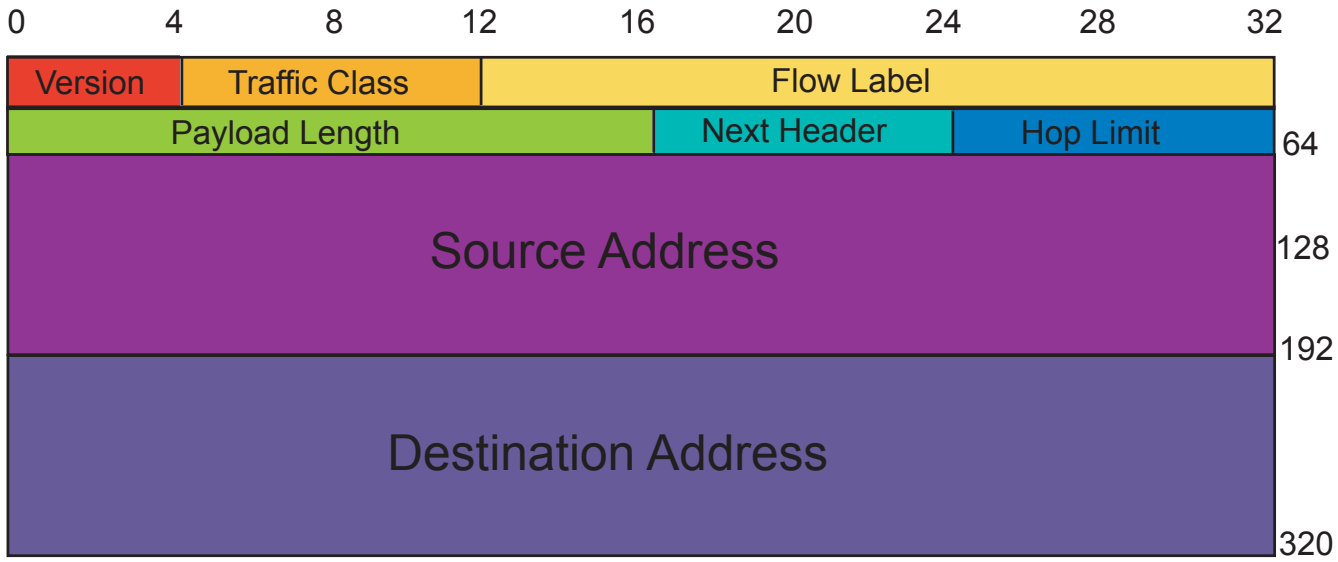
The IPv6 header has a fixed length of 40 bytes. This provides 16 bytes each for Source and Destination information, and 8 bytes for general header information. The IPv6 header includes the following fields:

- Version (4 bits)
- Traffic Class (8 bits)
- Flow Label (20 bits)
- Payload Length (16 bits)
- Next Header (8 bits)
- Hop Limit (8 bits)
- Source Address (128 bits)
- Destination Address (128 bits)

IPv6 provides for Extension Headers. Extension Headers are used only if necessary. There can be no extension headers, one extension header or more than one extension header in an IPv6 packet. Extension Headers are defined in the Next Header field of the preceding IPv6 header. IPv6 header fields

The 40 bytes of the IPv6 header are ordered as show in Figure 22-1.

Figure 22-1. IPv6 Header Fields



Version (4 bits)

The Version field always contains the number 6, referring to the packet's IP version.

Traffic Class (8 bits)

The Traffic Class field deals with any data that needs special handling. These bits define the packet priority and are defined by the packet Source. Sending and forwarding routers use this field to identify different IPv6 classes and priorities. Routers understand the priority settings and handle them appropriately during conditions of congestion.

Flow Label (20 bits)

The Flow Label field identifies packets requiring special treatment in order to manage real-time data traffic. The sending router can label sequences of IPv6 packets so that forwarding routers can process packets within the same flow without needing to reprocess each packet's head separately.

Note: All packets in the flow must have the same source and destination addresses.

Payload Length (16 bits)

The Payload Length field specifies the packet payload. This is the length of the data *following* the IPv6 header. IPv6 Payload Length only includes the data following the header, not the header itself.

The Payload Length limit of 2 bytes requires that the maximum packet payload be 64 KB. However, the Jumbogram option type Extension header supports larger packet sizes when required.

Next Header (8 bits)

The Next Header field identifies the next header's type. If an Extension header is used, this field contains the type of Extension header (Table 22-1). If the next header is a TCP or UDP header, the value in this field is the same as for IPv4. The Extension header is located between the IP header and the TCP or UDP header.

Table 22-1. Next Header field values

Value	Description
0	Hop-by Hop option header following
4	IPv4
6	TCP
8	Exterior Gateway Protocol (EGP)
41	IPv6
43	Routing header
44	Fragmentation header
50	Encrypted Security
51	Authentication header

Table 22-1. Next Header field values (continued)

Value	Description
59	No Next Header
60	Destinations option header

Note: This is not a comprehensive table of Next Header field values. Refer to the Internet Assigned Numbers Authority (IANA) web page http://www.iana.org/assignments/protocol-numbers for a complete and current listing.

Hop Limit (8 bits)

The Hop Limit field shows the number of hops remaining for packet processing. In IPv4, this is known as the Time to Live (TTL) field and uses seconds rather than hops.

Each time the packet moves through a forwarding router, this field decrements by 1. If a router receives a packet with a Hop Limit of 1, it decrements it to 0 (zero). The router discards the packet and sends an ICMPv6 message back to the sending router indicating that the Hop Limit was exceeded in transit.

Source Address (128 bits)

The Source Address field contains the IP address for the packet originator.

Destination Address (128 bits)

The Destination Address field contains the intended recipient's IP address. This can be either the ultimate destination or the address of the next hop router.

Extension Header fields

Extension headers are used only when necessary. Due to the streamlined nature of the IPv6 header, adding extension headers do not severely impact performance. Each Extension headers's lengths vary, but they are always a multiple of 8 bytes.

Each extension header is identified by the Next Header field in the IPv6 header that precedes it. Extension headers are viewed only by the destination router identified in the Destination Address field. If the Destination Address is a multicast address, the Extension headers are examined by all the routers in that multicast group.

However, if the Destination Address is a Hop-by-Hop options header, the Extension header is examined bye every forwarding router along the packet's route. The Hop-by-Hop options header must immediately follow the IPv6 header, and is noted by the value 0 (zero) in the Next Header field (Table 22-1).

Extension headers are processed in the order in which they appear in the packet header.

Hop-by-Hop Options header

The Hop-by-Hop options header contains information that is examined by every router along the packet's path. It follows the IPv6 header and is designated by the Next Header value 0 (zero) (Table 22-1).

When a Hop-by-Hop Options header is not included, the router knows that it does not have to process any router specific information and immediately processes the packet to its final destination.

When a Hop-by-Hop Options header is present, the router only needs this extension header and does not need to take the time to view further into the packet.

The Hop-by-Hop Options header contains:

- Next Header (1 byte)

  This field identifies the type of header following the Hop-by-Hop Options header and uses the same values shown in Table 22-1.
- Header Extension Length (1 byte)

  This field identifies the length of the Hop-by-Hop Options header in 8-byte units, but does not include the first 8 bytes. Consequently, if the header is less than 8 bytes, the value is 0 (zero).
- Options (size varies)

  This field can contain 1 or more options. The first byte if the field identifies the Option type, and directs the router how to handle the option.

00	Skip and continue processing
01	Discard the packet.
10	Discard the packet and send an ICMP Parameter Problem Code 2 message to the packet's Source IP Address identifying the unknown option type
11	Discard the packet and send an ICMP Parameter Problem, Code 2 message to the packet's Source IP Address only if the Destination IP Address is not a multicast address.

  The second byte contains the Option Data Length.
  The third byte specifies whether the information can change en route to the destination. The value is 1 if it can change; the value is 0 if it cannot change.

Addressing

IPv6 addresses are normally written as eight groups of four hexadecimal digits, where each group is separated by a colon (:). For example, 2001:0db8:0000:0000:0000:0000:1428:57ab is a valid IPv6 address. If one or more four-digit group(s) is 0000, the zeros may be omitted and replaced with two colons(::). For example, 2001:0db8:0000:0000:0000:0000:1428:57ab can be shortened to 2001:0db8::1428:57ab. Only one set of double colons is supported in a single address. Any number of consecutive 0000 groups may be reduced to two colons, as long as there is *only one double colon used in an address*. Leading zeros in a group can also be omitted (as in ::1 for localhost).

All the addresses in the following list are all valid and equivalent.

- 2001:0db8:0000:0000:0000:0000:1428:57ab
- 2001:0db8:0000:0000:0000::1428:57ab
- 2001:0db8:0:0:0:0:1428:57ab
- 2001:0db8:0:0::1428:57ab
- 2001:0db8::1428:57ab
- 2001:db8::1428:57ab

IPv6 networks are written using Classless Inter-Domain Routing (CIDR) notation. An IPv6 network (or subnet) is a contiguous group of IPv6 addresses the size of which must be a power of two; the initial bits of addresses, which are identical for all hosts in the network, are called the network's prefix.

A network is denoted by the first address in the network and the size in bits of the prefix (in decimal), separated with a slash. Since a single host is seen as a network with a 128-bit prefix, host addresses may be written with a following /128.

For example, 2001:0db8:1234::/48 stands for the network with addresses 2001:0db8:1234:0000:0000:0000:0000:0000 through 2001:0db8:1234:ffff:ffff:ffff:ffff:ffff

Link-local Addresses

Link-local addresses, starting with **fe80:**, are assigned only in the local link area. The addresses are generated usually automatically by the operating system's IP layer for each network interface. This provides instant automatic network connectivity for any IPv6 host and means that if several hosts connect to a common hub or switch, they have an instant communication path via their link-local IPv6 address. .

Link-local addresses cannot be routed to the public Internet.

Static and Dynamic Addressing

Static IP addresses are manually assigned to a computer by an administrator. Dynamic IP addresses are assigned either randomly or by a server using Dynamic Host Configuration Protocol (DHCP). Even though IP addresses assigned using DHCP may stay the same for long periods of time, they can change. In some cases, a network administrator may implement dynamically assigned static IP addresses. In this case, a DHCP server is used, but it is specifically configured to always assign the same IP address to a particular computer, and never to assign that IP address to another computer. This allows static IP addresses to be configured in one place, without having to specifically configure each computer on the network in a different way.

In IPv6, every interface, whether using static or dynamic address assignments, also receives a local-link address automatically in the fe80::/64 subnet.

Implementing IPv6 with FTOS

FTOS supports both IPv4 and IPv6, and both may be used simultaneously in your system.

Note: Dell Force10 recommends that you use FTOS version 7.6.1.0 or later when implementing IPv6 functionality on an E-Series system.

Table 22-2 lists the FTOS Version in which an IPv6 feature became available for each platform. The sections following the table give some greater detail about the feature. Specific platform support for each feature or functionality is designated by the C E S symbols.

Note: When both E-Series TeraScale and ExaScale are supported, only the E symbol is shown. If a feature is supported by one or the other chassis, the specific symbols are shown: e T for E-Series TeraScale or E X for E-Series ExaScale.

Table 22-2. FTOS and IPv6 Feature Support

Feature and/or Functionality	FTOS Release Introduction				Documentation and Chapter Location
	E-Series TeraScale	E-Series ExaScale	C-Series	S-Series	
Basic IPv6 Commands	7.4.1	8.2.1	7.8.1	7.8.1	IPv6 Basic Commands in the *FTOS Command Line Interface Reference Guide*
IPv6 Basic Addressing					
IPv6 address types: Unicast	7.4.1	8.2.1	7.8.1	7.8.1	Extended Address Space in this chapter
IPv6 neighbor discovery	7.4.1	8.2.1	7.8.1	7.8.1	IPv6 Neighbor Discovery in this chapter
IPv6 stateless autoconfiguration	7.4.1	8.2.1	7.8.1	7.8.1	Stateless Autoconfiguration in this chapter
IPv6 MTU path discovery	7.4.1	8.2.1	7.8.1	7.8.1	Path MTU Discovery in this chapter
IPv6 ICMPv6	7.4.1	8.2.1	7.8.1	7.8.1	ICMPv6 in this chapter
IPv6 ping	7.4.1	8.2.1	7.8.1	7.8.1	ICMPv6 in this chapter
IPv6 traceroute	7.4.1	8.2.1	7.8.1	7.8.1	ICMPv6 in this chapter
IPv6 Routing					
Static routing	7.4.1	8.2.1	7.8.1	7.8.1	Assign a Static IPv6 Route in this chapter
Route redistribution	7.4.1	8.2.1	7.8.1	8.4.2	OSPF, IS-IS, and IPv6 BGP chapters in the *FTOS Command Line Reference Guide*
Multiprotocol BGP extensions for IPv6	7.4.1	8.2.1	7.8.1	8.4.2	IPv6 BGP in the *FTOS Command Line Reference Guide*
IPv6 BGP MD5 Authentication	8.2.1.0	8.2.1.0	8.2.1.0	8.4.2	IPv6 BGP in the *FTOS Command Line Reference Guide*

Table 22-2. FTOS and IPv6 Feature Support (continued)

IS-IS for IPv6	7.5.1	8.2.1	8.4.2	8.4.2	Chapter 23, "Intermediate System to Intermediate System," on page 507 in the *FTOS Configuration Guide* IPv6 IS-IS in the *FTOS Command Line Reference Guide*
IS-IS for IPv6 support for redistribution	7.6.1	8.2.1	8.4.2	8.4.2	Chapter 23, "Intermediate System to Intermediate System," on page 507 in the *FTOS Configuration Guide* IPv6 IS-IS in the *FTOS Command Line Reference Guide*
ISIS for IPv6 support for distribute lists and administrative distance	7.6.1	8.2.1	8.4.2	8.4.2	Chapter 23, "Intermediate System to Intermediate System," on page 507 in the *FTOS Configuration Guide* IPv6 IS-IS in the *FTOS Command Line Reference Guide*
OSPF for IPv6 (OSPFv3)	7.4.1	8.2.1	7.8.1	8.4.2	OSPFv3 in the *FTOS Command Line Reference Guide*
Equal Cost Multipath for IPv6	7.4.1	8.2.1	7.8.1	7.8.1	
IPv6 Services and Management					
Telnet client over IPv6 (outbound Telnet)	7.5.1	8.2.1	7.8.1	7.8.1	Telnet with IPv6 in this chapter Control and Monitoring in the *FTOS Command Line Reference Guide*
Telnet server over IPv6 (inbound Telnet)	7.4.1	8.2.1	7.8.1	7.8.1	Telnet with IPv6 in this chapter Control and Monitoring in the *FTOS Command Line Reference Guide*
Secure Shell (SSH) client support over IPv6 (outbound SSH) Layer 3 only	7.5.1	8.2.1	7.8.1	7.8.1	SSH over an IPv6 Transport in this chapter
Secure Shell (SSH) server support over IPv6 (inbound SSH) Layer 3 only	7.4.1	8.2.1	7.8.1	7.8.1	SSH over an IPv6 Transport in this chapter
IPv6 Access Control Lists	7.4.1	8.2.1	7.8.1	8.2.1.0	IPv6 Access Control Lists in the *FTOS Command Line Reference Guide*
IPv6 Multicast					
PIM-SM for IPv6	7.4.1	8.2.1	8.4.2	8.4.2	IPv6 Multicast in this chapter; IPv6 PIM in the *FTOS Command Line Reference Guide*

Table 22-2. FTOS and IPv6 Feature Support (continued)

PIM-SSM for IPv6	7.5.1	8.2.1	8.4.2	8.4.2	IPv6 Multicast in this chapter IPv6 PIM in the *FTOS Command Line Reference Guide*
MLDv1/v2	7.4.1	8.2.1	8.4.2	8.4.2	IPv6 Multicast in this chapter Multicast IPv6 in the *FTOS Command Line Reference Guide*
MLDv1 Snooping	7.4.1	8.2.1	8.4.2	8.4.2	IPv6 Multicast in this chapter Multicast IPv6 in the *FTOS Command Line Reference Guide*
MLDv2 Snooping	8.3.1.0	8.3.1.0	8.4.2	8.4.2	IPv6 Multicast in this chapter Multicast IPv6 in the *FTOS Command Line Reference Guide*
IPv6 QoS					
trust DSCP values	7.4.1	8.2.1	8.4.2	8.4.2	QoS for IPv6 in this chapter

ICMPv6

ICMPv6 is supported on platforms C E S

ICMP for IPv6 combines the roles of ICMP, IGMP and ARP in IPv4. Like IPv4, it provides functions for reporting delivery and forwarding errors, and provides a simple echo service for troubleshooting. The FTOS implementation of ICMPv6 is based on RFC 2463.

Generally, ICMPv6 uses two message types:

- Error reporting messages indicate when the forwarding or delivery of the packet failed at the destination or intermediate node. These messages include Destination Unreachable, Packet Too Big, Time Exceeded and Parameter Problem messages.
- Informational messages provide diagnostic functions and additional host functions, such as Neighbor Discovery and Multicast Listener Discovery. These messages also include Echo Request and Echo Reply messages.

The FTOS **ping** and **traceroute** commands extend to support IPv6 addresses. These commands use ICMPv6 Type-2 messages.

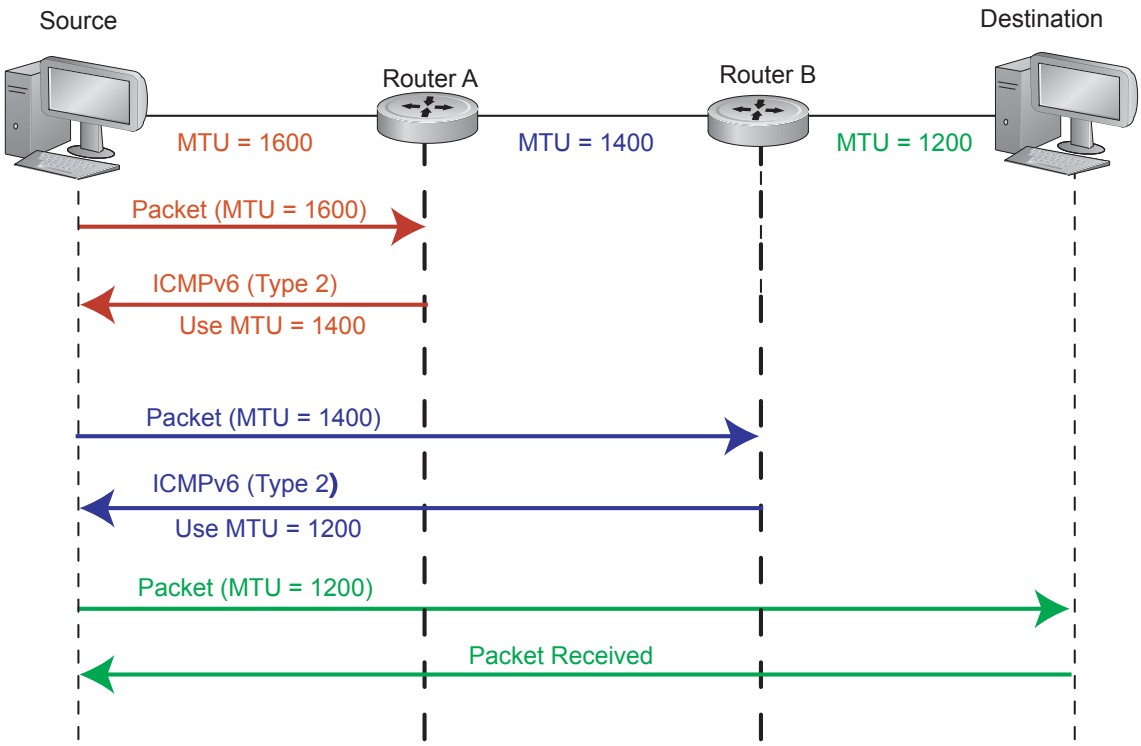
Path MTU Discovery

IPv6 MTU Discovery is supported on platforms C E S

Path MTU (Maximum Transmission Unit) defines the largest packet size that can traverse a transmission path without suffering fragmentation. Path MTU for IPv6 uses ICMPv6 Type-2 messages to discover the largest MTU along the path from source to destination and avoid the need to fragment the packet.

The recommended MTU for IPv6 is 1280. Greater MTU settings increase processing efficiency because each packet carries more data while protocol overheads (headers, for example) or underlying per-packet delays remain fixed.

Figure 22-2. MTU Discovery Path



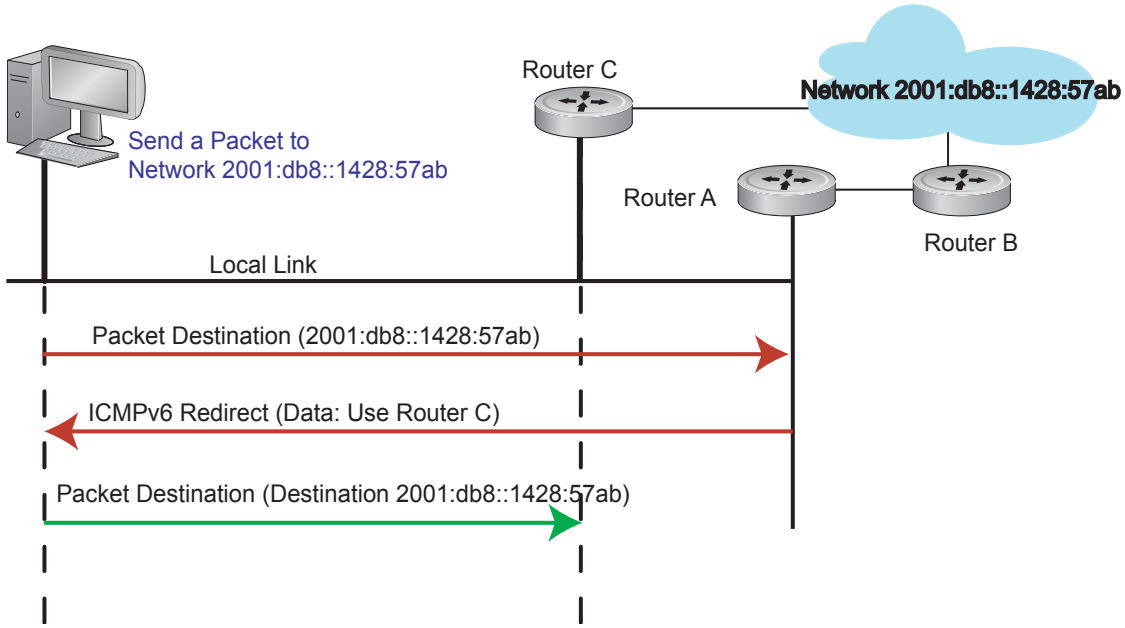
IPv6 Neighbor Discovery

IPv6 NDP is supported on platforms C E S

Neighbor Discovery Protocol (NDP) is a top-level protocol for neighbor discovery on an IPv6 network. In lieu of ARP, NDP uses "Neighbor Solicitation" and "Neighbor Advertisement" ICMPv6 messages for determining relationships between neighboring nodes. Using these messages, an IPv6 device learns the link-layer addresses for neighbors known to reside on attached links, quickly purging cached values that become invalid.

With ARP, each node broadcasts ARP requests on the entire link. This approach causes unnecessary processing by uninterested nodes. With NDP, each node sends a request only to the intended destination via a multicast address with the unicast address used as the last 24 bits. Other hosts on the link do not participate in the process, greatly increasing network bandwidth efficiency.

Figure 22-3. NDP Router Redistribution



IPv6 Neighbor Discovery of MTU packets

With FTOS 8.3.1.0, you can set the MTU advertised through the RA packets to incoming routers, without altering the actual MTU setting on the interface. The **ip nd mtu** command sets the value advertised to routers. It does not set the actual MTU rate. For example, if **ip nd mtu** is set to 1280, the interface will still pass 1500-byte packets, if that is what is set with the **mtu** command.

Advertise Neighbor Prefixes

Specify which IPv6 prefixes are include in Neighbor Advertisements. By default, all prefixes configured as addresses on the interface are advertised. You can control the advertise parameters per prefix; the **default** keyword can be used to use the default parameters for all prefixes.

Command Syntax	Command Mode	Purpose
ipv6 nd prefix {*ipv6-address/prefix-length*> | **default**} [**no-advertise**] | [**no-autoconfig**] [**no-rtr-address**] [**off-link**] [**lifetime** {*valid* | **infinite**} {*preferred* | **infinite**}]	INTERFACE	Specify which IPv6 prefixes are include in Neighbor Advertisements.

QoS for IPv6

IPv6 QoS is supported on platforms: C E S

FTOS IPv6 supports quality of service based on DSCP field. You can configure FTOS to honor the DSCP value on incoming routed traffic and forward the packets with the same value.

Refer to Chapter 41, Quality of Service for details. Refer also to the Honor DSCP values on ingress packets in the QoS chapter for information relating to the **trust diffserv** command.

IPv6 Multicast

IPv6 Multicast is supported on platforms: C E S

FTOS supports the following protocols to implement IPv6 multicast routing:

- Multicast Listener Discovery Protocol (MLD). MLD on a multicast router sends out periodic general MLD queries that the switch forwards through all ports in the VLAN. There are two versions of MLD: MLD version 1 is based on version 2 of the Internet Group Management Protocol (IGMP) for IPv4, and MLD version 2 is based on version 3 of the IGMP for IPv4. IPv6 multicast for FTOS supports versions 1 and 2
- PIM-SM. Protocol-Independent Multicast-Sparse Mode (PIM-SM) is a multicast protocol in which multicast receivers explicitly join to receive multicast traffic. The protocol uses a router as the root or Rendezvous Point (RP) of the share tree distribution tree to distribute multicast traffic to a multicast group. Messages to join the multicast group (Join messages) are sent towards the RP and data is sent from senders to the RP so receivers can discover who are the senders and begin receiving traffic destined to the multicast group.
- PIM in Source Specific Multicast (PIM-SSM). PIM-SSM protocol is based on the source specific model for forwarding Multicast traffic across multiple domains on the Internet. It is restricted to shortest path trees (SPTs) to specific sources described by hosts using MLD. PIM-SSM is essentially a subset of PIM-SM protocol, which has the capability to join SPTs. The only difference being register states and shared tree states for Multicast groups in SSM range are not maintained. End-hosts use MLD to register their interest in a particular source-group (S,G) pair. PIM-SSM protocol interacts with MLD to construct the multicast forwarding tree rooted at the source S.

Refer to *FTOS Command Line Interface Reference* document chapters Multicast IPv6, and Protocol Independent Multicast (IPv6) for configuration details.

SSH over an IPv6 Transport

IPv6 SSH is supported on platforms C E S

FTOS supports both inbound and outbound SSH sessions using IPv6 addressing. Inbound SSH supports accessing the system through the management interface as well as through a physical Layer 3 interface.

Refer to the **Security Commands** chapter in the *FTOS Command Line Interface Reference* document for SSH configuration details.

Configuration Task List for IPv6

This section contains information regarding the following:

- Change your CAM-Profile on an E-Series system (mandatory)
- Adjust your CAM-Profile on an C-Series or S-Series
- Assign an IPv6 Address to an Interface
- Assign a Static IPv6 Route
- Telnet with IPv6
- SNMP over IPv6
- Show IPv6 Information
- Clear IPv6 Routes

Change your CAM-Profile on an E-Series system

The **cam-profile** command is supported only on platform E

Change your CAM profile to the CAM **ipv6-extacl** before doing any further IPv6 configuration. Once the CAM profile is changed, save the configuration and reboot your router.

Command Syntax	Command Mode	Purpose
cam-profile ipv6-extacl microcode ipv6-extacl *chassis* | *linecard slot*	EXEC Privileged	Enable the CAM profile with IPv6 extended ACLs on the entire chassis or on a specific linecard *chassis* changes the CAM profile for all linecards in the chassis *linecard slot/port* changes the CAM profile only for the specified slot

Figure 22-4 displays the IPv6 CAM profile summary for a chassis that already has IPv6 CAM profile configured. Figure 22-5 shows the full IPv6 CAM profiles. Refer to Chapter 11, Content Addressable Memory, on page 281 for complete information regarding CAM configuration.

Figure 22-4. Command Example: show cam-profile summary (E-Series)

```
FTOS#show cam-profile summary

-- Chassis CAM Profile --
                 : Current Settings : Next Boot
Profile Name     : IPV6-ExtACL      : IPV6-ExtACL
MicroCode Name   : IPv6-ExtACL      : IPv6-ExtACL

-- Line card 1 --
                 : Current Settings : Next Boot
Profile Name     : IPV6-ExtACL      : IPV6-ExtACL
MicroCode Name   : IPv6-ExtACL      : IPv6-ExtACL

FTOS#
```

Figure 22-5. Command Example: show cam profile (E-Series)

```
FTOS#show cam-profile

-- Chassis CAM Profile --

CamSize          : 18-Meg
                 : Current Settings : Next Boot
Profile Name     : IPV6-ExtACL      : IPV6-ExtACL
L2FIB            : 32K entries      : 32K entries
L2ACL            : 1K entries       : 1K entries
IPv4FIB          : 192K entries     : 192K entries
IPv4ACL          : 12K entries      : 12K entries
IPv4Flow         : 8K entries       : 8K entries
EgL2ACL          : 1K entries       : 1K entries
EgIPv4ACL        : 1K entries       : 1K entries
Reserved         : 2K entries       : 2K entries
IPv6FIB          : 6K entries       : 6K entries
IPv6ACL          : 3K entries       : 3K entries
IPv6Flow         : 4K entries       : 4K entries
EgIPv6ACL        : 1K entries       : 1K entries
MicroCode Name   : IPv6-ExtACL      : IPv6-ExtACL

-- Line card 1 --
CamSize          : 18-Meg
                 : Current Settings : Next Boot
--More--
```

Adjust your CAM-Profile on an C-Series or S-Series

The **cam-acl** command is supported on platforms C S

If you plan to implement IPv6 ACLs, you **must** adjust your CAM settings.

The CAM space is allotted in FP blocks. The total space allocated must equal 13 FP blocks. Note that there are 16 FP blocks, but the System Flow requires 3 blocks that cannot be reallocated.

The **ipv6acl** allocation must be entered as a factor of 2 (2, 4, 6, 8, 10). All other profile allocations can use either even or odd numbered ranges.

The **default** option sets the CAM Profile as follows:

- L3 ACL (ipv4acl): 6
- L2 ACL(l2acl) : 5
- IPv6 L3 ACL (ipv6acl): 0
- L3 QoS (ipv4qos): 1
- L2 QoS (l2qos): 1

Save the new CAM settings to the startup-config (**write-mem or copy run start**) then reload the system for the new settings to take effect.

Command Syntax	Command Mode	Purpose
cam-acl {**default** | **l2acl** *number* **ipv4acl** *number* **ipv6acl** *number*, **ipv4qos** *number* **l2qos** *number*, **l2pt** *number* **ipmacacl** *number* **ecfmacl** *number* [**vman-qos** | **vman-dual-qos** *number*}	CONFIGURATION	Allocate space for IPV6 ACLs. Enter the CAM profile name followed by the amount to be allotted. When not selecting the default option, you must enter all of the profiles listed and a range for each. The total space allocated must equal 13. The **ipv6acl** range must be a factor of 2.
show cam-acl	EXEC EXEC Privilege	Show the current CAM settings.

Assign an IPv6 Address to an Interface

IPv6 Addresses are supported on platforms C E S

Essentially IPv6 is enabled in FTOS simply by assigning IPv6 addresses to individual router interfaces. IPv6 and IPv4 can be used together on a system, but be sure to differentiate that usage carefully. Use the **ipv6 address** command to assign an IPv6 address to an interface.

Command Syntax	Command Mode	Purpose
ipv6 address *ipv6 address/ mask*	CONFIG-INTERFACE	Enter the IPv6 Address for the device. *ipv6 address* : x:x:x:x::x *mask* : prefix length 0-128
	IPv6 addresses are normally written as eight groups of four hexadecimal digits, where each group is separated by a colon (:). Omitting zeros is accepted as described in Addressing earlier in this chapter.	

Assign a Static IPv6 Route

IPv6 Static Routes are supported on platforms C E S

Use the **ipv6 route** command to configure IPv6 static routes.

Command Syntax	Command Mode	Purpose
ipv6 route *prefix type {slot/ port} forwarding router tag*	CONFIGURATION	Set up IPv6 static routes *prefix*: IPv6 route prefix *type {slot/port}:* interface type and slot/port *forwarding router:* forwarding router's address *tag:* route tag Enter the keyword **interface** followed by the type of interface and slot/port information: • For a 10/100/1000 Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/ port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a loopback interface, enter the keyword **loopback** followed by the loopback number • For a linecard interface, enter the keyword **linecard** followed by the slot number • For a port-channel interface, enter the keyword **port-channel** followed by the port-channel number • For a VLAN interface, enter the keyword **vlan** followed by the VLAN ID • For a Null interface, enter the keyword **null** followed by the Null interface number

Telnet with IPv6

IPv6 Telnet is supported on platforms C E S

The Telnet client and server in FTOS support IPv6 connections. You can establish a Telnet session directly to the router using an IPv6 Telnet client, or an IPv6 Telnet connection can be initiated from the router.

Note: Telnet to link local addresses is not supported.

Command Syntax	Command Mode	Purpose
telnet *ipv6 address*	EXEC *or* EXEC Privileged	Enter the IPv6 Address for the device. *ipv6 address* : x:x:x:x::x *mask* : prefix length 0-128
	IPv6 addresses are normally written as eight groups of four hexadecimal digits, where each group is separated by a colon (:). Omitting zeros is accepted as described in Addressing earlier in this chapter.	

SNMP over IPv6

SNMP is supported on platforms C E S

Simple Network Management Protocol (SNMP) can be configured over IPv6 transport so that an IPv6 host can perform SNMP queries and receive SNMP notifications from a device running FTOS IPv6. The FTOS SNMP-server commands for IPv6 have been extended to support IPv6. Refer to the *SNMP and SYSLOG* chapter in the *FTOS Command Line Interface Reference* for more information regarding SNMP commands.

- snmp-server host
- snmp-server user ipv6
- snmp-server community ipv6
- snmp-server community access-list-name ipv6
- snmp-server group ipv6
- snmp-server group access-list-name ipv6

Show IPv6 Information

All of the following show commands are supported on platforms: C E S

View specific IPv6 configuration with the following commands.

Command Syntax	Command Mode	Purpose
show ipv6 ?	EXEC *or* EXEC Privileged	List the IPv6 show options

Command Syntax	Command Mode	Purpose

```
FTOS#show ipv6 ?
accounting      IPv6 accounting information
cam linecard    IPv6 CAM Entries for Line Card
fib linecard    IPv6 FIB Entries for Line Card
interface       IPv6 interface information
mbgproutes      MBGP routing table
mld             MLD information
mroute          IPv6 multicast-routing table
neighbors       IPv6 neighbor information
ospf            OSPF information
pim             PIM V6 information
prefix-list     List IPv6 prefix lists
route           IPv6 routing information
rpf             RPF table
FTOS#
```

Show an IPv6 Interface

View the IPv6 configuration for a specific interface with the following command.

Command Syntax	Command Mode	Purpose
show ipv6 interface *type {slot/port}*	EXEC	Show the currently running configuration for the specified interface Enter the keyword **interface** followed by the type of interface and slot/port information: • For all brief summary of IPv6 status and configuration , enter the keyword **brief.** • For all IPv6 configured interfaces, enter the keyword **configured.** • For a 10/100/1000 Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a loopback interface, enter the keyword **loopback** followed by the loopback number • For a linecard interface, enter the keyword **linecard** followed by the slot number • For a port-channel interface, enter the keyword **port-channel** followed by the port-channel number • For a VLAN interface, enter the keyword **vlan** followed by the VLAN ID

Figure 22-6 illustrates the **show ipv6 interface** command output.

Figure 22-6. Command Example: show ipv6 interface

```
FTOS#show ipv6 interface gi 2/2
GigabitEthernet 2/2 is down, line protocol is down
  IPV6 is enabled
  Link Local address: fe80::201:e8ff:fe06:95a3
  Global Unicast address(es):
    3:4:5:6::8, subnet is 3::/24
  Global Anycast address(es):
  Joined Group address(es):
    ff02::1
    ff02::2
    ff02::1:ff00:8
    ff02::1:ff06:95a3
  MTU is 1500
  ICMP redirects are not sent
  DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30 seconds
  ND advertised reachable time is 30 seconds
  ND advertised retransmit interval is 30 seconds
  ND router advertisements are sent every 200 seconds
  ND router advertisements live for 1800 seconds
```

Show IPv6 Routes

View the global IPv6 routing information with the following command.

Command Syntax	Command Mode	Purpose
show ipv6 route *type*	EXEC	Show IPv6 routing information for the specified route type. Enter the keyword: • To display information about a network, enter the **ipv6 address** (X:X:X:X::X). • To display information about a host, enter the **hostname.** • To display information about all IPv6 routes (including non-active routes), enter **all**. • To display information about all connected IPv6 routes, enter **connected**. • To display information about brief summary of all IPv6 routes, enter **summary**. • To display information about Border Gateway Protocol (BGP) routes, enter **bgp**. • To display information about ISO IS-IS routes, enter **isis**. • To display information about Open Shortest Path First (OSPF) routes, enter **ospf**. • To display information about Routing Information Protocol (RIP), enter **rip**. • To display information about static IPv6 routes, enter **static**. • To display information about an IPv6 Prefix lists, enter **list** and the prefix-list name.

Figure 22-7 illustrates the **show ipv6 route** command output.

Figure 22-7. Command Example: show ipv6 route

```
FTOS#show ipv6 route

Codes: C - connected, L - local, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
Gateway of last resort is not set

       Destination  Dist/Metric, Gateway, Last Change
       ------------------------------------------------------
  C    2001::/64 [0/0]
        Direct, Gi 1/1, 00:28:49
```

Figure 22-8 illustrates the **show ipv6 route** summary command output.

Figure 22-8. Command Example: show ipv6 route summary

```
FTOS#show ipv6 route summary

Route Source              Active Routes   Non-active Routes
connected                 5               0
static                    0               0
```

Figure 22-9 illustrates the **show ipv6 route static** command output.

Figure 22-9. Command Example: show ipv6 route static

```
FTOS#show ipv6 route static
Destination Dist/Metric, Gateway, Last Change
-----------------------------------------------------
       S      8888:9999:5555:6666:1111:2222::/96 [1/0]
                     via    2222:2222:3333:3333::1, Gi 9/1, 00:03:16
       S      9999:9999:9999:9999::/64 [1/0]
                     via 8888:9999:5555:6666:1111:2222:3333:4444, 00:03:16
```

Show the Running-Configuration for an Interface

View the configuration for any interface with the following command.

Command Syntax	Command Mode	Purpose
show running-config interface *type {slot/port}*	EXEC	Show the currently running configuration for the specified interface Enter the keyword **interface** followed by the type of interface and slot/port information: • For a 10/100/1000 Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For the Management interface on the RPM, enter the keyword **ManagementEthernet** followed by the slot/port information. • For a 10 Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information.

Figure 22-10 illustrates the **show running-config** command output. Note the IPv6 address listed.

Figure 22-10. Command Example: show running-config interface

```
FTOS#show run int gi 2/2
!
interface GigabitEthernet 2/2
 no ip address
 ipv6 address 3:4:5:6::8/24
 shutdown
FTOS#
```

Clear IPv6 Routes

Use the clear IPv6 route command to clear routes from the IPv6 routing table.

Command Syntax	Command Mode	Purpose
clear ipv6 route {***** | *ipv6 address prefix-length*}	EXEC	Clear (refresh) all or a specific routes from the IPv6 routing table. * : all routes *ipv6 address* : x:x:x:x::x *mask* : prefix length 0-128

Command Syntax	Command Mode	Purpose
		IPv6 addresses are normally written as eight groups of four hexadecimal digits, where each group is separated by a colon (:). Omitting zeros is accepted as described in Addressing earlier in this chapter.

23

Intermediate System to Intermediate System

Intermediate System to Intermediate System is supported on platform: E

Intermediate System to Intermediate System (IS-IS) protocol is an interior gateway protocol (IGP) that uses a shortest-path-first algorithm. Dell Force10 supports both IPv4 and IPv6 versions of IS-IS, as described in this chapter.

IS-IS protocol standards are listed in the Appendix 63, Standards Compliance chapter.

Protocol Overview

The intermediate-system-to-intermediate-system (IS-IS) protocol, developed by the International Organization for Standardization (ISO), is an interior gateway protocol (IGP) that uses a shortest-path-first algorithm.

Note: This protocol supports routers passing both IP and OSI traffic, though the Dell Force10 implementation supports only IP traffic.

IS-IS is organized hierarchally into routing domains, and each router or system resides in at least one area. In IS-IS, routers are designated as Level 1, Level 2 or Level 1-2 systems. Level 1 routers only route traffic within an area, while Level 2 routers route traffic between areas. At its most basic, Level 1 systems route traffic within the area and any traffic destined for outside the area is sent to a Level 1-2 system. Level 2 systems manage destination paths for external routers. Only Level 2 routers can exchange data packets or

routing information directly with external routers located outside of the routing domains. Level 1-2 systems manage both inter-area and intra-area traffic by maintaining two separate link databases; one for Level 1 routes and one for Level 2 routes. A Level 1-2 router does not advertise Level 2 routes to a Level 1 router.

To establish adjacencies, each IS-IS router sends different Protocol Data Units (PDU). For IP traffic, the IP addressing information is included in the IS-IS hello PDUs and the Link State PDUs (LSPs).

This brief overview is not intended to provide a complete understanding of IS-IS; for that, consult the documents listed in Multi-Topology IS-IS on page 509.

IS-IS Addressing

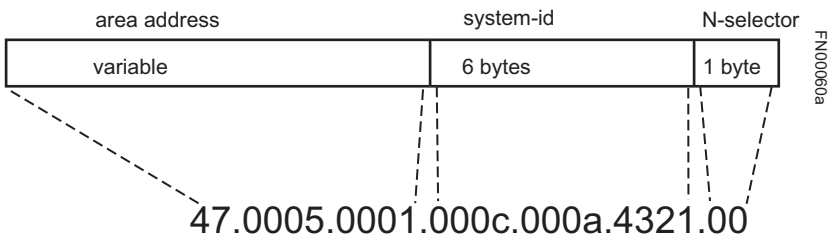
IS-IS PDUs require ISO-style addressing called Network Entity Title (NET). For those familiar with NSAP addresses, the composition of the NET is identical to an NSAP address, except the last byte is always 0. The NET is composed of IS-IS area address, system ID, and the N-selector. The last byte is the N-selector. All routers within an area have the same area portion. Level 1 routers route based on the system address portion of the address, while the Level 2 routers route based on the area address.

The NET length is variable, with a maximum of 20 bytes and a minimum of 8 bytes. It is composed of the following:

- area address. Within your routing domain or area, each area must have a unique area value. The first byte is called the authority and format indicator (AFI).
- system address. This is usually the router's MAC address.
- N-selector. This is always 0.

Figure 23-1 is an example of the ISO-style address to illustrate the address format used by IS-IS. In this example, the first five bytes (47.0005.0001) are the area address. The system portion is 000c.000a.4321 and the last byte is always 0.

Figure 23-1. ISO Address Format



Multi-Topology IS-IS

FTOS 7.8.1.0 and later support Multi-Topology Routing IS-IS.

E-Series ExaScale platform (E)(X) supports Multi-Topology IS-IS with FTOS 8.2.1.0 and later.

Multi-Topology IS-IS (MT IS-IS) allows you to create multiple IS-IS topologies on a single router with separate databases. This feature is used to place a virtual physical topology into logical routing domains, which can each support different routing and security policies.

All routers on a LAN or point-to-point must have at least one common supported topology when operating in Multi-Topology IS-IS mode. If IPv4 is the common supported topology between those two routers, adjacency can be formed. All topologies must share the same set of L1-L2 boundaries.

You must implement a **wide metric-style** globally on the Autonomous System to run Multi-Topology IS-IS for IPv6 because the TLVs used to advertise IPv6 information in link-state packets (LSPs) are defined to use only extended metrics.

The Multi-Topology ID is shown in the first octet of the IS-IS packet. Certain MT topologies are assigned to serve predetermined purposes:

- MT ID #0: Equivalent to the "standard" topology.
- MT ID #1: Reserved for IPv4 in-band management purposes.
- MT ID #2: Reserved for IPv6 routing topology.
- MT ID #3: Reserved for IPv4 multicast routing topology.
- MT ID #4: Reserved for IPv6 multicast routing topology.
- MT ID #5: Reserved for IPv6 in-band management purposes.

Transition Mode

All routers in the area or domain must use the same type of IPv6 support, either single-topology or multi-topology. A router operating in multi-topology mode will not recognize the ability of the single-topology mode router to support IPv6 traffic, which will lead to holes in the IPv6 topology.

While in transition mode, both types of TLVs (single-topology and multi-topology) are sent in LSPs for all configured IPv6 addresses, but the router continues to operate in single-topology mode (that is, the topological restrictions of the single-topology mode remain in effect). Transition mode stops after all routers in the area or domain have been upgraded to support multi-topology IPv6. Once all routers in the area or domain are operating in multi-topology IPv6 mode, the topological restrictions of single-topology mode are no longer in effect.

Interface support

MT IS-IS is supported on physical Ethernet interfaces, physical Sonet interfaces, port-channel interfaces (static & dynamic using LACP), and VLAN interfaces.

Adjacencies

Adjacencies on point-to-point interfaces are formed as usual, where IS-IS routers do not implement Multi-Topology (MT) extensions. If a local router does not participate in certain MTs, it will not advertise those MT IDs in its IIHs and so will not include that neighbor within its LSPs. If an MT ID is not detected in the remote side's IIHs, the local router does not include that neighbor within its LSPs. The local router will not form an adjacency if both routers don't have at least one common MT over the interface.

Graceful Restart

Graceful Restart is supported on E platforms for both Helper and Restart modes.

Graceful Restart is a protocol-based mechanism that preserves the forwarding table of the restarting router and its neighbors for a specified period to minimize the loss of packets. A graceful-restart router does not immediately assume that a neighbor is permanently down and so does not trigger a topology change.

Normally, when an IS-IS router is restarted, temporary disruption of routing occurs due to events in both the restarting router and the neighbors of the restarting router. When a router goes down without a Graceful Restart, there is a potential to lose access to parts of the network due to the necessity of network topology changes.

IS-IS Graceful Restart recognizes the fact that in a modern router, the control plane and data plane are functionality separate. Restarting the control plane functionality (such as the failover of the active RPM to the backup in a redundant configuration) should not necessarily interrupt data packet forwarding. This behavior is supported because the forwarding tables previously computed by an active RPM have been downloaded into the Forwarding Information Base on the line cards (the data plane) and are still resident. For packets that have existing FIB/CAM entries, forwarding between ingress and egress ports can continue uninterrupted while the control plane IS-IS process comes back to full functionality and rebuilds its routing tables.

A new TLV (the Restart TLV) is introduced in the IIH PDUs, indicating that the router supports Graceful Restart.

Timers

Three timers are used to support IS-IS Graceful Restart functionality. Once Graceful Restart is enabled, these timers manage the the Graceful Restart process.

- The T1 timer specifies the wait time before unacknowledged restart requests are generated. This is the interval before the system sends a Restart Request (an IIH with RR bit set in Restart TLV) until the CSNP is received from the helping router. The duration can be set to a specific amount of time (seconds) or a number of attempts.
- The T2 timer is the maximum time that the system will wait for LSP database synchronization. This timer applies to the database type (level-1, level-2 or both).

- The T3 timer sets the overall wait time after which the router determines that it has failed to achieve database synchronization (by setting the overload bit in its own LSP). This timer can be based on adjacency settings with the value derived from adjacent routers that are engaged in graceful restart recovery (the minimum of all the Remaining Time values advertised by the neighbors) or by setting a specific amount of time manually.

Implementation Information

IS-IS implementation supports one instance of IS-IS and six areas. The system can be configured as a Level 1 router, a Level 2 router, or a Level 1-2 router. For IPv6, the IPv4 implementation has been expanded to include two new type-length-values (TLV) in the protocol data unit (PDU) that carry information required for IPv6 routing. The new TLVs are *IPv6 Reachability* and *IPv6 Interface Address*. Also, a new IPv6 protocol identifier has also been included in the supported TLVs. The new TLVs use the extended metrics and up/down bit semantics.

Multi-Topology IS-IS adds TLVs:

- The *Multi-Topology TLV* contains one or more Multi-Topology IDs in which the router participates. This TLV is included in IIH and the first fragment of an LSP.
- The *MT Intermediate Systems TLV* appears for every topology a node supports. An MT ID is added to the extended IS reachability TLV type 22.
- The *Multi-Topology Reachable IPv4 Prefixes TLV* appears for each IPv4 announced by an IS for a given MT ID. Its structure is aligned with the extended IS Reachability TLV Type 236 and it adds an MT ID.
- The *Multi-Topology Reachable IPv6 Prefixes TLV* appears for each IPv6 announced by an IS for a given MT ID. Its structure is aligned with the extended IS Reachability TLV Type 236 and add a MT ID.

By default, FTOS supports dynamic hostname exchange to assist with troubleshooting and configuration. By assigning a name to an IS-IS NET address, you can track IS-IS information on that address easier. FTOS does not support ISO CLNS routing; however, the ISO NET format is supported for addressing.

To support IPv6, the Dell Force10 implementation of IS-IS performs the following tasks:

- Advertise IPv6 information in the PDUs
- Process IPv6 information received in the PDUs
- Compute routes to IPv6 destinations
- Download IPv6 routes to RTM for installing in the FIB
- Accept external IPv6 information and advertise this information in the PDUs

Table 23-1 displays the default values for IS-IS.

Table 23-1. IS-IS Default Values

IS-IS Parameter	Default Value
Complete Sequence Number PDU (CSNP) interval	10 seconds
IS-to-IS hello PDU interval	10 seconds
IS-IS interface metric	10
Metric style	Narrow
Designated Router priority	64
Circuit Type	Level 1 and Level 2
IS Type	Level 1 and Level 2
Equal Cost Multi Paths	16

Configuration Information

To use IS-IS, you must configure and enable IS-IS in two or three modes: CONFIGURATION ROUTER ISIS, CONFIGURATION INTERFACE, and (when configuring for IPv6) ADDRESS-FAMILY mode. Commands in ROUTER ISIS mode configure IS-IS globally, while commands executed in INTERFACE mode enable and configure IS-IS features on that interface only. Commands in the ADDRESS-FAMILY mode are specific to IPv6.

Note that by using the IS-IS routing protocol to exchange IPv6 routing information and to determine destination reachability, you can route IPv6 along with IPv4 while using a single intra-domain routing protocol. The configuration commands allow you to enable and disable IPv6 routing and to configure or remove IPv6 prefixes on links.

Except where identified, the commands discussed in this chapter apply to both IPv4 and IPv6 versions of IS-IS.

Configuration Task List for IS-IS

The following list includes the configuration tasks for IS-IS:

Enable IS-IS

By default, IS-IS is not enabled.

The system supports one instance of IS-IS. To enable IS-IS globally, create an IS-IS routing process and assign a NET address. To exchange protocol information with neighbors, enable IS-IS on an interface, instead of on a network as with other routing protocols.

In IS-IS, neighbors form adjacencies only when they are same IS type. For example, a Level 1 router never forms an adjacency with a Level 2 router. A Level 1-2 router will form Level 1 adjacencies with a neighboring Level 1 router and will form Level 2 adjacencies with a neighboring Level 2 router.

Note: Even though you enable IS-IS globally, you must enable the IS-IS process on an interface for the IS-IS process to exchange protocol information and form adjacencies.

Use these commands in the following sequence to configure IS-IS globally.

Step	Task	Command Syntax	Command Mode
1	Create an IS-IS routing process. • *tag* is optional and identifies the name of the IS-IS process.	**router isis** [*tag*]	CONFIGURATION
2	Configure an IS-IS network entity title (NET) for a routing process. Specify the area address and system ID for an IS-IS routing process. The last byte must be 00. Refer to IS-IS Addressing for more information on configuring a NET.	**net** *network-entity-title*	ROUTER ISIS

Step	Task	Command Syntax	Command Mode
3	Enter the interface configuration mode. Enter the keyword **interface** followed by the type of interface and slot/port information: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For the Loopback interface on the RPM, enter the keyword **loopback** followed by a number from 0 to 16383. • For a port channel, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale. • For a SONET interface, enter the keyword **sonet** followed by slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.	**interface** *interface*	CONFIGURATION
4	Enter an IPv4 Address. Assign an IP address and mask to the interface. The IP address must be on the same subnet as other IS-IS neighbors, but the IP address does not need to relate to the NET address.	**ip address** *ip-address mask*	INTERFACE
5	Enter an IPv6 Address. *ipv6 address* : x:x:x:x::x *mask* : prefix length 0-128 The IPv6 address must be on the same subnet as other IS-IS neighbors, but the IP address does not need to relate to the NET address.	**ipv6 address** *ip-address mask*	INTERFACE
6	Enable IS-IS on the interface. If you configure a *tag* variable, it must be the same as the *tag* variable assigned in step 1.	{**ip | ipv6**} **router isis** [*tag*]	INTERFACE

The default IS type is level-1-2. To change the IS type to Level 1 only or Level 2 only, use the **is-type** command in ROUTER ISIS mode.

Enter the **show isis protocol** command in EXEC Privilege mode or the **show config** command in ROUTER ISIS mode to view the IS-IS configuration.

Figure 23-2. Command Example: show isis protocol

```
FTOS#show isis protocol
IS-IS Router: <Null Tag>
  System Id: EEEE.EEEE.EEEE  IS-Type: level-1-2
  Manual area address(es):
   47.0004.004d.0001
  Routing for area address(es):
   21.2223.2425.2627.2829.3031.3233
   47.0004.004d.0001
  Interfaces supported by IS-IS:
   Vlan 2
   GigabitEthernet 4/22
   Loopback 0
  Redistributing:
  Distance: 115
  Generate narrow metrics: level-1-2
  Accept narrow metrics:   level-1-2
  Generate wide metrics:   none
  Accept wide metrics:     none
FTOS#
```

Use the **show isis traffic** command in EXEC Privilege mode to view IS-IS protocol statistics.

Figure 23-3. Command Example: show isis traffic

```
FTOS#show isis traffic
 IS-IS: Level-1 Hellos (sent/rcvd) : 4272/1538
 IS-IS: Level-2 Hellos (sent/rcvd) : 4272/1538
 IS-IS: PTP Hellos (sent/rcvd)     : 0/0
 IS-IS: Level-1 LSPs sourced (new/refresh) : 0/0
 IS-IS: Level-2 LSPs sourced (new/refresh) : 0/0
 IS-IS: Level-1 LSPs flooded (sent/rcvd) : 32/19
 IS-IS: Level-2 LSPs flooded (sent/rcvd) : 32/17
 IS-IS: Level-1 LSPs CSNPs (sent/rcvd) : 1538/0
 IS-IS: Level-2 LSPs CSNPs (sent/rcvd) : 1534/0
 IS-IS: Level-1 LSPs PSNPs (sent/rcvd) : 0/0
 IS-IS: Level-2 LSPs PSNPs (sent/rcvd) : 0/0
 IS-IS: Level-1 DR Elections : 2
 IS-IS: Level-2 DR Elections : 2
 IS-IS: Level-1 SPF Calculations : 29
 IS-IS: Level-2 SPF Calculations : 29
 IS-IS: LSP checksum errors received : 0
 IS-IS: LSP authentication failures : 0
FTOS#
```

You can assign additional NET addresses, but the System ID portion of the NET address must remain the same. FTOS supports up to six area addresses.

Some address considerations are:

- In order to be neighbors, Level 1 routers must be configured with at least one common area address.
- A Level 2 router becomes a neighbor with another Level 2 router regardless of the area address configured. However, if the area addresses are different, the link between the Level 2 routers is only at Level 2.

Configure Multi-Topology IS-IS (MT IS-IS)

Step	Task	Command Syntax	Command Mode
1	Enable Multi-Topology IS-IS for IPv6. Enter the **transition** keyword to allow an IS-IS IPv6 user to continue to use single-topology mode while upgrading to multi-topology mode.After every router has been configured with the transition keyword, and all the routers are in MT IS-IS IPv6 mode users can remove the transition keyword on each router.	**multi-topology** [*transition*]	ROUTER ISIS AF IPV6
	Note: When transition mode is not enabled, you will not have IPv6 connectivity between routers operating in single-topology mode and routers operating in multi-topology mode.		
2	Excluded this router from other router's SPF calculations.	**set-overload-bit**	ROUTER ISIS AF IPV6
3	Set the minimum interval between SPF calculations.	**spf-interval [level-l | level-2 | interval] [initial_wait_interval [second_wait_interval]]**	ROUTER ISIS AF IPV6
	This command is used for IPv6 route computation *only* when multi-topology is enabled. If using single-topology mode, use the **spf-interval** command in CONFIG ROUTER ISIS mode to apply to both IPv4 and IPv6 route computations.		
4	Implement a **wide metric-style** globally. To configure wide or wide transition metric style, the cost can be a between 0 and 16,777,215.	**isis ipv6 metric metric-value [level-1 | level-2 | level-1-2]**	ROUTER ISIS AF IPV6

Configure Multi-Topology IS-IS (MT IS-IS)

Step	Task	Command Syntax	Command Mode
1	Enable Multi-Topology IS-IS for IPv6. Enter the **transition** keyword to allow an IS-IS IPv6 user to continue to use single-topology mode while upgrading to multi-topology mode.After every router has been configured with the transition keyword, and all the routers are in MT IS-IS IPv6 mode users can remove the transition keyword on each router.	**multi-topology** [*transition*]	ROUTER ISIS AF IPV6
	Note: When transition mode is not enabled, you will not have IPv6 connectivity between routers operating in single-topology mode and routers operating in multi-topology mode.		
2	Excluded this router from other router's SPF calculations.	**set-overload-bit**	ROUTER ISIS AF IPV6
3	Set the minimum interval between SPF calculations.	**spf-interval [level-l | level-2 | interval] [initial_wait_interval [second_wait_interval]]**	ROUTER ISIS AF IPV6
	This command is used for IPv6 route computation *only* when multi-topology is enabled. If using single-topology mode, use the **spf-interval** command in CONFIG ROUTER ISIS mode to apply to both IPv4 and IPv6 route computations.		
4	Implement a **wide metric-style** globally. To configure wide or wide transition metric style, the cost can be a between 0 and 16,777,215.	**isis ipv6 metric metric-value [level-1 | level-2 | level-1-2]**	ROUTER ISIS AF IPV6

Configure IS-IS Graceful Restart

To enable IS-IS Graceful Restart globally, use the following command in ROUTER-ISIS mode. Additional, optional commands can be implemented to enable the Graceful Restart settings.

Command Syntax	Command Mode	Purpose
graceful-restart ietf	ROUTER-ISIS	Enable Graceful Restart on ISIS processes
graceful-restart interval *minutes*	ROUTER-ISIS	Configure the period of time during which the Graceful Restart attempt will be prevented. Range: 1-120 minutes Default: 5 minutes

Command Syntax	Command Mode	Purpose
graceful-restart restart- wait *seconds*	ROUTER-ISIS	Enable the Graceful Restart maximum wait time before a restarting peer comes up. Be sure to set the **t3** timer to adjacency on the restarting router when implementing this command. Range: 5-120 seconds Default: 30 seconds
graceful-restart t1 {interval *seconds* **| retry-times** *value***}**	ROUTER-ISIS	Configure the time that the Graceful Restart timer T1 defines for a restarting router to use for each interface, as an interval before regenerating Restart Request (an IIH with RR bit set in Restart TLV) after waiting for an acknowledgement. **interval**: wait time (Range: 5-120, default: 5) **retry-times**: number of times an unacknowledged restart request will be sent before the restarting router gives up the graceful restart engagement with the neighbor. (Range: 1-10 attempts, default: 1)
graceful-restart t2 {level-1 | level-2} *seconds*	ROUTER-ISIS	Configure the time for Graceful Restart timer T2 that a restarting router will use as the wait time for each database to synchronize. **level-1, level-2**: identifies the database instance type to which the wait interval applies. Range:5-120 seconds Default: 30 seconds
graceful-restart t3 {adjacency | manual *seconds***}**	ROUTER-ISIS	Configure Graceful Restart timer T3 to set the time used by the restarting router as an overall maximum time to wait for database synchronization to complete. **adjacency**: the restarting router receives the remaining time value from its peer and adjusts its T3 value accordingly if user has configured configured this option. **manual**: allows you to specify a fixed value that the restarting router should use. Range: 50-120 seconds Default: 30 seconds
		Note: If this timer expires before the synchronization has completed, the restarting router sends the overload bit in the LSP. The 'overload' bit is an indication to the receiving router that database synchronization did not complete at the restarting router.

Use the **show isis graceful-restart detail** command in EXEC Privilege mode to view all Graceful Restart related configuration.

Figure 23-4. Command Example: show isis graceful-restart detail

```
FTOS#show isis graceful-restart detail
Configured Timer Value
======================
Graceful Restart         : Enabled
Interval/Blackout time   : 1 min
T3 Timer                 : Manual
T3 Timeout Value         : 30
T2 Timeout Value         : 30 (level-1), 30 (level-2)
T1 Timeout Value         : 5, retry count: 1
Adjacency wait time      : 30

Operational Timer Value
=======================
Current Mode/State       : Normal/RUNNING
T3 Time left             : 0
T2 Time left             : 0 (level-1), 0 (level-2)
Restart ACK rcv count    : 0 (level-1), 0 (level-2)
Restart Req rcv count    : 0 (level-1), 0 (level-2)
Suppress Adj rcv count   : 0 (level-1), 0 (level-2)
Restart CSNP rcv count   : 0 (level-1), 0 (level-2)
Database Sync count      : 0 (level-1), 0 (level-2)

Circuit GigabitEthernet 2/10:
  Mode: Normal L1-State:NORMAL, L2-State: NORMAL

  L1: Send/Receive: RR:0/0, RA: 0/0, SA:0/0
      T1 time left: 0, retry count left:0

  L2: Send/Receive: RR:0/0, RA: 0/0, SA:0/0
      T1 time left: 0, retry count left:0
FTOS#
```

Use the **show isis interface** command in EXEC Privilege mode to view all interfaces configured with IS-IS routing along with the defaults.

Figure 23-5. Command Example: show isis interface

```
show isis interface G1/34
GigabitEthernet 2/10 is up, line protocol is up
  MTU 1497, Encapsulation SAP
  Routing Protocol: IS-IS
    Circuit Type: Level-1-2
    Interface Index 0x62cc03a, Local circuit ID 1
    Level-1 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000B.01
            Hello Interval: 10, Hello Multiplier: 3, CSNP Interval: 10
    Number of active level-1 adjacencies: 1
    Level-2 Metric: 10, Priority: 64, Circuit ID: 0000.0000.000B.01
            Hello Interval: 10, Hello Multiplier: 3, CSNP Interval: 10
    Number of active level-2 adjacencies: 1
    Next IS-IS LAN Level-1 Hello in 4 seconds
    Next IS-IS LAN Level-2 Hello in 6 seconds
    LSP Interval: 33    Next IS-IS LAN Level-1 Hello in 4 seconds
    Next IS-IS LAN Level-2 Hello in 6 seconds
    LSP Interval: 33
Restart Capable Neighbors: 2, In Start: 0, In Restart: 0
FTOS#
```

Change LSP attributes

IS-IS routers flood Link state PDUs (LSPs) to exchange routing information. LSP attributes include the generation interval, maximum transmission unit (MTU) or size, and the refresh interval. You can modify the LSP attribute defaults, but it is not necessary.

To change the defaults, use any or all of the following commands in ROUTER ISIS mode:

Command Syntax	Command Mode	Purpose
lsp-gen-interval [**level-1** | **level-2**] *seconds*	ROUTER ISIS	Set interval between LSP generation. • *seconds* range: 0 to 120 Default is 5 seconds. Default level is Level 1.
lsp-mtu *size*	ROUTER ISIS	Set the LSP size. • *size* range: 128 to 9195. Default is 1497.
lsp-refresh-interval *seconds*	ROUTER ISIS	Set the LSP refresh interval. • *seconds* range: 1 to 65535. Default is 900 seconds.
max-lsp-lifetime *seconds*	ROUTER ISIS	Set the maximum time LSPs lifetime. • *seconds* range: 1 to 65535 Default is 1200 seconds.

To view the configuration, use the **show config** command in ROUTER ISIS mode or the **show running-config isis** command in EXEC Privilege mode (Figure 475).

Figure 23-6. Command Example: show running-config isis

```
FTOS#show running-config isis
!
router isis
 lsp-refresh-interval 902
 net 47.0005.0001.000C.000A.4321.00
 net 51.0005.0001.000C.000A.4321.00
FTOS#
```

Configure IS-IS metric style and cost

All IS-IS links or interfaces are associated with a cost that is used in the SPF calculations. The possible cost varies depending on the metric style supported. If you configure narrow, transition or narrow transition metric style, the cost can be a number between 0 and 63. If you configure wide or wide transition metric style, the cost can be a number between 0 and 16,777,215. FTOS supports five different metric styles: narrow, wide, transition, narrow transition, and wide transition.

By default, FTOS generates and receives narrow metric values. Metrics or costs higher than 63 are not supported. To accept or generate routes with a higher metric, you must change the metric style of the IS-IS process. For example, if metric is configured as narrow, and an LSP with wide metrics is received, the route is not installed.

FTOS supports the following IS-IS metric styles:

Table 23-2. Metric Styles

Metric Style	Characteristics	Cost Range Supported on IS-IS Interfaces
narrow	Sends and accepts narrow or old TLVs (Type Length Value).	0 to 63
wide	Sends and accepts wide or new TLVs	0 to 16777215
transition	Sends both wide (new) and narrow (old) TLVs.	0 to 63
narrow transition	Sends narrow (old) TLVs and accepts both narrow (old) and wide (new) TLVs	0 to 63
wide transition	Sends wide (new) TLVs and accepts both narrow (old) and wide (new) TLVs.	0 to 16777215

Use the following command in ROUTER ISIS mode to change the IS-IS metric style of the IS-IS process.

Command Syntax	Command Mode	Purpose
metric-style {**narrow** [**transition**] | **transition** | **wide** [**transition**]} [**level-1** | **level-2**]	ROUTER ISIS	Set the metric style for the IS-IS process. Default: narrow Default: Level 1 and Level 2 (level-1-2)

Use the **show isis protocol** command (Figure 476) in EXEC Privilege mode to view which metric types are generated and received.

Figure 23-7. Command Example: show isis protocol

```
FTOS#show isis protocol
IS-IS Router: <Null Tag>
  System Id: EEEE.EEEE.EEEE  IS-Type: level-1-2
  Manual area address(es):
   47.0004.004d.0001
  Routing for area address(es):
   21.2223.2425.2627.2829.3031.3233
   47.0004.004d.0001
  Interfaces supported by IS-IS:
   Vlan 2
   GigabitEthernet 4/22
   Loopback 0
  Redistributing:
  Distance: 115
  Generate narrow metrics: level-1-2
  Accept narrow metrics:   level-1-2
  Generate wide metrics:   none
  Accept wide metrics:     none
FTOS#
```

IS-IS metrics settings.

When you change from one IS-IS metric style to another, the IS-IS metric value could be affected. For each interface with IS-IS enabled, you can assign a cost or metric that is used in the link state calculation.

Use the following command in INTERFACE mode to change the metric or cost of the interface.

Command Syntax	Command Mode	Purpose
isis metric *default-metric* [**level-1** | **level-2**]	INTERFACE	*default-value* range: 0 to 63 if the metric-style is narrow, narrow-transition or transition. 0 to 16777215 if the metric style is wide or wide transition. Default: 10.
isis ipv6 metric *default-metric* [**level-1** | **level-2**]	INTERFACE	Assign a metric for an IPv6 link or interface. • *default-metric* range: 0 to 63 for narrow and transition metric styles; 0 to 16777215 for wide metric styles. Default is 10. Default level is level-1. Refer to Configure IS-IS metric style and cost for more information on this command.

Use the **show config** command in INTERFACE mode or the **show isis interface** command in EXEC Privilege mode to view the interface's current metric.

Table 23-3. Correct Value Range for the isis metric command

Metric Style	Correct Value Range
wide	0 to 16777215
narrow	0 to 63
wide transition	0 to 16777215

Table 23-3. Correct Value Range for the isis metric command

Metric Style	Correct Value Range
narrow transition	0 to 63
transition	0 to 63

Configuring the distance of a route

Configure the distance for a route using the **distance** command from ROUTER ISIS mode.

Change the IS-type

You can configure the system to act as one of the following:

- Level 1 router
- Level 1-2 router
- Level 2 router

Use the following command in ROUTER ISIS mode to change the IS-type for the router, u

Command Syntax	Command Mode	Purpose
is-type {**level-1** | **level-1-2** | **level-2-only**}	ROUTER ISIS	Configure IS-IS operating level for a router. Default is level-1-2.

Command Syntax	Command Mode	Purpose
is-type {**level-1** | **level-1-2** | **level-2**}	ROUTER ISIS	Change the IS-type for the IS-IS process.

Use the **show isis protocol** command in EXEC Privilege mode (Figure 476) to view which IS-type is configured. The **show config** command in ROUTER ISIS mode displays only non-default information, so if you do not change the IS-type, the default value (level-1-2) is not displayed.

The default is Level 1-2 router. When the IS-type is Level 1-2, the software maintains two Link State databases, one for each level. Use the **show isis database** command to view the Link State databases (Figure 477).

Figure 23-8. Command Example: show isis database

```
FTOS#show isis database
IS-IS Level-1 Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime      ATT/P/OL
B233.00-00            0x00000003   0x07BF        1088              0/0/0
eljefe.00-00        * 0x00000009   0xF76A        1126              0/0/0
eljefe.01-00        * 0x00000001   0x68DF        1122              0/0/0
eljefe.02-00        * 0x00000001   0x2E7F        1113              0/0/0
Force10.00-00         0x00000002   0xD1A7        1102              0/0/0
IS-IS Level-2 Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime      ATT/P/OL
B233.00-00            0x00000006   0xC38A        1124              0/0/0
eljefe.00-00        * 0x0000000D   0x51C6        1129              0/0/0
eljefe.01-00        * 0x00000001   0x68DF        1122              0/0/0
eljefe.02-00        * 0x00000001   0x2E7F        1113              0/0/0
Force10.00-00         0x00000004   0xCDA9        1107              0/0/0

FTOS#
```

Control routing updates

Use the following commands in ROUTER ISIS mode to control the source of IS-IS route information.

Command Syntax	Command Mode	Purpose
passive-interface *interface*	ROUTER ISIS	Disable a specific interface from sending or receiving IS-IS routing information. Enter the type of interface and slot/port information: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For the Loopback interface on the RPM, enter the keyword **loopback** followed by a number from 0 to 16383. • For a port channel, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale. • For a SONET interface, enter the keyword **sonet** followed by slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

Distribute Routes

Another method of controlling routing information is to filter the information through a prefix list. Prefix lists are applied to incoming or outgoing routes and routes must meet the conditions of the prefix lists or FTOS does not install the route in the routing table. The prefix lists are globally applied on all interfaces running IS-IS.

Configure the prefix list in the PREFIX LIST mode prior to assigning it to the IS-IS process. For configuration information on prefix lists, see Chapter 8, IP Access Control Lists (ACL), Prefix Lists, and Route-maps.

IPv4 routes

Use the following commands in ROUTER ISIS mode to apply prefix lists to incoming or outgoing IPv4 routes.

Note: These commands apply to IPv4 IS-IS only. Use the ADDRESS-FAMILY IPV6 mode shown later to apply prefix lists to IPv6 routes

Command Syntax	Command Mode	Purpose
distribute-list *prefix-list-name* **in** [*interface*]	ROUTER ISIS	Apply a configured prefix list to all incoming IPv4 IS-IS routes. Enter the type of interface and slot/port information: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For the Loopback interface on the RPM, enter the keyword **loopback** followed by a number from 0 to 16383. • For a port channel, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale. • For a SONET interface, enter the keyword **sonet** followed by slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

Command Syntax	Command Mode	Purpose
distribute-list *prefix-list-name* **out** [**bgp** *as-number* | **connected** | **ospf** *process-id* | **rip** | **static**]	ROUTER ISIS	Apply a configured prefix list to all outgoing IPv4 IS-IS routes. You can configure one of the optional parameters: • **connected:** for directly connected routes. • **ospf** *process-id:* for OSPF routes only. • **rip:** for RIP routes only. • **static:** for user-configured routes. • **bgp**: for BGP routes only
distribute-list redistributed-override in	ROUTER ISIS	Deny RTM download for pre-existing redistributed IPv4 routes

IPv6 routes

Use these commands in ADDRESS-FAMILY IPV6 mode to apply prefix lists to incoming or outgoing IPv6 routes. =

Note: These commands apply to IPv6 IS-IS only. Use the ROUTER ISIS mode previously shown to apply prefix lists to IPv4 routes.

Command Syntax	Command Mode	Purpose
distribute-list *prefix-list-name* **in** [*interface*]	ROUTER ISIS-AF IPV6	Apply a configured prefix list to all incoming IPv6 IS-IS routes. Enter the type of interface and slot/port information: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For the Loopback interface on the RPM, enter the keyword **loopback** followed by a number from 0 to 16383. • For a port channel, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale, 1 to 32 for EtherScale. • For a SONET interface, enter the keyword **sonet** followed by slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

Command Syntax	Command Mode	Purpose
distribute-list *prefix-list-name* **out** [**bgp** *as-number* | **connected** | **ospf** *process-id* | **rip** | **static**]	ROUTER ISIS-AF IPV6	Apply a configured prefix list to all outgoing IPv6 IS-IS routes. You can configure one of the optional parameters: • **connected:** for directly connected routes. • **ospf** *process-id:* for OSPF routes only. • **rip:** for RIP routes only. • **static:** for user-configured routes. • **bgp**: for BGP routes only
distribute-list redistributed-override in	ROUTER ISIS-AF IPV6	Deny RTM download for pre-existing redistributed IPv6 routes

Redistribute routes

In addition to filtering routes, you can add routes from other routing instances or protocols to the IS-IS process. With the **redistribute** command syntax, you can include BGP, OSPF, RIP, static, or directly connected routes in the IS-IS process.

Note: Do not route iBGP routes to IS-IS unless there are route-maps associated with the IS-IS redistribution.

IPv4 routes

Use any of the following commands in ROUTER ISIS mode to add routes from other routing instances or protocols.

Note: These commands apply to IPv4 IS-IS only. Use the ADDRESS-FAMILY IPV6 mode shown later to apply prefix lists to IPv6 routes.

Command Syntax	Command Mode	Purpose
redistribute {**bgp** *as-number* | **connected** | **rip** | **static**} [**level-1 level-1-2** | **level-2**] [**metric** *metric-value*] [**metric-type** {**external** | **internal**}] [**route-map** *map-name*]	ROUTER ISIS	Include BGP, directly connected, RIP, or user-configured (static) routes in IS-IS. Configure the following parameters: • **level-1**, l**evel-1-2**, or **level-2**: Assign all redistributed routes to a level. Default is **level-2**. • *metric* range: 0 to 16777215. Default is 0. • **metric-type:** choose either **external** or **internal**. Default is internal. • *map-name*: name of a configured route map.

Command Syntax	Command Mode	Purpose
redistribute ospf *process-id* [**level-1**| **level-1-2** | **level-2**] [**metric** *value*] [**match external** {**1 | 2**} | **match internal**] [**metric-type** {**external** | **internal**}] [**route-map** *map-name*]	ROUTER ISIS	Include specific OSPF routes in IS-IS. Configure the following parameters: • *process-id* range: 1 to 65535 • **level-1**, l**evel-1-2**, or **level-2**: Assign all redistributed routes to a level. Default is **level-2**. • *metric* range: 0 to 16777215. Default is 0. • **match external** range: 1 or 2 • **match internal** • **metric-type**: external or internal. • *map-name*: name of a configured route map.

IPv6 routes

Use any of the these commands in ROUTER ISIS ADDRESS-FAMILY IPV6 mode to add routes from other routing instances or protocols.

Note: These commands apply to IPv6 IS-IS only. Use the ROUTER ISIS mode previously shown to apply prefix lists to IPv4 routes.

Command Syntax	Command Mode	Purpose
redistribute {**bgp** *as-number* | **connected** | **rip** | **static**} [**level-1 level-1-2** | **level-2**] [**metric** *metric-value*] [**metric-type** {**external** | **internal**}] [**route-map** *map-name*]	ROUTER ISIS	Include BGP, directly connected, RIP, or user-configured (static) routes in IS-IS. Configure the following parameters: • **level-1**, l**evel-1-2**, or **level-2**: Assign all redistributed routes to a level. Default is **level-2**. • *metric* range: 0 to 16777215. Default is 0. • **metric-type:** choose either **external** or **internal**. Default is internal. • *map-name*: name of a configured route map.
redistribute ospf *process-id* [**level-1**| **level-1-2** | **level-2**] [**metric** *value*] [**match external** {**1 | 2**} | **match internal**] [**metric-type** {**external** | **internal**}] [**route-map** *map-name*]	ROUTER ISIS	Include specific OSPF routes in IS-IS. Configure the following parameters: • *process-id* range: 1 to 65535 • **level-1**, l**evel-1-2**, or **level-2**: Assign all redistributed routes to a level. Default is **level-2**. • *metric* range: 0 to 16777215. Default is 0. • **match external** range: 1 or 2 • **match internal** • **metric-type**: external or internal. • *map-name*: name of a configured route map.

Use the **show running-config isis** command in EXEC Privilege mode to view IS-IS configuration globally (including both IPv4 and IPv6 settings), or the **show config** command in ROUTER ISIS mode to view the current IPv4 IS-IS configuration, or the **show config** command in ROUTER ISIS-ADDRESS FAMILY IPV6 mode to view the current IPv6 IS-IS configuration

Configure authentication passwords

You can assign an authentication password for routers in Level 1 and for routers in Level 2. Since Level 1 and Level 2 routers do not communicate with each other, you can assign different passwords for Level 1 routers and for Level 2 routers. If you want the routers in the level to communicate with each other, though, they must be configured with the same password.

Use either or both of the commands in ROUTER ISIS mode to configure a simple text password.

Command Syntax	Command Mode	Purpose
area-password [hmac-md5] *password*	ROUTER ISIS	Configure authentication password for an area. FTOS supports HMAC-MD5 authentication. This password is inserted in Level 1 LSPs, Complete SNPs, and Partial SNPs.
domain-password [*encryption-type* | **hmac-md5**] *password*	ROUTER ISIS	Set the authentication password for a routing domain. FTOS supports both DES and HMAC-MD5 authentication methods. This password is inserted in Level 2 LSPs, Complete SNPs, and Partial SNPs.

Use the **show config** command in ROUTER ISIS mode or the **show running-config isis** command in EXEC Privilege mode to view the passwords.

Remove a password by using either the **no area-password** or **no domain-password** command in ROUTER ISIS mode.

Set the overload bit

Another use for the overload bit is to prevent other routers from using this router as an intermediate hop in their shortest path first (SPF) calculations. For example, if the IS-IS routing database is out of memory and cannot accept new LSPs, FTOS sets the overload bit and IS-IS traffic continues to transit the system.

Use this command the following command in ROUTER ISIS mode to set the overload bit manually.

Command Syntax	Command Mode	Purpose
set-overload-bit	ROUTER ISIS	Set the overload bit in LSPs. This prevents other routers from using it as an intermediate hop in their shortest path first (SPF) calculations.

Enter **no set-overload-bit** to remove the overload bit.

When the bit is set, a 1 is placed in the OL column in the show isis database command output. In Figure 23-9, the overload bit is set in both the Level-1 and Level-2 database because the IS type for the router is Level-1-2

Figure 23-9. Command Example: show isis database

```
FTOS#show isis database
IS-IS Level-1 Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime      ATT/P/OL
B233.00-00              0x00000003   0x07BF        1074              0/0/0
eljefe.00-00        *   0x0000000A   0xF963        1196              0/0/1
eljefe.01-00        *   0x00000001   0x68DF        1108              0/0/0
eljefe.02-00        *   0x00000001   0x2E7F        1099              0/0/0
Force10.00-00           0x00000002   0xD1A7        1088              0/0/0
IS-IS Level-2 Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime      ATT/P/OL
B233.00-00              0x00000006   0xC38A        1110              0/0/0
eljefe.00-00        *   0x0000000E   0x53BF        1196              0/0/1
eljefe.01-00        *   0x00000001   0x68DF        1108              0/0/0
eljefe.02-00        *   0x00000001   0x2E7F        1099              0/0/0
Force10.00-00           0x00000004   0xCDA9        1093              0/0/0
FTOS#
```

when overload bit is set, 1 is listed in the OL column.

Debug IS-IS

Enter the **debug isis** command in EXEC Privilege mode to debug all IS-IS processes.

Use the following commands for specific IS-IS debugging.

Command Syntax	Command Mode	Purpose
debug isis	EXEC Privilege	View all IS-IS information.
debug isis adj-packets [*interface*]	EXEC Privilege	View information on all adjacency-related activity (for example, hello packets that are sent and received). To view specific information, enter one of the following optional parameters: • *interface:* Enter the type of interface and slot/port information to view IS-IS information on that interface only.
debug isis local-updates [*interface*]	EXEC Privilege	View information about IS-IS local update packets. To view specific information, enter one of the following optional parameters: • *interface:* Enter the type of interface and slot/port information to view IS-IS information on that interface only.

Command Syntax	Command Mode	Purpose
debug isis snp-packets [*interface*]	EXEC Privilege	View IS-IS SNP packets, include CSNPs and PSNPs. To view specific information, enter one of the following optional parameters: • *interface:* Enter the type of interface and slot/port information to view IS-IS information on that interface only.
debug isis spf-triggers	EXEC Privilege	View the events that triggered IS-IS shortest path first (SPF) events for debugging purposes.
debug isis update-packets [*interface*]	EXEC Privilege	View sent and received LSPs. To view specific information, enter one of the following optional parameters: • *interface:* Enter the type of interface and slot/port information to view IS-IS information on that interface only.

FTOS displays debug messages on the console. Use the **show debugging** command in EXEC Privilege mode to view which debugging commands are enabled.

Enter the keyword no followed by the debug command to disable a specific debug command. For example, to disable debugging of IS-IS updates, enter **no debug isis updates-packets**.

Enter **no debug isis** to disable all IS-IS debugging.

Enter **undebug all** to disable all debugging, e.

IS-IS Metric Styles

The following sections provide additional information on IS-IS Metric Styles.

- IS-IS Metric Styles on page 531
- Configure Metric Values on page 532

FTOS supports the following IS-IS metric styles:

- narrow (supports only type, length, and value (TLV) up to 63)
- wide (supports TLV up to 16777215)
- transition (supports both narrow and wide and uses a TLV up to 63)
- narrow transition (accepts both narrow and wide and sends only narrow or old-style TLV)
- wide transition (accepts both narrow and wide and sends only wide or new-style TLV)

Configure Metric Values

The following topics are covered in this section:

For any level (Level-1, Level-2, or Level-1-2), the value range possible in the **isis metric** command in INTERFACE mode changes depending on the metric style.

Table 23-4. Correct Value Range for the isis metric Command

Metric Style	Correct Value Range for the isis metric Command
wide	0 to 16777215
narrow	0 to 63
wide transition	0 to 16777215
narrow transition	0 to 63
transition	0 to 63

Maximum Values in the Routing Table

IS-IS metric styles support different cost ranges for the route. The cost range for the narrow metric style is 0 to 1023, while all other metric styles support a range of 0 to 0xFE000000.

Changing the IS-IS Metric Style in One Level Only

By default, the IS-IS metric style is narrow. When you change from one IS-IS metric style to another, the IS-IS metric value (configured with the **isis metric** command) could be affected.

In the following scenarios, the IS-type is either Level-1 or Level-2 or Level-1-2 and the metric style changes.

Table 23-5. Metric Value when Metric Style Changes

Beginning metric style	Final metric style	Resulting IS-IS metric value
wide	narrow	default value (10) if the original value is greater than 63. A message is sent to the console.
wide	transition	truncated value[1] (the truncated value appears in the LSP only.) The original **isis metric** value is displayed in the **show config** and **show running-config** commands and is used if you change back to transition metric style.

Table 23-5. Metric Value when Metric Style Changes

Beginning metric style	Final metric style	Resulting IS-IS metric value
wide	narrow transition	default value (10) if the original value is greater than 63. A message is sent to the console.
wide	wide transition	original value
narrow	wide	original value
narrow	transition	original value
narrow	narrow transition	original value
narrow	wide transition	original value
transition	wide	original value
transition	narrow	original value
transition	narrow transition	original value
transition	wide transition	original value
narrow transition	wide	original value
narrow transition	narrow	original value
narrow transition	wide transition	original value
narrow transition	transition	original value
wide transition	wide	original value
wide transition	narrow	default value (10) if the original value is greater than 63. A message is sent to the console.
wide transition	narrow transition	default value (10) if the original value is greater than 63. A message is sent to the console.
wide transition	transition	truncated value (the truncated value appears in the LSP only.) The original **isis metric** value is displayed in the **show config** and **show running-config** commands and is used if you change back to transition metric style.

1 a truncated value is a value that is higher than 63, but set back to 63 because the higher value is not supported.

Moving to transition and then to another metric style produces different results (Table 23-6).

Table 23-6. Metric Value when Metric Style Changes Multiple Times

Beginning metric style	next isis metric style	resulting isis metric value	Next metric style	final isis metric value
wide	transition	truncated value	wide	original value is recovered
wide transition	transition	truncated value	wide transition	original value is recovered

Table 23-6. Metric Value when Metric Style Changes Multiple Times

Beginning metric style	next isis metric style	resulting isis metric value	Next metric style	final isis metric value
wide	transition	truncated value	narrow	default value (10) A message is sent to the logging buffer
wide transition	transition	truncated value	narrow transition	default value (10) A message is sent to the logging buffer

Leaking from One Level to Another

In the following scenarios, each IS-IS level is configured with a different metric style.

Table 23-7. Metric Value with Different Levels Configured with Different Metric Styles

Level-1 metric style	Level-2 metric style	Resulting isis metric value
narrow	wide	original value
narrow	wide transition	original value
narrow	narrow transition	original value
narrow	transition	original value
wide	narrow	truncated value
wide	narrow transition	truncated value
wide	wide transition	original value
wide	transition	truncated value
narrow transition	wide	original value
narrow transition	narrow	original value
narrow transition	wide transition	original value
narrow transition	transition	original value
transition	wide	original value
transition	narrow	original value
transition	wide transition	original value
transition	narrow transition	original value
wide transition	wide	original value
wide transition	narrow	truncated value
wide transition	narrow transition	truncated value
wide transition	transition	truncated value

Sample Configuration

The following configurations are examples for enabling IPv6 IS-IS. These are not comprehensive directions. They are intended to give you a some guidance with typical configurations.

Note: Only **one** IS-IS process can run on the router, even if both IPv4 and IPv6 routing is being used.

You can copy and paste from these examples to your CLI. Be sure you make the necessary changes to support your own IP Addresses, Interfaces, Names, etc.

Note: Whenever ISIS configuration changes are made, the IS-IS process must be cleared (re-started) using **clear isis**. The command **clear isis** must include the tag for the ISIS process. The example below shows the response from the router:

```
FTOS#clear isis *
% ISIS not enabled.
FTOS#clear isis 9999 *
```

Figure 23-10 is a sample configuration for enabling IPv6 IS-IS. Figure 23-13 illustrates the topology created with that CLI configuration.

Figure 23-10. IS-IS Sample Configuration

Router 1

```
R1(conf)#interface Loopback 0
R1(conf-if-lo-0)#ip address 192.168.1.1/24
R1(conf-if-lo-0)#ipv6 address 2001:db8:9999:1::/48
R1(conf-if-lo-0)#ip router isis 9999
R1(conf-if-lo-0)#no shutdown

R1(conf-if-lo-0)#router isis 9999
R1(conf-router_isis)#is-type level-1
R1(conf-router_isis)#net FF.F101.0002.0C00.1111.00
R1(conf-router_isis)#ipv6 route 2001:db8:9999:2::/128 2001:db8:1021:2::
R1(conf)#ipv6 route 2001:db8:9999:3::/128 2001:db8:1022:3::
R1(conf)#ip route 192.168.1.2/32 10.0.12.2
R1(conf)#ip route 192.168.1.3/32 10.0.13.3

R1(conf)#interface GigabitEthernet 1/21
R1(conf-if-gi-1/21)#ip address 10.0.12.1/24
R1(conf-if-gi-1/21)#ipv6 address 2001:db8:1022:1::/48
R1(conf-if-gi-1/21)#isis circuit-type level-1
R1(conf-if-gi-1/21)#isis network point-to-point
R1(conf-if-gi-1/21)#ip router isis 9999
R1(conf-if-gi-1/21)#no shutdown

R1(conf-if-gi-1/21)#interface GigabitEthernet 1/34
R1(conf-if-gi-1/34)# ip address 10.0.13.1/24
R1(conf-if-gi-1/34)#ipv6 address 2001:db8:1021:1::/48
R1(conf-if-gi-1/34)#ip router isis 9999
R1(conf-if-gi-1/34)#no shutdown
R1(conf-if-gi-1/34)#end

R1#show ip route
Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set
       Destination        Gateway                      Dist/Metric Last Change
       -----------        -------                      ----------- -----------
  C    10.0.12.0/24       Direct, Gi 1/21                     0/0    00:00:57
  C    192.168.1.0/24     Direct, Lo 0                        0/0    00:04:19
  S    192.168.1.2/32     via 10.0.12.2, Gi 1/21              1/0    00:00:57

R1#show isis data
IS-IS Level-1 Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime      ATT/P/OL
R1.00-00            * 0x0000000F   0x3A6C        1176              0/0/0
R1.02-00            * 0x00000002   0x90AC        1076              0/0/0
R1.03-00            * 0x00000002   0x67C3        1176              0/0/0
R2.00-00              0x0000000C   0x5418        1183              0/0/0
R2.00-00              0x00000009   0x1E39        1183              0/0/0
R2.03-00              0x00000002   0x589D        1180              0/0/0

R1#show isis neigh
System Id      Interface State     Type Priority Uptime           Circuit Id
R2             Gi 1/21   Up        L1    64      00:02:28         R1.02
R2             Gi 1/34   Up        L1    64      00:00:42         R1.03
R1#
```

Figure 23-11. IS-IS Sample Configuration continued

Router 2

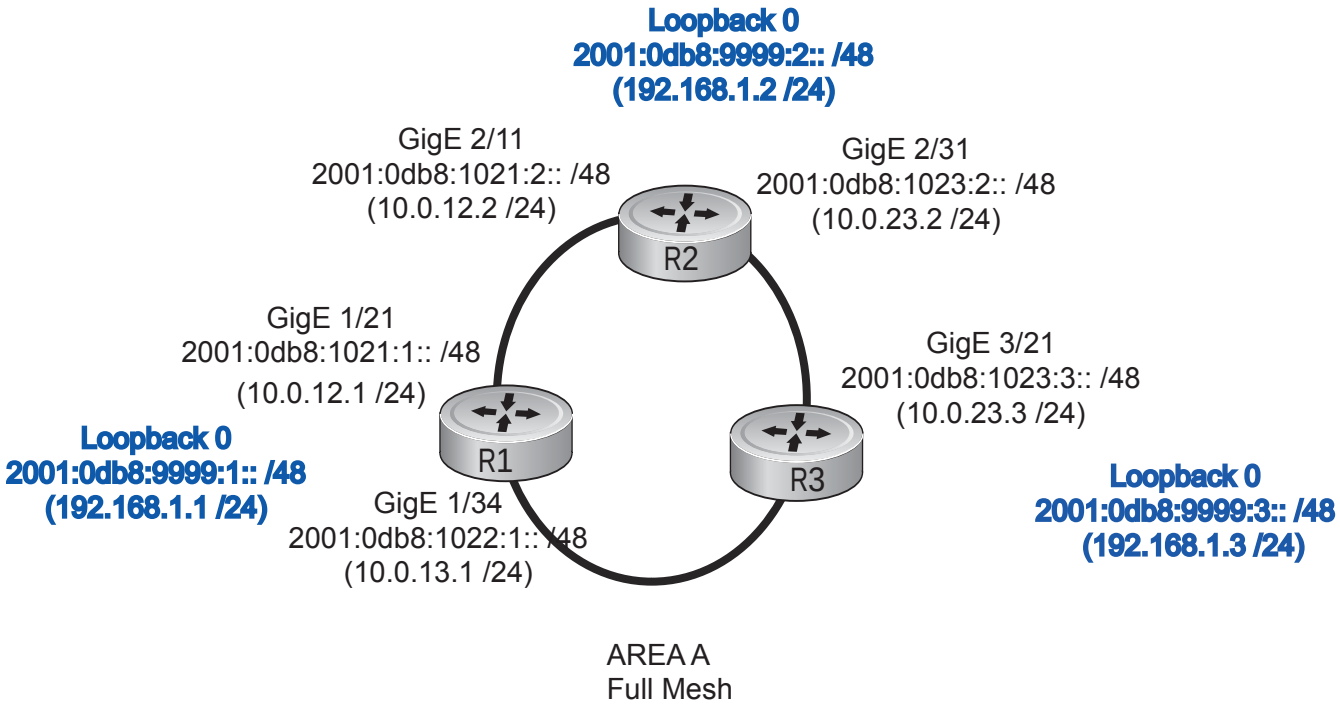
```
R2(conf)#interface Loopback 0
R2(conf-if-lo-0)#ip address 192.168.1.1/24
R2(conf-if-lo-0)#ipv6 address 2001:db8:9999:1::/48
R2(conf-if-lo-0)#ip router isis 9999
R2(conf-if-lo-0)#no shutdown

R2(conf-if-lo-0)#router isis 9999
R2(conf-router_isis)#int gi 2/11
R2(conf-if-gi-2/11)#ip address 10.0.12.2/24
R2(conf-if-gi-2/11)#ipv6 address 2001:db8:9999:2::/48
R2(conf-if-gi-2/11)#ip router isis 9999
R2(conf-if-gi-2/11)#isis network point-to-point
R2(conf-if-gi-2/11)#no shutdown
R2(conf-if-gi-2/11)#int gi 2/31
R2(conf-if-gi-2/31)#ip address 10.0.23.2/24
R2(conf-if-gi-2/31)#ipv6 address 2001:db8:1021:2::/48
R2(conf-if-gi-2/31)#ip router isis 9999
R2(conf-if-gi-2/31)#isis network point-to-point
R2(conf-if-gi-2/31)#no shutdown
R2(conf-if-gi-2/31)#router isis 9999

R2(conf-router_isis)#ipv6 route 2001:db8:9999:1::/128 2001:db8:1021:1::
R2(conf)#ipv6 route 2001:db8:9999:3::/128 2001:db8:1023:3::
R2(conf)#ip route 192.168.1.1/32 10.0.12.1
R2(conf)#ip route 192.168.1.3/32 10.0.23.3
R2(conf)#ex

R2#show ip route
Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is 172.21.212.1 to network 0.0.0.0

       Destination        Gateway                     Dist/Metric Last Change
       -----------        -------                     ----------- -----------
 *S    0.0.0.0/0          via 172.21.212.1, Vl 212            1/0    00:59:50
  C    10.0.12.0/24       Direct, Gi 2/11                     0/0    00:02:25
  C    10.0.23.0/24       Direct, Gi 2/31                     0/0    00:01:53
  C    10.10.92.0/24      Direct, Po 4                        0/0        6d9h
  C    172.21.212.0/24    Direct, Vl 212                      0/0       2d20h
  C    192.168.1.0/24     Direct, Lo 0                        0/0    01:11:48
  S    192.168.1.1/32     via 10.0.12.1, Gi 2/11              1/0    00:00:51
  S    192.168.1.3/32     via 10.0.23.3, Gi 2/31              1/0    00:00:39
R2#show isis data
IS-IS Level-1 Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime      ATT/P/OL
R2.00-00            * 0x0000000F   0x0174        1088              0/0/0

R2#show isis neigh
A2#show isis neigh
System Id      Interface State     Type Priority Uptime            Circuit Id
R1             Gi 2/11   Up        L1    64      00:02:19          A102.02
R3             Gi 2/31   Up        L1    64      00:00:25          A121.03
R2#
```

Figure 23-12. IS-IS Sample Configuration continued

Router 3

```
R3(conf)#interface Loopback 0
R3(conf-if-lo-0)#ip address 192.168.1.3/24
R3(conf-if-lo-0)#ipv6 address 2001:db8:9999:3::/48
R3(conf-if-lo-0)#ip router isis 9999
R3(conf-if-lo-0)#no shutdown
R3(conf-if-lo-0)#router isis 9999
R3(conf-router_isis)#net FF.F101.0002.0C00.1133.00
R3(conf-router_isis)#ipv6 route 2001:db8:9999:1::/128 2001:db8:1022:1::
R3(conf)#ipv6 route 2001:db8:9999:2::/128 2001:db8:1023:2::
R3(conf)#ip route 192.168.1.1/32 10.0.13.1

R3(conf)#interface GigabitEthernet 3/14
R3(conf-if-gi-3/14)#ip address 10.0.13.3/24
R3(conf-if-gi-3/14)#ipv6 address 2001:db8:1022:3::/48
R3(conf-if-gi-3/14)#ip router isis 9999
R3(conf-if-gi-3/14)#isis circuit-type level-1
R3(conf-if-gi-3/14)#isis network point-to-point
R3(conf-if-gi-3/14)#no shutdown
R3(conf-if-gi-3/14)#interface GigabitEthernet 3/21
R3(conf-if-gi-3/21)#ip address 10.0.23.3/24
R3(conf-if-gi-3/21)#ipv6 address 2001:db8:1023:3::/48
R3(conf-if-gi-3/21)#ip router isis 9999
R3(conf-if-gi-3/21)#isis circuit-type level-1
R3(conf-if-gi-3/21)#isis network point-to-point
R3(conf-if-gi-3/21)#no shutdown
R3(conf-if-gi-3/21)#end

R3#show ip route
Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route
Gateway of last resort is not set
       Destination        Gateway                     Dist/Metric Last Change
       -----------        -------                     ----------- -----------
  C    10.0.13.0/24       Direct, Gi 3/14                    0/0    00:00:10
  C    10.0.23.0/24       Direct, Gi 3/21                    0/0    00:00:03
  C    192.168.1.0/24     Direct, Lo 0                       0/0    00:00:32
  S    192.168.1.1/32     via 10.0.13.1, Gi 3/14             1/0    00:00:10
  S    192.168.1.2/32     via 10.0.23.2, Gi 3/21             1/0    00:00:03

R2#show isis data
IS-IS Level-1 Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime      ATT/P/OL
R1.00-00            0x0000000F   0x3A6C        1198              0/0/0
R1.03-00            0x00000001   0x69C2        1193              0/0/0
R2.00-00          * 0x00000007   0x51F6        1198              0/0/0
R2.03-00          * 0x00000001   0x5A9C        1200              0/0/0

IS-IS Level-2 Link State Database
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime      ATT/P/OL
R3.00-00          * 0x00000008   0xC09C        1199              0/0/0
R3#show isis neigh
System Id      Interface State     Type Priority Uptime           Circuit Id
R1           Gi 3/14   Init      L1    64      00:00:02         R1.03
R2           Gi 3/21   Up        L1    64      00:00:14         A101.03
```

Figure 23-13. IPv6 IS-IS Sample Topography



24

Link Aggregation Control Protocol

Link Aggregation Control Protocol is supported on platforms C E S

LACP addressing is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

The major sections in the chapter are:

Introduction to Dynamic LAGs and LACP

A *Link Aggregation Group* (*LAG*), referred to as a *port channel* by FTOS, can provide both load-sharing and port redundancy across line cards. LAGs can be enabled as static or dynamic. The benefits and constraints are basically the same, as described in Port Channel Interfaces on page 428 in Chapter 20, Interfaces.

The unique benefit of a dynamic LAG is that its ports can toggle between participating in the LAG or acting as dedicated ports, whereas ports in a static LAG must be specifically removed from the LAG in order to act alone.

FTOS uses LACP to create dynamic LAGs. LACP provides a standardized means of exchanging information between two systems (also called Partner Systems) and automatically establishes the LAG between the systems. LACP permits the exchange of messages on a link to allow their LACP instances to:

- Reach agreement on the identity of the LAG to which the link belongs.
- Move the link to that LAG.
- Enable the transmission and reception functions in an orderly manner.

The FTOS implementation of LACP is based on the standards specified in the IEEE 802.3: “Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications.”

LACP functions by constantly exchanging custom MAC PDUs across LAN Ethernet links. The protocol packets are only exchanged between ports that are configured as LACP capable.

Important Points to Remember

- On ExaScale, LACP is supported on 200 physical ports. Use static LAGs for the remaining ports to avoid unpredictable results.
- LACP enables you to add members to a port channel (LAG) as long as it has no static members. Conversely, if the LAG already contains a statically defined member (**channel-member** command), the **port-channel mode** command is not permitted.
- A static LAG cannot be created if a dynamic LAG using the selected number already exists.
- **No dual membership in static and dynamic LAGs**:
  - If a physical interface is a part of a static LAG, then the command **port-channel-protocol lacp** will be rejected on that interface.
  - If a physical interface is a part of a dynamic LAG, it cannot be added as a member of a static LAG. The command **channel-member gigabitethernet** *x/y* will be rejected in the static LAG interface for that physical interface.
- A dynamic LAG can be created with any type of configuration.
- There is a difference between the **shutdown** and **no interface port-channel**:
  - — The **shutdown** command on LAG “xyz” disables the LAG and retains the user commands. However, the system does not allow the channel number “xyz” to be statically created.
  - — The command **no interface port-channel** *channel-number* deletes the specified LAG, including a dynamically created LAG. This command causes all LACP-specific commands on the member interfaces to be removed. The interfaces are restored to a state that is ready to be configured.

  Note: There will be no configuration on the interface since that condition is required for an interface to be part of a LAG.
- Link dampening can be configured on individual members of a LAG. See Link Debounce Timer on page 446 for more information.
- LACP cannot add an interface to a LAG if one of the LAG members is shut down on the remote interface. If a remote LAG member is shut down, Message 1 appears on the local system when you attempt to add a member. In this case, enable all members of the LAG on the remote system, and then add any new members on the local system.
- FTOS might not synch connected and static routes learned via LACP to the secondary RPM. Configure **redundancy protocol lacp** to synch LACP states between RPMs.

Message 1 LACP Remote Port Down Error Message

```
% Error: This port property does not match with other LAG member.
```

LACP modes

FTOS provides the following three modes for configuration of LACP:

- **Off**—In this state, an interface is not capable of being part of a dynamic LAG. LACP does not run on any port that is configured to be in this state.
- **Active**—In this state, the interface is said to be in the "active negotiating state." LACP runs on any link that is configured to be in this state. A port in Active state also automatically initiates negotiations with other ports by initiating LACP packets.
- **Passive**—In this state, the interface is not in an active negotiating state, but LACP will run on the link. A port in Passive state also responds to negotiation requests (from ports in Active state). Ports in Passive state respond to LACP packets.

FTOS supports LAGs in the following cases:

- A port in Active state can set up a port channel (LAG) with another port in Active state.
- A port in Active state can set up a LAG with another port in Passive state.

A port in Passive state cannot set up a LAG with another port in Passive state.

LACP Configuration Commands

If aggregated ports are configured with compatible LACP modes (Off, Active, Passive), LACP can automatically link them, as defined in IEEE 802.3, Section 43. The following commands configure LACP:

Command Syntax	Command Mode	Purpose
[**no**] **lacp system-priority** *priority-value*	CONFIGURATION	Configure the system priority. Range: 1– 65535 (the higher the number, the lower the priority) Default: 32768
[**no**] **port-channel-protocol lacp**	INTERFACE	Enable or disable LACP on any LAN port: • Default is "LACP disabled" • This command creates a new context.
[**no**] **port-channel** *number* **mode** [**active** | **passive** | **off**]	LACP	Configure LACP mode. • Default is "LACP active" • **number** cannot statically contain any links
[**no**] **lacp port-priority** *priority-value*	LACP	Configure port priority. • Ranges: 1 – 65535 (the higher the number, the lower the priority) • Default: 32768

LACP Configuration Tasks

The tasks covered in this section are:

- Create a LAG
- Configure the LAG interfaces as dynamic on page 544
- Set the LACP long timeout on page 545
- Monitor and Debugging LACP on page 546
- Configure Shared LAG State Tracking on page 547

Create a LAG

To create a dynamic port channel (LAG), define the LAG and then the LAG interfaces. Use the **interface port-channel** and **switchport** commands, as shown in Figure 24-1, which uses the example of LAG 32:

Figure 24-1. Placing a LAG into the Default VLAN

```
FTOS(conf)#interface port-channel 32
FTOS(conf-if-po-32)#no shutdown
FTOS(conf-if-po-32)#switchport
```

The LAG is in the default VLAN. To place the LAG into a non-default VLAN, use the **tagged** command on the LAG (Figure 24-2):

Figure 24-2. Placing a LAG into a Non-default VLAN

```
FTOS(conf)#interface vlan 10
FTOS(conf-if-vl-10)#tagged port-channel 32
```

Configure the LAG interfaces as dynamic

After creating a LAG, configure the dynamic LAG interfaces. Figure 24-3 shows ports 3/15, 3/16, 4/15, and 4/16 added to LAG 32 in LACP mode with the command **port-channel-protocol lacp**.

Figure 24-3. Creating a Dynamic LAG Example

```
FTOS(conf)#interface Gigabitethernet 3/15
FTOS(conf-if-gi-3/15)#no shutdown
FTOS(conf-if-gi-3/15)#port-channel-protocol lacp
FTOS(conf-if-gi-3/15-lacp)#port-channel 32 mode active
...
FTOS(conf)#interface Gigabitethernet 3/16
FTOS(conf-if-gi-3/16)#no shutdown
FTOS(conf-if-gi-3/16)#port-channel-protocol lacp
FTOS(conf-if-gi-3/16-lacp)#port-channel 32 mode active
...
FTOS(conf)#interface Gigabitethernet 4/15
FTOS(conf-if-gi-4/15)#no shutdown
FTOS(conf-if-gi-4/15)#port-channel-protocol lacp
FTOS(conf-if-gi-4/15-lacp)#port-channel 32 mode active
...
FTOS(conf)#interface Gigabitethernet 4/16
FTOS(conf-if-gi-4/16)#no shutdown
FTOS(conf-if-gi-4/16)#port-channel-protocol lacp
FTOS(conf-if-gi-4/16-lacp)#port-channel 32 mode active
```

The **port-channel 32 mode active** command shown above may be successfully issued as long as there is no existing static channel-member configuration in LAG 32.

Set the LACP long timeout

PDUs are exchanged between port channel (LAG) interfaces to maintain LACP sessions. PDUs are transmitted at either a slow or fast transmission rate, depending upon the LACP timeout value. The timeout value is the amount of time that a LAG interface waits for a PDU from the remote system before bringing the LACP session down. The default timeout value is 1 second; it can be configured to be 30 seconds. Invoking the longer timeout might prevent the LAG from flapping if the remote system is up but temporarily unable to transmit PDUs due to a system interruption.

Note: The 30-second timeout is available for dynamic LAG interfaces only. The **lacp long-timeout** command can be entered for static LAGs, but it has no effect.

To configure the LACP long timeout (Figure 196):

Step	Task	Command Syntax	Command Mode
1	Set the LACP timeout value to 30 seconds.	**lacp long-timeout**	CONFIG-INT-PO

Figure 24-4. Invoking the LACP Long Timeout

```
FTOS(conf)# interface port-channel 32
FTOS(conf-if-po-32)#no shutdown
FTOS(conf-if-po-32)#switchport
FTOS(conf-if-po-32)#lacp long-timeout
FTOS(conf-if-po-32)#end
FTOS# show lacp 32
Port-channel 32 admin up, oper up, mode lacp
Actor System ID: Priority 32768, Address 0001.e800.a12b
Partner System ID: Priority 32768, Address 0001.e801.45a5
Actor Admin Key 1, Oper Key 1, Partner Oper Key 1
LACP LAG 1 is an aggregatable link
A - Active LACP, B - Passive LACP, C - Short Timeout, D - Long Timeout
E - Aggregatable Link, F - Individual Link, G - IN_SYNC, H - OUT_OF_SYNC
I - Collection enabled, J - Collection disabled, K - Distribution enabled L - Distribution disabled,
M - Partner Defaulted, N - Partner Non-defaulted, O - Receiver is in expired state,
P - Receiver is not in expired state
Port Gi 10/6 is enabled, LACP is enabled and mode is lacp
Actor Admin: State ADEHJLMP Key 1 Priority 128
```

Note: View PDU exchanges and the timeout value using the command **debug lacp**. See Monitor and Debugging LACP on page 546.

Monitor and Debugging LACP

The system log (syslog) records faulty LACP actions.

To debug LACP, use the following command:

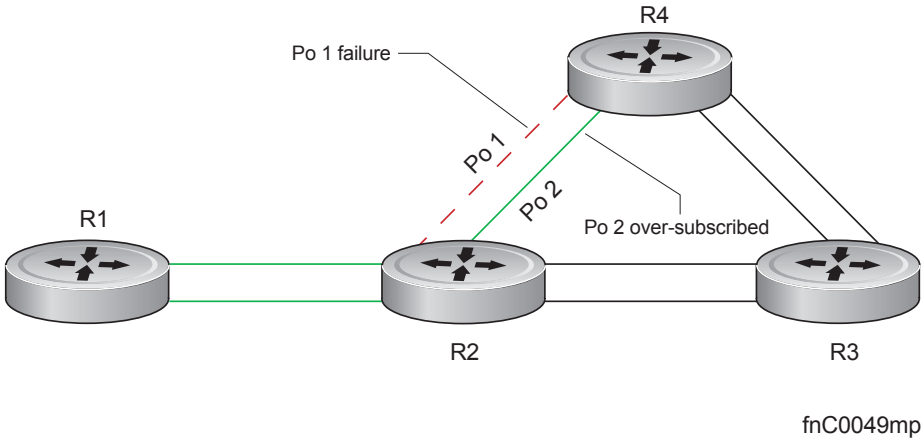
Command Syntax	Command Mode	Purpose
[**no**] **debug lacp** [**config** | **events** | **pdu** [**in** | **out** | [*interface* [**in** | **out**]]]]	EXEC	Debug LACP, including configuration and events.

Shared LAG State Tracking

Shared LAG State Tracking provides the flexibility to bring down a port channel (LAG) based on the operational state of another LAG. At any time, only two LAGs can be a part of a group such that the fate (status) of one LAG depends on the other LAG.

In Figure 24-5, line-rate traffic from R1 destined for R4 follows the lowest-cost route via R2, as shown. Traffic is equally distributed between LAGs 1 and 2. If LAG 1 fails, all traffic from R1 to R4 flows across LAG 2 only. This condition over-subscribes the link, and packets are dropped.

Figure 24-5. LAGs using ECMP without Shared LAG State Tracking



To avoid packet loss, traffic must be re-directed through the next lowest-cost link (R3 to R4). FTOS has the ability to bring LAG 2 down in the event that LAG 1 fails, so that traffic can be re-directed, as described. This is what is meant by Shared LAG State Tracking. To achieve this functionality, you must group LAG 1 and LAG 2 into a single entity, called a *failover group*.

Configure Shared LAG State Tracking

To configure Shared LAG State Tracking, you configure a failover group:

Step	Task	Command	Command Mode
1	Enter port-channel failover group mode.	**port-channel failover-group**	CONFIGURATION
2	Create a failover group and specify the two port-channels that will be members of the group.	**group** *number* **port-channel** *number* **port-channel** *number*	CONFIG-PO-FAILOVER-GRP

In Figure 24-6, LAGs 1 and 2 have been placed into to the same failover group.

Figure 24-6. Configuring Shared LAG State Tracking

```
R2#config
R2(conf)#port-channel failover-group
R2(conf-po-failover-grp)#group 1 port-channel 1 port-channel 2
```

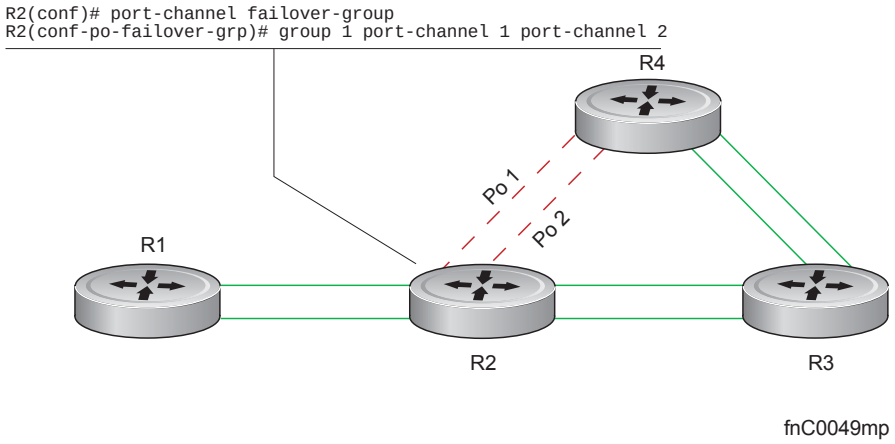
View the failover group configuration using the **show running-configuration po-failover-group** command, as shown in Figure 24-7.

Figure 24-7. Viewing Shared LAG State Tracking in the Running-configuration

```
R2#show running-config po-failover-group
!
port-channel failover-group
 group 1 port-channel 1 port-channel 2
```

In Figure 24-8, LAGs 1 and 2 are members of a failover group. LAG 1 fails and LAG 2 is brought down upon the failure. This effect is logged by Message 2, in which a console message declares both LAGs down at the same time.

Figure 24-8. Shared LAG State Tracking



Message 2 Shared LAG State Tracking Console Message

```
2d1h45m: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Po 1
2d1h45m: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Po 2
```

View the status of a failover group member using the command **show interface port-channel,** as shown in Figure 24-9.

Figure 24-9. Viewing Status of a Failover Group Member

```
R2#show interface Port-channel 2
Port-channel 2 is up, line protocol is down (Failover-group 1 is down)
Hardware address is 00:01:e8:05:e8:4c, Current address is 00:01:e8:05:e8:4c
Interface index is 1107755010
Minimum number of links to bring Port-channel up is 1
Port-channel is part of failover-group 1
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 1000 Mbit
Members in this channel:  Gi 1/17(U)
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:01:28
Queueing strategy: fifo
```

Note: The set of console messages shown in Message 2 appear only if Shared LAG State Tracking is configured on that router (the feature can be configured on one or both sides of a link). For example, in Figure 24-8, if Shared LAG State Tracking is configured on R2 only, then no messages appear on R4 regarding the state of LAGs in a failover group.

Important Points about Shared LAG State Tracking

- This feature is available for static and dynamic LAGs.

- Only a LAG can be a member of a failover group.
- Shared LAG State Tracking can be configured on one side of a link or on both sides.
- If a LAG that is part of a failover group is deleted, the failover group is deleted.
- If a LAG moves to the down state due to this feature, its members may still be in the up state.

Configure LACP as Hitless

Configure LACP as Hitless is supported only on platforms: C E

LACP on Dell Force10 systems can be configured to be hitless. When configured as hitless, there is no noticeable impact on dynamic LAG state upon an RPM failover. Critical LACP state information is synchronized between the two RPMs. See Hitless Behavior on page 389.

Configure LACP to be hitless using the command **redundancy protocol lacp** from CONFIGURATION mode, as shown in Figure 24-10.

Figure 24-10. Enabling Hitless LACP

```
FTOS(conf)#redundancy protocol lacp


FTOS#show running-config redundancy
!
redundancy protocol lacp
FTOS#
FTOS#show running-config interface gigabitethernet 0/12
!
interface GigabitEthernet 0/12
 no ip address
!
 port-channel-protocol LACP
  port-channel 200 mode active
 no shutdown
```

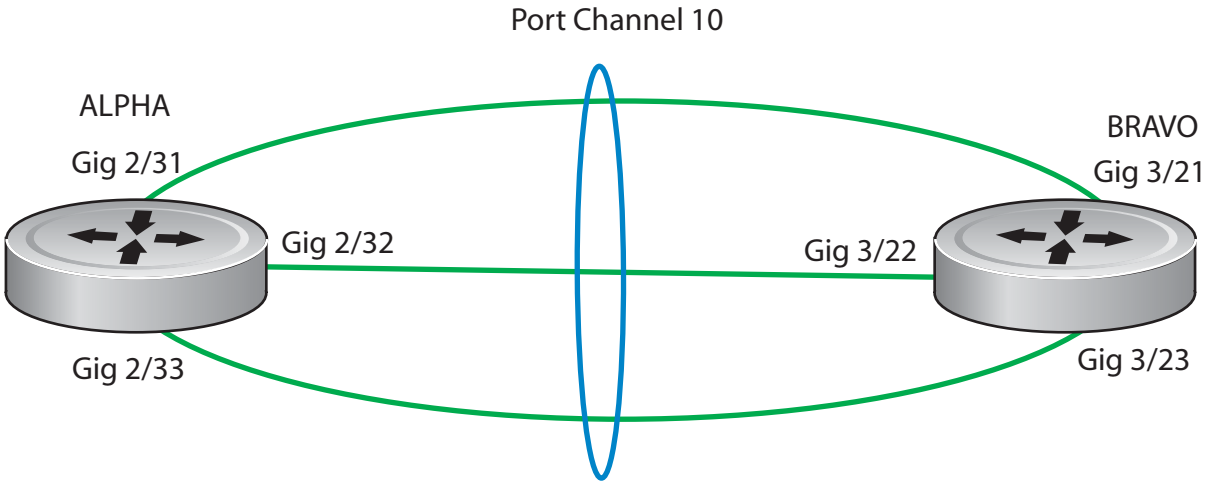
LACP Basic Configuration Example

The screenshots in this section are based on the example topology shown in Figure 24-11. Two routers are named ALPHA and BRAVO, and their hostname prompts reflect those names.

The sections are:

- Configuring a LAG on ALPHA on page 550
- Summary of the configuration on ALPHA on page 554
- Summary of the configuration on BRAVO on page 555

Figure 24-11. LACP Sample Topology



Configuring a LAG on ALPHA

Figure 24-12. Creating a LAG on ALPHA

```
Alpha(conf)#interface port-channel 10
Alpha(conf-if-po-10)#no ip address
Alpha(conf-if-po-10)#switchport
Alpha(conf-if-po-10)#no shutdown
Alpha(conf-if-po-10)#show config
!
interface Port-channel 10
 no ip address
 switchport
 no shutdown
!
Alpha(conf-if-po-10)#
```

Figure 24-13. Inspecting a LAG Port Configuration on ALPHA

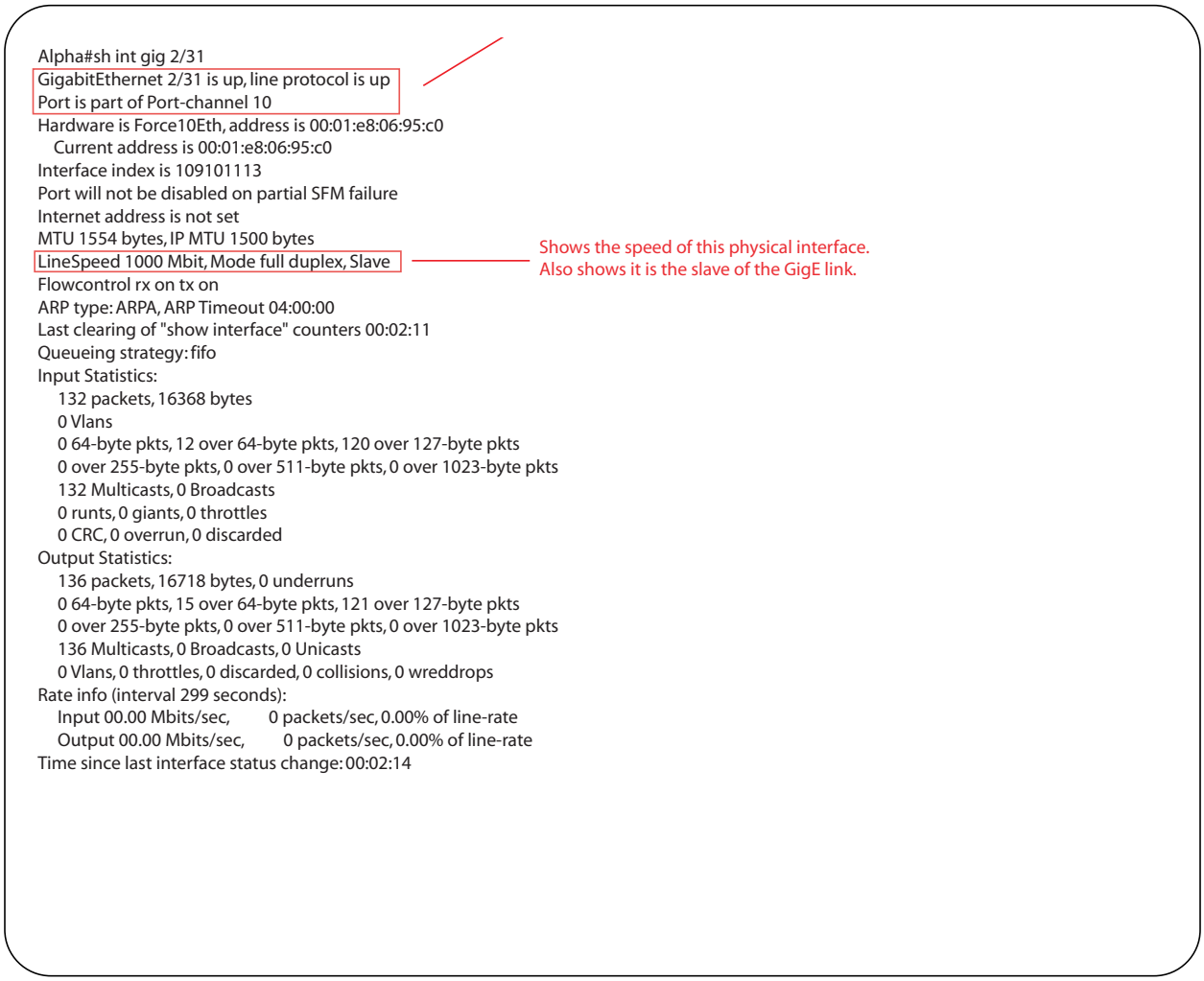


Figure 24-14. Inspecting Configuration of LAG 10 on ALPHA

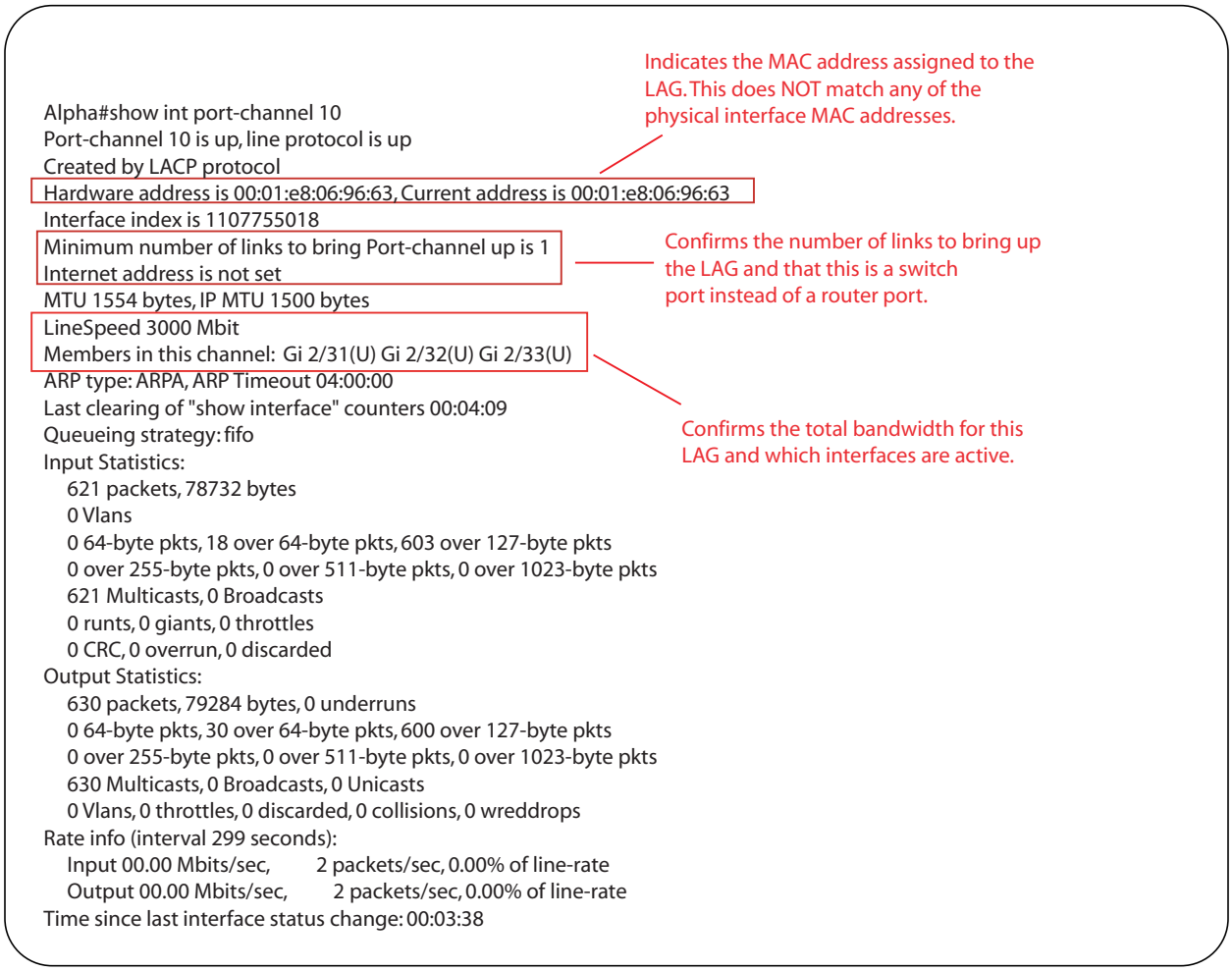
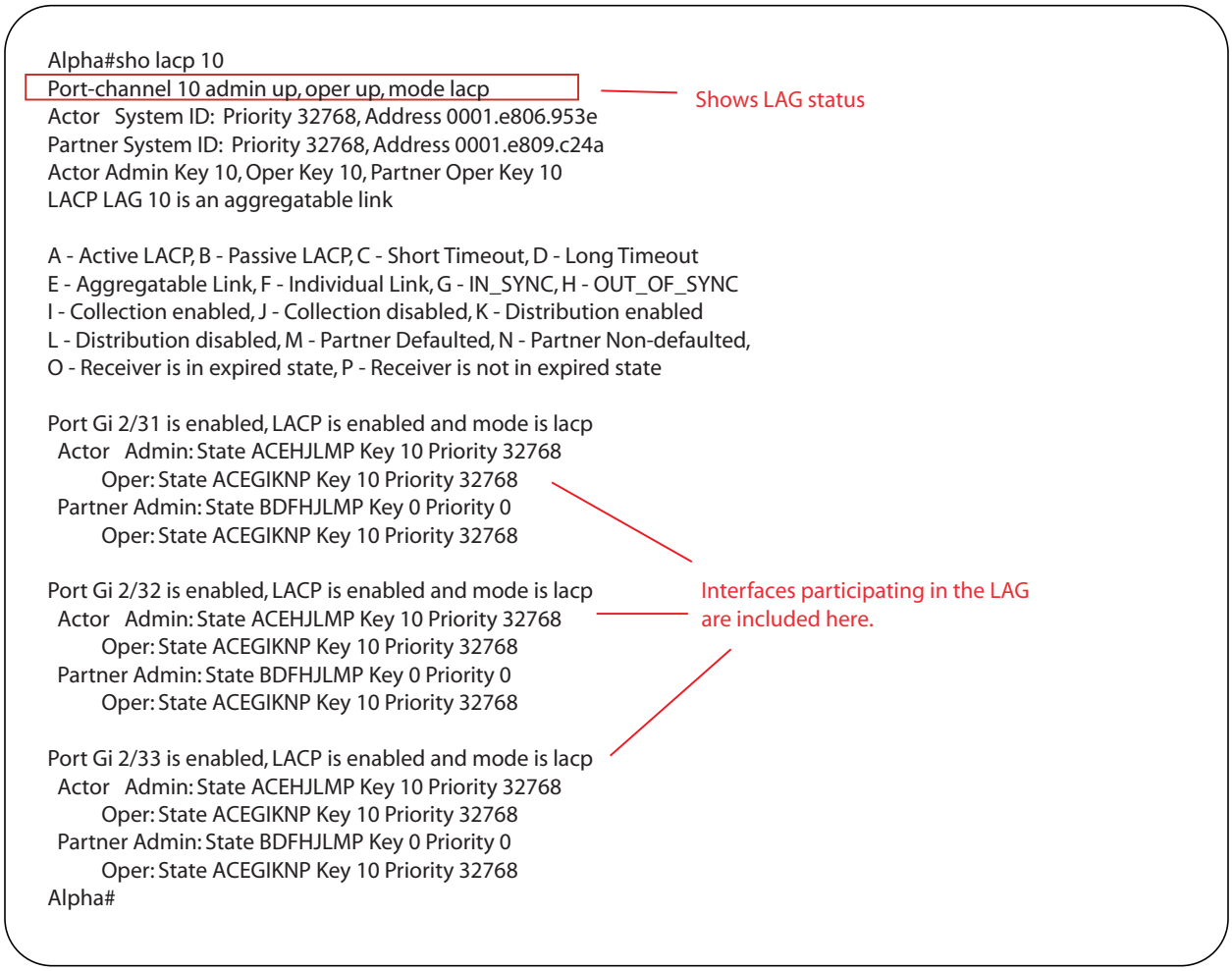


Figure 24-15. Using the show lacp Command to Verify LAG 10 Status on ALPHA



Summary of the configuration on ALPHA

Figure 24-16. Summary of the configuration on ALPHA

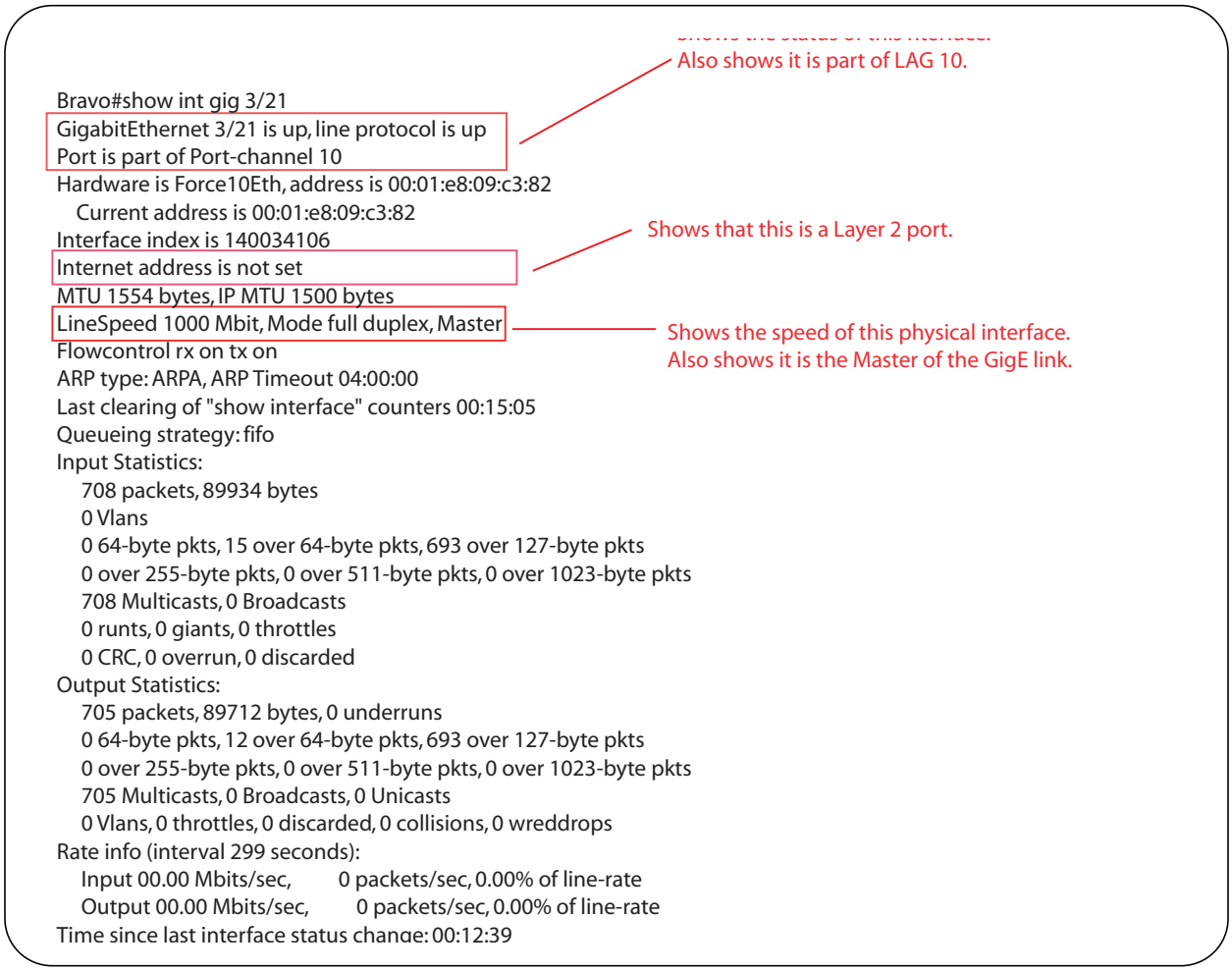
```
Alpha(conf-if-po-10)#int gig 2/31
Alpha(conf-if-gi-2/31)#no ip address
Alpha(conf-if-gi-2/31)#no switchport
Alpha(conf-if-gi-2/31)#shutdown
Alpha(conf-if-gi-2/31)#port-channel-protocol lacp
Alpha(conf-if-gi-2/31-lacp)#port-channel 10 mode active
Alpha(conf-if-gi-2/31-lacp)#no shut
Alpha(conf-if-gi-2/31)#show config

!
interface GigabitEthernet 2/31
 no ip address
!
 port-channel-protocol LACP
  port-channel 10 mode active
 no shutdown
!
Alpha(conf-if-gi-2/31)#

interface Port-channel 10
no ip address
switchport
no shutdown

interface GigabitEthernet 2/31
no ip address
no switchport
switchport
port-channel-protocol LACP
port-channel 10 mode active
no shutdown
```

Summary of the configuration on BRAVO

Figure 24-17. Summary of the configuration on BRAVO

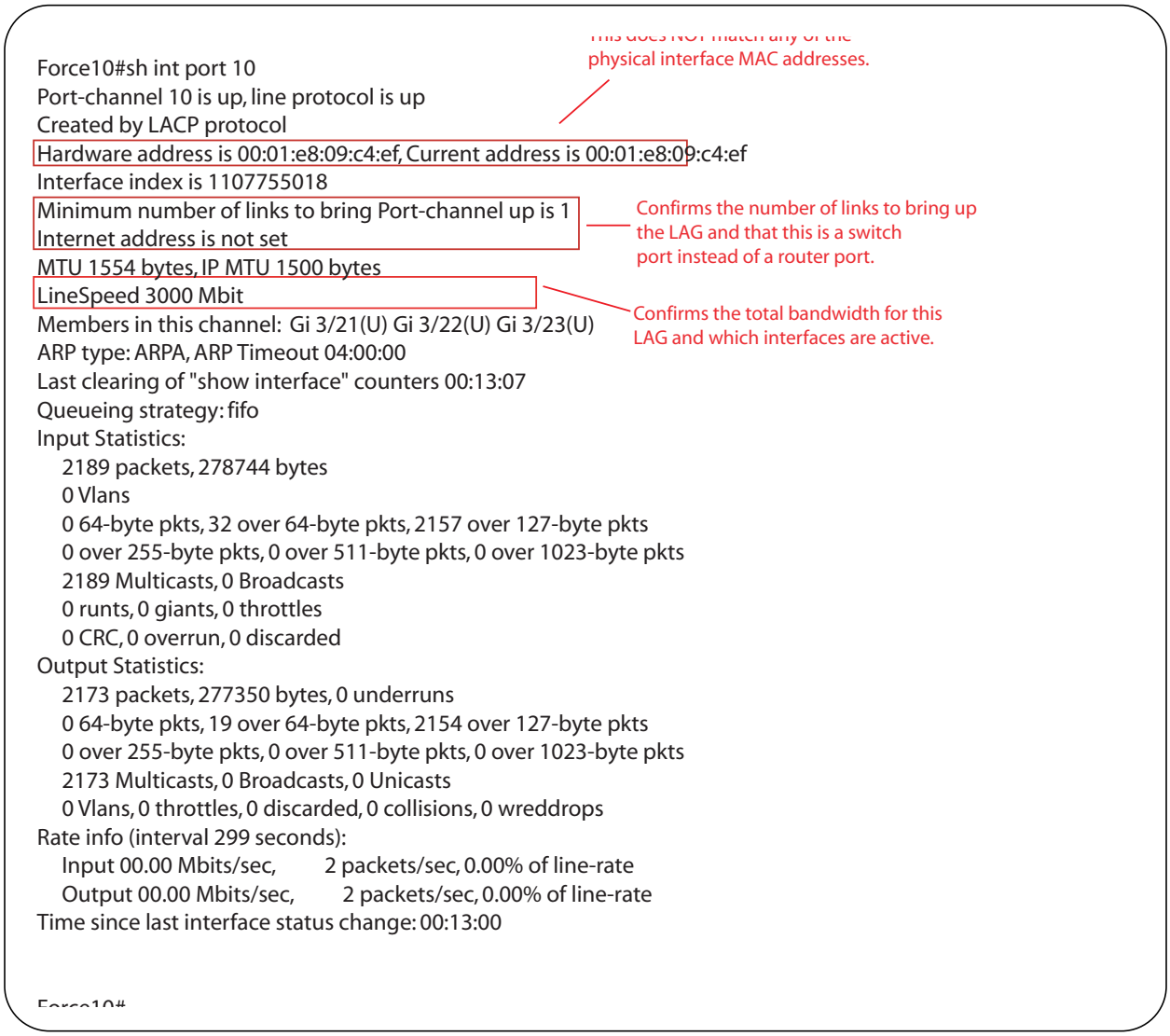
```
Bravo(conf-if-gi-3/21)#int port-channel 10
Bravo(conf-if-po-10)#no ip add
Bravo(conf-if-po-10)#switch
Bravo(conf-if-po-10)#no shut
Bravo(conf-if-po-10)#show config
!
interface Port-channel 10
 no ip address
 switchport
 no shutdown
!
Bravo(conf-if-po-10)#exit

Bravo(conf)#int gig 3/21
Bravo(conf)#no ip address
Bravo(conf)#no switchport
Bravo(conf)#shutdown
Bravo(conf-if-gi-3/21)#port-channel-protocol lacp
Bravo(conf-if-gi-3/21-lacp)#port-channel 10 mode active
Bravo(conf-if-gi-3/21-lacp)#no shut
Bravo(conf-if-gi-3/21)#end

!
interface GigabitEthernet 3/21
 no ip address
!
 port-channel-protocol LACP
  port-channel 10 mode active
 no shutdown
Bravo(conf-if-gi-3/21)#end

int port-channel 10
no ip address
switchport
no shutdown
show config

int gig 3/21
no ip address
no switchport
shutdown
port-channel-protocol lacp
port-channel 10 mode active
no shut
show config
end
```

Figure 24-18. Using the show interface Command to Inspect a LAG Port on BRAVO



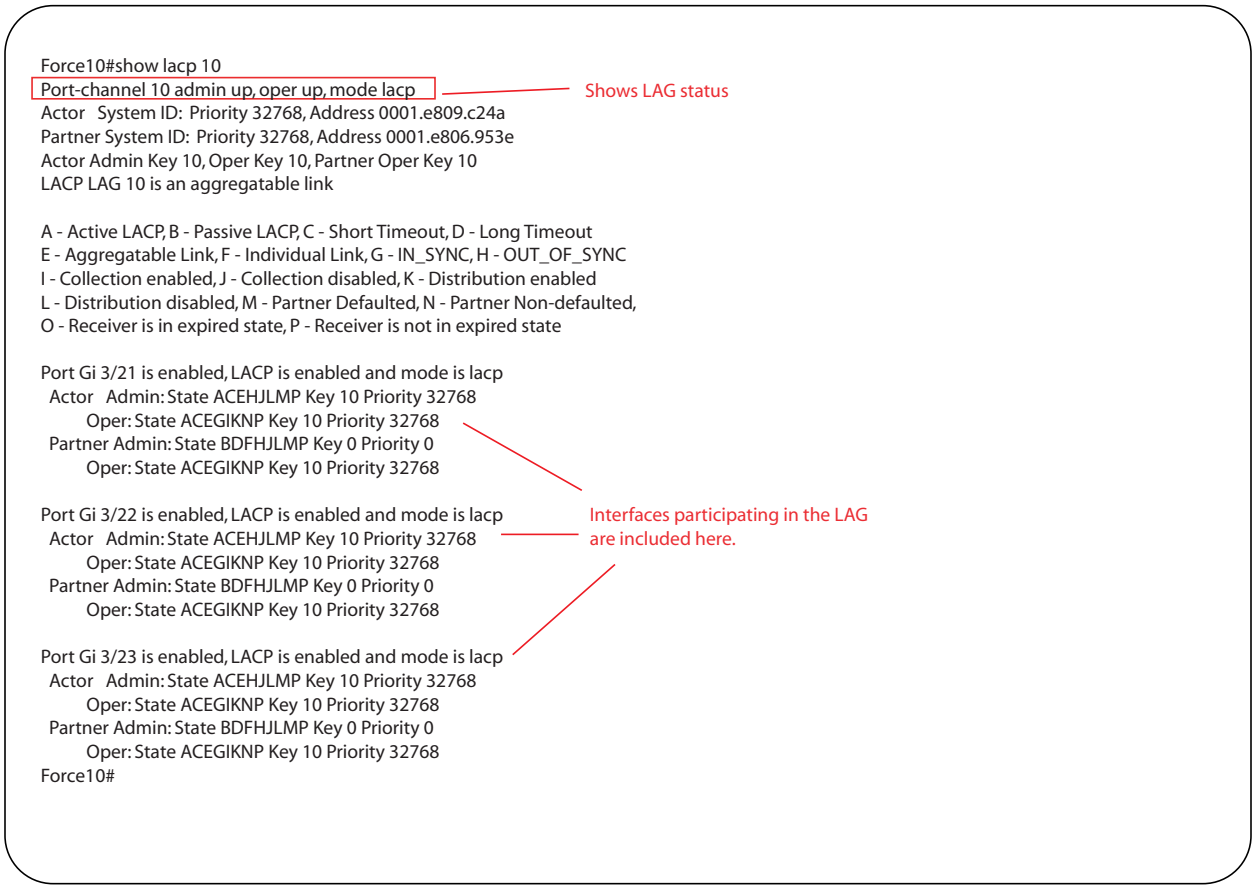

```
Bravo#show int gig 3/21
GigabitEthernet 3/21 is up, line protocol is up
Port is part of Port-channel 10
Hardware is Force10Eth, address is 00:01:e8:09:c3:82
   Current address is 00:01:e8:09:c3:82
Interface index is 140034106
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 1000 Mbit, Mode full duplex, Master
Flowcontrol rx on tx on
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:15:05
Queueing strategy: fifo
Input Statistics:
   708 packets, 89934 bytes
   0 Vlans
   0 64-byte pkts, 15 over 64-byte pkts, 693 over 127-byte pkts
   0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
   708 Multicasts, 0 Broadcasts
   0 runts, 0 giants, 0 throttles
   0 CRC, 0 overrun, 0 discarded
Output Statistics:
   705 packets, 89712 bytes, 0 underruns
   0 64-byte pkts, 12 over 64-byte pkts, 693 over 127-byte pkts
   0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
   705 Multicasts, 0 Broadcasts, 0 Unicasts
   0 Vlans, 0 throttles, 0 discarded, 0 collisions, 0 wreddrops
Rate info (interval 299 seconds):
   Input 00.00 Mbits/sec,        0 packets/sec, 0.00% of line-rate
   Output 00.00 Mbits/sec,       0 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:12:39
```

Figure 24-19. Using the show interfaces port-channel Command to Inspect LAG 10




```
Force10#sh int port 10
Port-channel 10 is up, line protocol is up
Created by LACP protocol
Hardware address is 00:01:e8:09:c4:ef, Current address is 00:01:e8:09:c4:ef
Interface index is 1107755018
Minimum number of links to bring Port-channel up is 1
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 3000 Mbit
Members in this channel: Gi 3/21(U) Gi 3/22(U) Gi 3/23(U)
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:13:07
Queueing strategy: fifo
Input Statistics:
    2189 packets, 278744 bytes
    0 Vlans
    0 64-byte pkts, 32 over 64-byte pkts, 2157 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    2189 Multicasts, 0 Broadcasts
    0 runts, 0 giants, 0 throttles
    0 CRC, 0 overrun, 0 discarded
Output Statistics:
    2173 packets, 277350 bytes, 0 underruns
    0 64-byte pkts, 19 over 64-byte pkts, 2154 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    2173 Multicasts, 0 Broadcasts, 0 Unicasts
    0 Vlans, 0 throttles, 0 discarded, 0 collisions, 0 wreddrops
Rate info (interval 299 seconds):
    Input 00.00 Mbits/sec,        2 packets/sec, 0.00% of line-rate
    Output 00.00 Mbits/sec,       2 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:13:00


Force10#
```

Figure 24-20. Using the show lacp Command to Inspect LAG Status



PPP is a connection-oriented protocol that enables layer two links over a variety of different physical layer connections. It is supported on both synchronous and asynchronous lines, and can operate in half-duplex or full-duplex mode. It was designed to carry IP traffic but is general enough to allow any type of network layer datagram to be sent over a PPP connection. As its name implies, it is for point-to-point connections between exactly two devices, and assumes that frames are sent and received in the same order.

25

Layer 2

Layer 2 features are supported on platforms C E S

The E-Series ExaScale platform is supported with FTOS 8.1.1.0 and later.

This chapter describes the following Layer 2 features:

Managing the MAC Address Table

FTOS provides the following management activities for the MAC address table:

Clear the MAC Address Table

You may clear the MAC address table of dynamic entries:

Task	Command Syntax	Command Mode
Clear a MAC address table of dynamic entries. • **address** deletes the specified entry • **all** deletes all dynamic entries • **interface** deletes all entries for the specified interface • **vlan** deletes all entries for the specified VLAN	**clear mac-address-table dynamic** {*address* | **all | interface | vlan**}	EXEC Privilege

Set the Aging Time for Dynamic Entries

Learned MAC addresses are entered in the table as dynamic entries, which means that they are subject to aging. For any dynamic entry, if no packet arrives on the switch with the MAC address as the source or destination address within the timer period, the address is removed from the table. The default aging time is 1800 seconds.

Task	Command Syntax	Command Mode
Disable MAC address aging for all dynamic entries.	**mac-address-table aging-time 0**	CONFIGURATION
Specify an aging time.	**mac-address-table aging-time** *seconds* Range: 10-1000000	CONFIGURATION

Set the Aging Time for Dynamic Entries on a VLAN

Set the Aging Time for Dynamic Entries on a VLAN is available only on platform: E

Task	Command Syntax	Command Mode
Specify an aging time.	**mac-address-table aging-time** *seconds* Range: 1-1000000	INTERFACE VLAN

FTOS Behavior: The time elapsed before the configured MAC aging time expires is not precisely as configured. For example, the VLAN configuration **mac-address-table aging-time 1**, does not remove dynamic entries from the CAM after precisely 1 second. The actual minimum aging time for entries is approximately 5 seconds because this is the default MAC address table scanning interval. Therefore, MAC aging configurations of less than 5 seconds, as in this example, might be ineffective. Configuring **mac-address-table station-move time-interval 500**, solves this limitation. Reducing the scanning interval to the minimum, 500 milliseconds, increases the detection speed, which results in FTOS clearing entries closer to the actual desired aging time.

Configure a Static MAC Address

A static entry is one that is not subject to aging. Static entries must be entered manually:

Task	Command Syntax	Command Mode
Create a static MAC address entry in the MAC address table.	**mac-address-table static**	CONFIGURATION

Display the MAC Address Table

To display the contents of the MAC address table:

Task	Command Syntax	CommandMode
Display the contents of the MAC address table. • **address** displays the specified entry. • **aging-time** displays the configured aging-time. • **count** displays the number of dynamic and static entries for all VLANs, and the total number of entries. • **dynamic** displays only dynamic entries • **interface** displays only entries for the specified interface. • **static** displays only static entries. • **vlan** displays only entries for the specified VLAN.	**show mac-address-table** [**address** | **aging-time** [**vlan** *vlan-id*]| **count** | **dynamic** | **interface** | **static | vlan**]	EXEC Privilege

MAC Learning Limit

This section has the following sub-sections:

MAC Address Learning Limit is a method of port security on Layer 2 physical, port-channel, and VLAN interfaces. It enables you to set an upper limit on the number of MAC addresses learned on an interface/ VLAN. After the limit is reached, the system drops all traffic from a device with an unlearned MAC address.

FTOS Behavior: When configuring MAC Learning Limit on a port or VLAN the configuration is accepted (becomes part of running-config and **show mac learning-limit interface**) before the system verifies that sufficient CAM space exists. If the CAM check fails, the a message is displayed:

```
%E90MH:5 %ACL_AGENT-2-ACL_AGENT_LIST_ERROR: Unable to apply  access-list
Mac-Limit  on GigabitEthernet 5/84
```

In this case, the configuration is still present in the running-config and **show** output. Remove the configuration before re-applying a MAC learning limit with lower value. Also, ensure that Syslog messages can be viewed on your session.

Note: The CAM-check failure message beginning in FTOS version 8.3.1.0 is different from versions 8.2.1.1 and earlier, which read:

```
% Error: ACL returned error
% Error: Remove existing limit configuration if it was configured before
```

To set a MAC learning limit on an interface:

Task	Command Syntax	Command Mode
Specify the number of MAC addresses that the system can learn off a Layer 2 interface.	**mac learning-limit** *address_limit*	INTERFACE

Three options are available with the mac learning-limit command: dynamic, no-station-move, and station-move,

Note: An SNMP trap is available for **mac learning-limit station-move**. No other SNMP traps are available for MAC Learning Limit, including limit violations.

mac learning-limit dynamic

After you enable a MAC learning limit, MAC addresses learned on the port and entered in the MAC address table are static by default. If you configure the MAC learning **dynamic** option, learned MAC addresses are stored in the dynamic region of the table and are subject to aging. Entries created before this option is set are not affected.

On the C-Series and S-Series, the MAC address table is stored in the Layer 2 FIB region of CAM. The Layer 2 FIB region allocates space for static MAC address entries and dynamic MAC address entries.

On the E-Series, the MAC address table is stored in the Layer 2 ACL region. All MAC address entries on the E-Series are dynamic.

FTOS Behavior: If you do not configure the **dynamic** option, the C-Series and S-Series do not detect station moves in which a MAC address learned off of a MAC-limited port is learned on another port on same line card. Therefore, FTOS does not take any configured station-move violation action. When a MAC address is relearned on any other linecard (any line card except the one to which the original MAC-limited port belongs), the station-move is detected, and the system takes the configured the violation action.

mac learning-limit station-move

mac learning-limit station-move is available only on platforms: C S

The **station-move** option, allows a MAC address already in the table to be learned off of another interface. For example, if you disconnect a network device from one interface and reconnect it to another interface, the MAC address is learned on the new interface. When the system detects this "station move," the system clears the entry learned on the original interface, and installs a new entry on the new interface.

mac learning-limit no-station-move

Note: Sticky MAC is not supported on the S25 or S50 in FTOS release 8.4.2.6.

The **no-station-move** option, also known as "sticky MAC," provides additional port security by preventing a station move. When this option is configured, the first entry in the table is maintained instead of creating a new entry on the new interface. **no-station-move** is the default behavior. Entries created before this option is set are not affected.

FTOS Behavior: The C-Series and S-Series do not generate a station-move violation log entry for physical interfaces or port-channels when you configure **mac learning-limit** or when you configure **mac learning-limit station-move-violation log**. FTOS detects a station-move violation only when you configure **mac learning-limit dynamic**, and logs the violation only when you configure the **mac learning-limit station-move-violation log**, as shown below:

```
FTOS(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
 no ip address
 switchport
 mac learning-limit 1 dynamic no-station-move
 mac learning-limit station-move-violation log
 no shutdown
```

mac learning-limit sticky

The sticky-MAC learning feature is supported on platforms: C S.

Note: Sticky MAC is not supported on the S25 or S50 in FTOS release 8.4.2.6.

You can provide security for the dynamically-learned MAC addresses of trusted devices that are allowed to access a port by configuring the **sticky** option. This MAC learning option allows a switch to maintain the association of a trusted MAC address with a port and prevents a device from accessing the switch on another interface until the option is disabled.

Trusted MAC addresses are added to the running configuration and "stick" to the port on which they are learned even if an interface goes down and comes back up. If you save sticky MAC addresses to the start-up configuration file by entering the **write config** command, the addresses are deleted from the running-configuration, do not have to be dynamically relearned, and do not change when the switch reboots. Any sticky MAC addresses learned after the **write config** is performed are not saved after a reboot.

The sticky MAC address option is supported on physical port and port-channel interfaces; it is not supported on VLAN interfaces.

Static MAC addresses have a higher preference than sticky MAC addresses and are therefore not converted with sticky-MAC learning.

FTOS Behavior: The following conditions apply when you enable the sticky-MAC address option for MAC learning on an interface:

- When you enable the sticky MAC learning option, all dynamically-learned MAC addresses that you save to the start-up configuration are converted to statically-configured MAC addresses when you reboot the switch.
- When the switch reboots, the interface deletes all previously learned dynamic MAC addresses and regenerates the list of sticky MAC addresses in the running-configuration with statically-configured and newly learned dynamic MAC addresses. During the new session, MAC addresses that are dynamically-learned are automatically converted to sticky MAC addresses until the configured limit is reached.
- The aging out of dynamically-learned MAC addresses on the interface is disabled and restarts only when you disable the **sticky** option.
- A "station move" is not supported and a trusted MAC address in the table cannot be learned off another interface.
    - The list of sticky MAC addresses is converted back to their former dynamic addresses.
    - New dynamically-learned MAC addresses are no longer converted to sticky MAC addresses.
- After a line card reset, a port or port-channel interface enabled for sticky-MAC learning dynamically learns the MAC addresses of devices attached to ports on other line cards only if the attached devices are transmitting continuous traffic on default VLAN 1.

Note: A Sticky MAC configuration limits the number of MAC addresses that can be learned on a port/port-channel interface. Because a trunk port receives trusted MAC addresses not from a single user or VLAN but from multiple users and VLANs, it is difficult to specify the exact MAC address limit on a trunk port. As a result, traffic from MAC addresses that exceed the limit may be dropped.

It is recommended that you configure the sticky MAC option only on an access port, which is directly connected to a host and on which you want to limit the number of learned MAC addresses.

It is not recommended that you configure sticky MAC learning on inter-bridge ports. If there is a topology change, traffic may be blocked because a sticky MAC address cannot be moved across the ports.

To enable and display sticky MAC address learning on a Layer 2 physical port or port-channel interface, enter the following commands:

Task	Command Syntax	Command Mode
Converts dynamically-learned MAC addresses to sticky MAC addresses to prevent trusted devices from moving to different interfaces.	**mac learning-limit** *address_limit* **sticky**	INTERFACE
Display the MAC addresses with sticky MAC address learning.	**show mac-address-table**	EXEC EXEC Privilege

Displaying MAC Learning-Limited Interfaces

To display a list of all interfaces with a MAC learning limit:

Task	Command Syntax	Command Mode
Display a list of all interfaces with a MAC learning limit.	**show mac learning-limit**	EXEC Privilege

Learning Limit Violation Actions

Learning Limit Violation Actions are supported only on platform: E

You can configure the system to take an action when the MAC learning limit is reached on an interface and a new address is received using one of the following options with the **mac learning-limit** command:

Task	Command Syntax	Command Mode
Generate a system log message when the MAC learning limit is exceeded.	**learn-limit-violation log**	INTERFACE
Shut down the interface and generate a system log message when the MAC learning limit is exceeded.	**learn-limit-violation shutdown**	INTERFACE

Station Move Violation Actions

Station Move Violation Actions are supported only on platform: E C S

Note: On a C-Series or S-Series switch, Station Move Violation actions are supported on interfaces on different line cards; they are not supported on interfaces on the same line card.

no-station-move is the default behavior (see mac learning-limit no-station-move on page 564). You can configure the system to take an action if a station move occurs using one the following options with the **mac learning-limit** command:.

Task	Command Syntax	Command Mode
Generate a system log message indicating a station move.	**station-move-violation log**	INTERFACE
Shut down the first port to learn the MAC address.	**station-move-violation shutdown-original**	INTERFACE
Shut down the second port to learn the MAC address.	**station-move-violation shutdown-offending**	INTERFACE
Shut down both the first and second port to learn the MAC address.	**station-move-violation shutdown-both**	INTERFACE

To display a list of interfaces configured with MAC learning limit or station move violation actions:

Task	Command Syntax	Command Mode
Display a list of all of the interfaces configured with MAC learning limit or station move violation.	**show mac learning-limit violate-action**	CONFIGURATION

Recovering from Learning Limit and Station Move Violations

After a learning-limit or station-move violation shuts down an interface, you must manually reset it:

Task	Command Syntax	Command Mode
Reset interfaces in ERR_Disabled state caused by a learning limit violation or station move violation.	**mac learning-limit reset**	CONFIGURATION
Reset interfaces in ERR_Disabled state caused by a learning limit violation.	**mac learning-limit reset learn-limit-violation** [**interface | all**]	CONFIGURATION
Reset interfaces in ERR_Disabled state caused by a station move violation.	**mac learning-limit reset station-move-violation** [**interface | all**]	CONFIGURATION

Note: You can also reset an interface by shutting it down with the **shutdown** command, and then reenabling it with the **no shutdown** command.

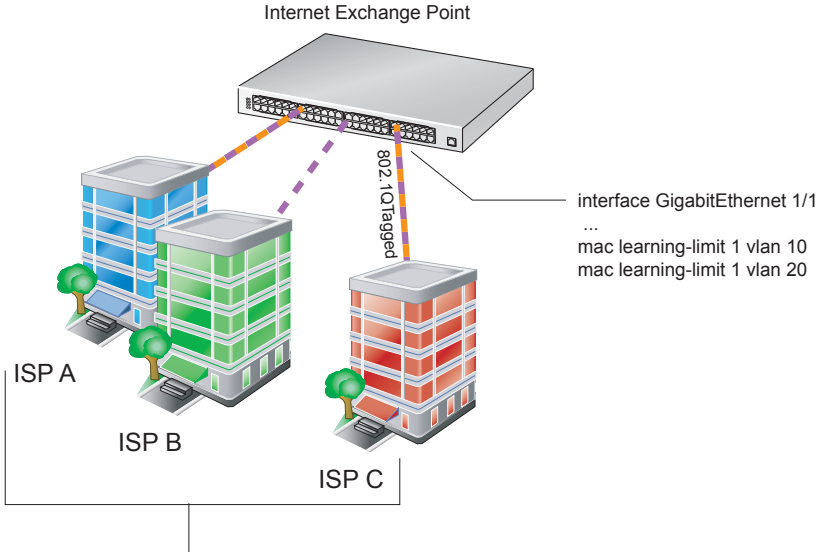
Per-VLAN MAC Learning Limit

Per-VLAN MAC Learning Limit is available only on platform: E

An individual MAC learning limit can be configured for each VLAN using Per-VLAN MAC Learning Limit.

One application of Per-VLAN MAC Learning Limit is on access ports. In Figure 25-1, an Internet Exchange Point (IXP) connects multiple Internet Service Provider (ISP). An IXP can provide several types of services to its customers including public an private peering. Public peering means that all customers are connected to one VLAN, and if one ISP wants to peer with another ISP, it establishes a BGP peering session over this VLAN. Private Peering means that the IXP sets up a separate VLAN between two customers that want to peer privately; only the ports of these two ISPs would belong to this VLAN, and they would peer via BGP. In Figure 25-1, Per-VLAN MAC Learning Limit is used on the access ports for the ISPs that have subscribed to private and public peering since these access ports are members of multiple VLANs.

Figure 25-1. Per-VLAN MAC Learning Limit



Task	Command Syntax	Command Mode
Configure a MAC learning limit on a VLAN.	**mac learning-limit** *limit* **vlan** *vlan-id*	INTERFACE
Display the MAC learning limit counters for a VLAN.	**show mac learning-limit** [**interface** *slot/port* [**vlan** *vlan-id*]]	EXEC Privilege

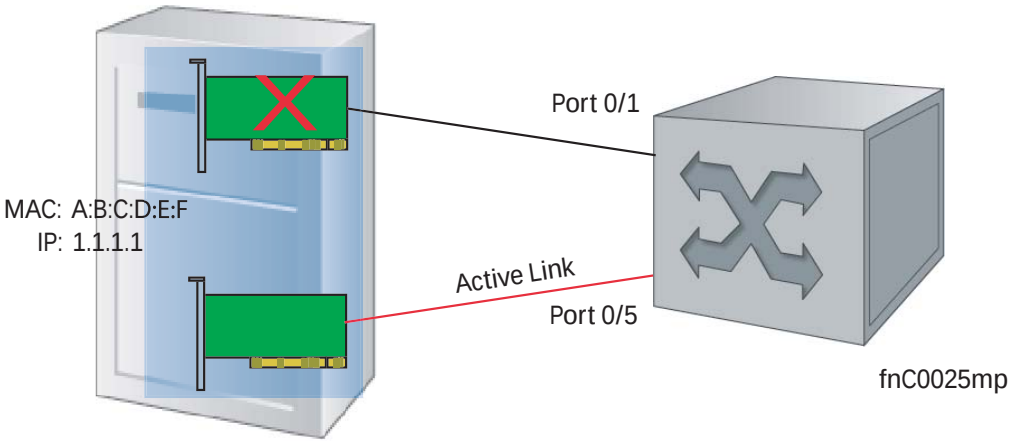
```
FTOS#show mac learning-limit
Interface      Vlan      Learning       Dynamic        Static         Unknown SA
Slot/port      Id        Limit          MAC count      MAC count      Drops
Gi 5/84        2         2                    0                0              0
Gi 5/84        *         5                    0                0              0
Gi 5/85        3         3                    0                0              0
Gi 5/85        *         10                   0                0              0
FTOS#show mac learning-limit interface gig 5/84
Interface      Vlan      Learning       Dynamic        Static         Unknown SA
Slot/port      Id        Limit          MAC count      MAC count      Drops
Gi 5/84        2         2                    0                0              0
Gi 5/84        *         5                    0                0              0
FTOS#show mac learning-limit interface gig 5/84 vlan 2
Interface      Vlan      Learning       Dynamic        Static         Unknown SA
Slot/port      Id        Limit          MAC count      MAC count      Drops
Gi 5/84        2         2                    0                0              0
```

NIC Teaming

NIC teaming is a feature that allows multiple network interface cards in a server to be represented by one MAC address and one IP address in order to provide transparent redundancy, balancing, and to fully utilize network adapter resources.

Figure 25-2 shows a topology where two NICs have been teamed together. In this case, if the primary NIC fails, traffic switches to the secondary NIC, since they are represented by the same set of addresses.

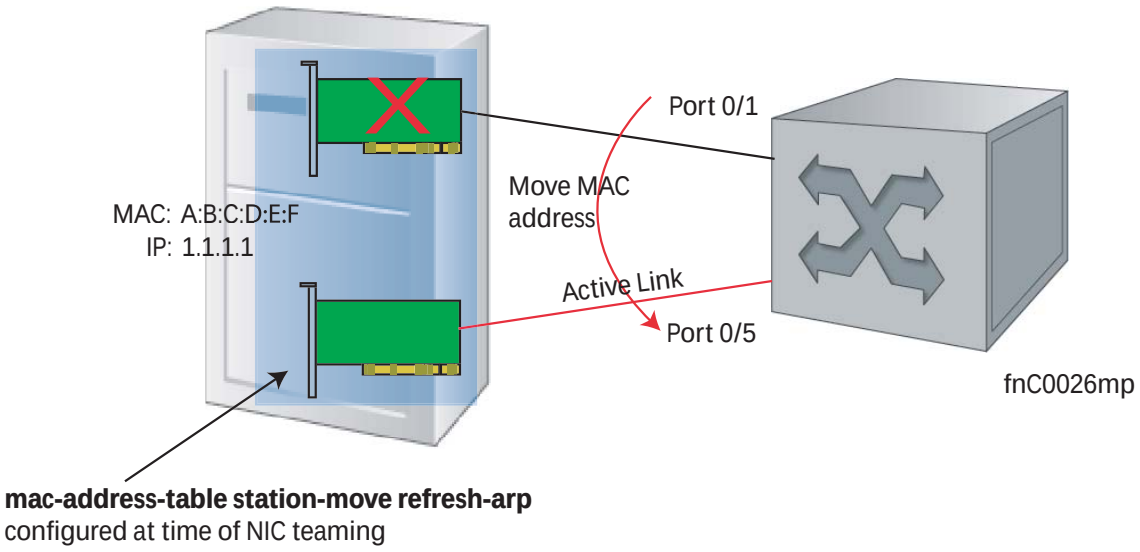
Figure 25-2. Redundant NICs with NIC Teaming



When NIC teaming is employed, consider that the server MAC address is originally learned on Port 0/1 of the switch (Figure 25-3). When the NIC fails, the same MAC address is learned on Port 0/5 of the switch. The MAC address must be disassociated with the one port and re-associated with another in the ARP table; in other words, the ARP entry must be “moved”. To ensure that this happens, you must configure the command **mac-address-table station-move refresh-arp** on the Dell Force10 switch at the time that NIC teaming is being configured on the server.

Note: If this command is not configured, traffic continues to be forwarded to the failed NIC until the ARP entry on the switch times out.

Figure 25-3. Configuring mac-address-table station-move refresh-arp Command



MAC Move Optimization

MAC Move Optimization is supported only on platform: E

Station-move detection takes 5000ms because this is the interval at which the detection algorithm runs. On the E-Series, you can reduce detection time to as little as 500ms using the command **mac-address-table station-move threshold time-interval** (though at the expense of CPU resources).

threshold is the number of times a station move must be detected in a single interval in order to trigger a system log message. For example, if you configure **mac-address-table station-move threshold 2 time-interval 5000**, and 4 station moves occur in 5000ms, then two log messages are generated.

Microsoft Clustering

Microsoft Clustering is supported only on platform: E

Microsoft Clustering allows multiple servers using Microsoft Windows to be represented by one MAC address and IP address in order to provide transparent failover or balancing. FTOS does not recognize server clusters by default; it must be configured to do so.

Default Behavior

When an ARP request is sent to a server cluster, either the active server or all of the servers send a reply, depending on the cluster configuration. If the active server sends a reply, the Dell Force10 switch learns the active server's MAC address. If all servers reply, the switch registers only the last received ARP reply, and the switch learns one server's actual MAC address (Figure 25-4); the virtual MAC address is never learned.

Since the virtual MAC address is never learned, traffic is forwarded to only one server rather than the entire cluster, and failover and balancing are not preserved (Figure 25-5).

Figure 25-4. Server Clustering: Multiple ARP Replies

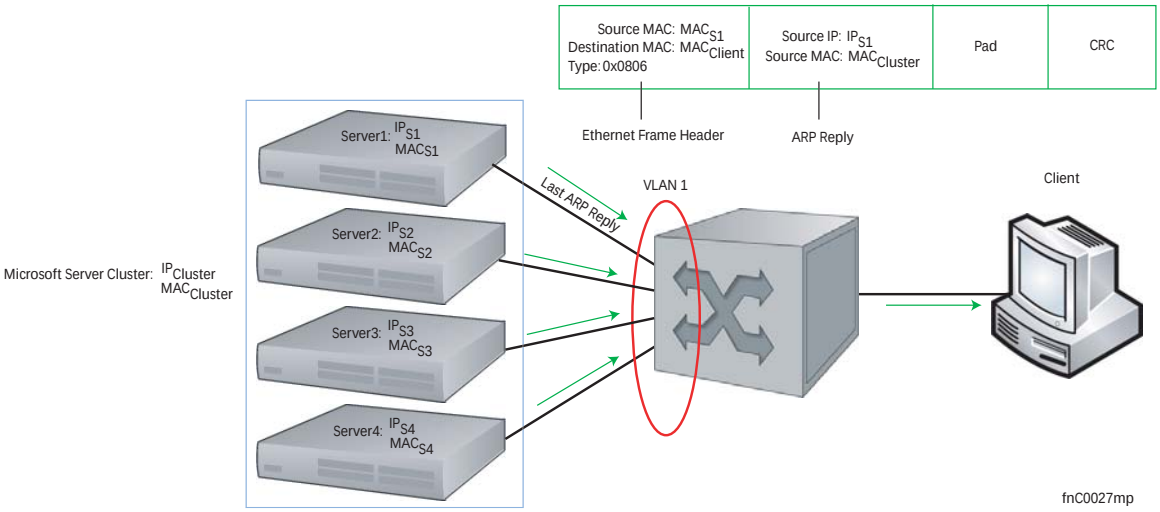
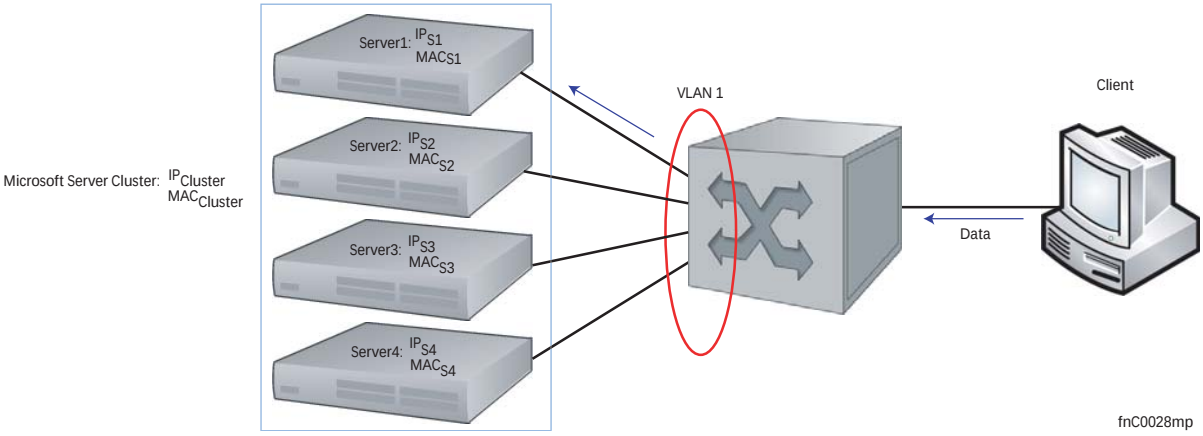


Figure 25-5. Server Clustering: Failover and Balancing Not Preserved

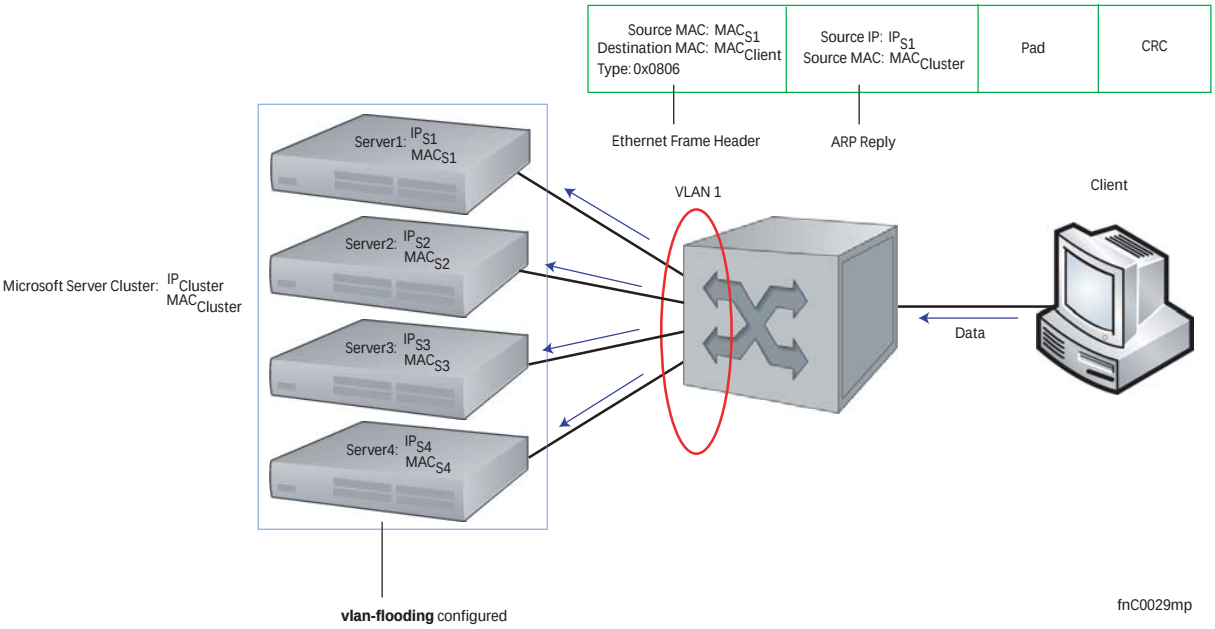


Configuring the Switch for Microsoft Server Clustering

To preserve failover and balancing, the Dell Force10 switch must learn the cluster's virtual MAC address, and it must forward traffic destined for the server cluster out all member ports in the VLAN connected to the cluster. To ensure that this happens, you must configure the command **ip vlan-flooding** on the Dell Force10 switch at the time that the Microsoft cluster is configured (Figure 25-6).

As shown in Figure 25-6, the server MAC address is given in the Ethernet frame header of the ARP reply, while the virtual MAC address representing the cluster is given in the payload. The **ip vlan-flooding** command directs the system to discover that there are different MAC addresses in an ARP reply and associate the virtual MAC address with the VLAN connected to the cluster. Then, all traffic destined for the cluster is flooded out of all member ports. Since all of the servers in the cluster receive traffic, failover and balancing are preserved.

Figure 25-6. Server Cluster: Failover and Balancing Preserved with the vlan-flooding Command



Enable and Disable VLAN Flooding

- ARP entries already resolved through the VLAN are deleted when the feature is enabled. This ensures that ARP entries across the VLAN are consistent.
- All ARP entries learned after the feature is enabled are deleted when the feature is disabled, and RP2 triggers ARP resolution. The feature is disabled with the command **no vlan-flooding**.
- When a port is added to the VLAN, the port automatically receives traffic if the feature is enabled. Old ARP entries are not deleted or updated.
- When a member port is deleted, its ARP entries are also deleted from the CAM.
- Port channels in the VLAN also receive traffic.
- There is no impact on the configuration from saving the configuration.
- The feature is not reflected in the output of the **show arp** command but is reflected in the output of the command **show ipf fib**.

The ARP entries exist in the secondary RPM CAM, so failover has no effect on the feature.

Configuring Redundant Pairs

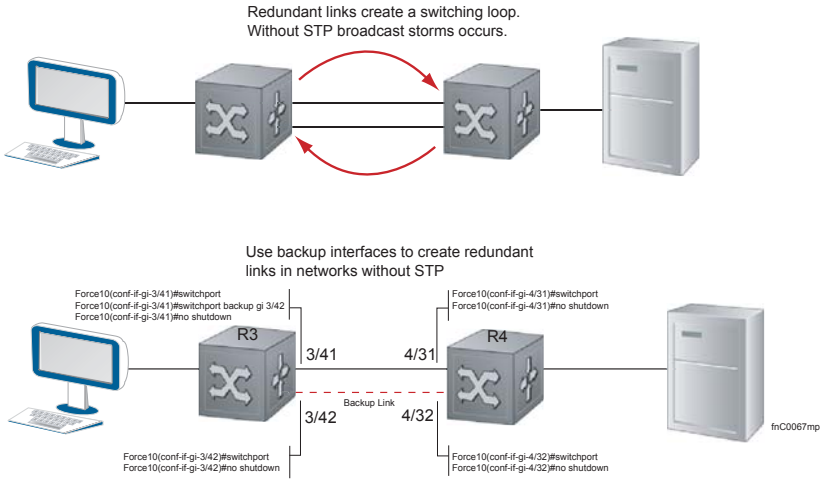
Configuring Redundant Pairs is supported:

- On physical interfaces on platforms C E S
- On static and dynamic port-channel interfaces on platforms C E S

The Redundant Pairs feature allows you to provide redundancy for Layer 2 links without using Spanning Tree (STP). You create redundant links by configuring pairs of Layer 2 (physical or port-channel) interfaces so that only one interface is up and carries user traffic at any time. Interfaces on either side of a link provide backup for the primary link. Redundant pairs are useful in service-provider and enterprise networks in which you do not run STP on switches—for example, networks with Digital Subscriber Line Access Mutiplexers (DSLAM)—to avoid creating switching loops (see Figure 25-7).

Note: For details on STP, see Chapter 52, "Spanning Tree Protocol," on page 1049.

Figure 25-7. Configuring Redundant Layer 2 Pairs without Spanning Tree



You configure a redundant pair by assigning a backup interface to a primary interface with the **switchport backup interface** command. Initially, the primary interface is active and transmits traffic and the backup interface remains down. If the primary fails for any reason, the backup transitions to an active UP state. If the primary interface fails and later comes back up, it remains as the backup interface for the redundant pair.

FTOS supports only Gigabit and 10-Gigabit ports and port channels as primary/backup interfaces in redundant pairs. (A port channel is also referred to as a Link Aggregation Group (LAG). See Port Channel Interfaces on page 428 for more information.)

In a redundant pair, any combination of physical and port-channel interfaces is supported as the two interfaces in a redundant pair. For example, you can configure a static (without LACP) or dynamic (with LACP) port-channel interface as either the primary or backup link in a redundant pair with a Gigabit interface.

To ensure that existing network applications see no difference when a primary interface in a redundant pair transitions to the backup interface, be sure to apply *identical* configurations of other traffic parameters to each interface.

If you remove an interface in a redundant link (remove the line card of a physical interface or delete a port channel with the **no interface port-channel** command), the redundant pair configuration is also removed.

Important Points about Configuring Redundant Pairs

- An interface cannot be used as a backup for more than one interface; an interface can have no more than one backup. A backup interface cannot have a backup interface.
- The active and standby/backup interfaces do *not* have to be of the same type (1G, 10G, etc).
- You may not enable any Layer 2 protocol on an interface in a redundant pair or on a port connected to a redundant interface.
- If the active interface in a redundant pair fails, you may have to synchronize the MAC addresses stored in the MAC and ARP tables to ensure that the backup interface can access the same MAC addresses that were learned on the failed active interface. To do so, enter the **mac-address-table station-move refresh-arp** command (global configuration mode).
- When you use a static or dynamic port channel as the active or backup interface in a redundant pair, the following conditions apply:
  - If you use two port-channel interfaces with different configurations in a redundant pair, traffic is transmitted in the same way following a transition to the backup interface. There is no difference in performance. For example, two port channels in a redundant pair can contain a different number and type of member ports or use different LACP modes.
  - There are no requirements on the number or type of links in a port channel.
  - There are no requirements on the location of the member links in a port channel, such as on the same line card or stacked device.
  - If you use a dynamic port channel in a redundant pair, LACP operation on the port channel is not affected by its status as the active or backup interface.
  - If you manually shut down a port channel that is the active interface in a redundant pair (**shutdown** command), the status of the port channel transitions to DOWN and the backup interface becomes active.
  - If the number of member links with an “oper up” status is less than the minimum number of required links configured for a port channel that is the active interface in a redundant pair, the status of the port channel transitions to DOWN and the backup interface becomes active.
  - If a static or dynamic port channel is used in a redundant pair, the port channel cannot be used as a member of a failover group in shared LAG state tracking (see in LACP chapter).
  - A dynamic port channel that is used in a redundant pair can participate in hitless LACP (see Configure LACP as Hitless in LACP chapter).

In Figure 25-8, interface 3/41 is a backup interface for 3/42, and 3/42 is DOWN as shown in message Message 1. If 3/41 fails, 3/42 transitions to the UP state, which makes the backup link active. A message similar to Message 1 appears whenever you configure a backup port.

Message 1 Configuring a Backup Layer 2 Port

```
02:28:04: %RPM0-P:CP %IFMGR-5-L2BKUP_WARN: Do not run any Layer2 protocols on Gi 3/41 and Gi
3/42
02:28:04: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Gi 3/42
02:28:04: %RPM0-P:CP %IFMGR-5-STATE_ACT_STBY: Changed interface state to standby: Gi 3/42
```

Figure 25-8. CLI for Configuring Redundant Layer 2 Pairs without Spanning Tree

```
FTOS(conf-if-range-gi-3/41-42)#switchport backup interface GigabitEthernet 3/42
FTOS(conf-if-range-gi-3/41-42)#show config
!
interface GigabitEthernet 3/41
 no ip address
 switchport
 switchport backup interface GigabitEthernet 3/42
 no shutdown
!
interface GigabitEthernet 3/42
 no ip address
 switchport
 no shutdown
FTOS(conf-if-range-gi-3/41-42)#
FTOS(conf-if-range-gi-3/41-42)#do show ip int brief | find 3/41
GigabitEthernet 3/41     unassigned      YES Manual up                    up
GigabitEthernet 3/42     unassigned      NO  Manual up                    down
[output omitted]
FTOS(conf-if-range-gi-3/41-42)#interface gig 3/41
FTOS(conf-if-gi-3/41)#shutdown
00:24:53: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Gi 3/41
FTOS(conf-if-gi-3/41)#00:24:55: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to
down: Gi 3/41
00:24:55: %RPM0-P:CP %IFMGR-5-INACTIVE: Changed Vlan interface state to inactive: Vl 1
00:24:55: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Gi 3/42
00:24:55: %RPM0-P:CP %IFMGR-5-ACTIVE: Changed Vlan interface state to active: Vl 1
00:24:55: %RPM0-P:CP %IFMGR-5-STATE_STBY_ACT: Changed interface state from standby to active:
Gi 3/42

FTOS(conf-if-gi-3/41)#do show ip int brief | find 3/41
GigabitEthernet 3/41     unassigned      NO  Manual administratively down down
GigabitEthernet 3/42     unassigned      YES Manual up                    up
[output omitted]

FTOS(conf-if-range-gi-0/23)#switchport backup interface port-channel 2
FTOS(conf-if-range-gi-0/23)#show config
!
interface GigabitEthernet 0/23
 no ip address
 switchport
 switchport backup interface port-channel 2
 shutdown
```

Restricting Layer 2 Flooding

Restricting Layer 2 Flooding is supported only on platform: E T

When Layer 2 multicast traffic must be forwarded on a VLAN that has multiple ports with different speeds on the same port-pipe, forwarding is limited to the speed of the slowest port. Restricted Layer 2 Flooding prevents slower ports from lowering the throughput of multicast traffic on faster ports by restricting flooding to ports with a speed equal to or above a link speed you specify.

For example, if a VLAN that has an (auto-negotiated) 100M port and a 1G port on the same port-pipe, and you enable Restricted Layer 2 Flooding with a minimum speed of 1G, multicast traffic is only flooded on the 1G port.

Enable Restricted Layer 2 Flooding using the command **restrict-flooding** from INTERFACE VLAN mode.

In combination with **restrict-flooding**, you can use the command **mac-flood-list** from CONFIGURATION mode, without the **min-speed** option, to allow some specific multicast traffic (identified using a MAC address range you specify) to be flooded on all ports regardless of the **restrict-flooding** configuration.

Conversely, if you want all multicast traffic to be flooded on all ports, but some specific traffic to be restricted, use **mac-flood-list** with the **min-speed** option, but without **restrict-flooding** configured. This configuration restricts flooding only for traffic with destination multicast MAC addresses within the multicast MAC address range you specify.

In Figure 25-9, flooding of unknown multicast traffic is restricted to 1G ports on VLAN100 using the command **restrict-flooding**. However, the command **mac-flood-list** allows traffic with MAC addresses 01:01:e8:00:00:00 to 01:01:e8:ff:ff:ff to be flooded on all ports regardless of link speed.

Figure 25-9. Restricting Layer 2 Multicast Flooding over Low Speed Ports

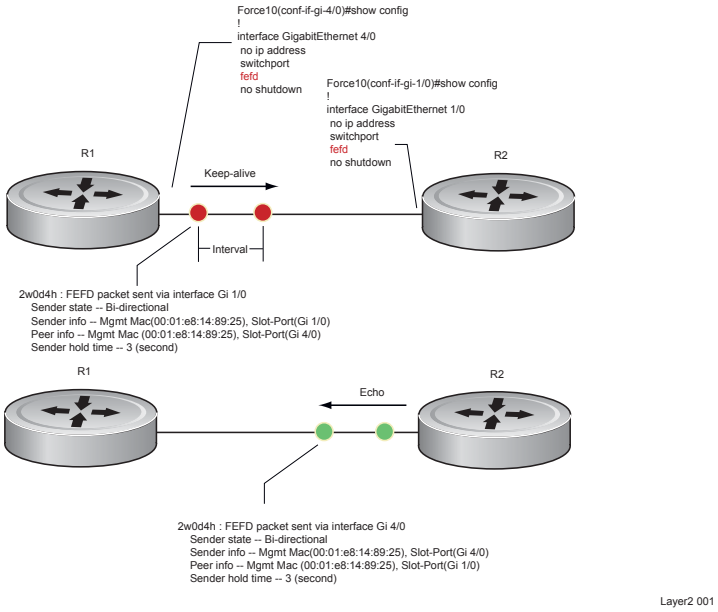
```
FTOS(conf)#$1:01:e8:00:00:00 ff:ff:ff:00:00:00 vlan 100-200,300
FTOS#show run | find mac-flood-list
mac-flood-list 01:01:e8:00:00:00 ff:ff:ff:00:00:00 vlan 100-200,300
[output omitted]
FTOS(conf)#interface vlan 100
FTOS(conf-if-vl-100)#restrict-flooding multicast min-speed 1000
FTOS(conf-if-vl-100)#show config
!
interface Vlan 100
restrict-flooding multicast min-speed 1000
 no shutdown
FTOS(conf-if-vl-100)#
```

Far-end Failure Detection

Far-end Failure Detection is supported only on platform: E

Far-end Failure Detection (FEFD) is a protocol that senses remote data link errors in a network. It responds by sending a unidirectional report that triggers an echoed response after a specified time interval.

Figure 25-10. Configuring Far-end Failure Detection



The report consists of several packets in SNAP format that are sent to the nearest known MAC address.

In the event of a far-end failure, the device stops receiving frames, and after the specified time interval assumes that the far-end is not available. The connecting line protocol is brought down so that upper layer protocols can detect the neighbor unavailability faster.

FEFD state changes

FEFD enabled systems (comprised of one or more interfaces) will automatically switch between four different modes: Idle, Unknown, Bi-directional, and Err-disabled.

1. An interface on which FEFD is not configured is in Idle state.
2. Once FEFD is enabled on an interface, it transitions to the Unknown state, and sends an FEFD packet to the remote end of the link.
3. When the local interface receives the echoed packet from the remote end, the local interface transitions to the Bi-directional state.
4. If the FEFD enabled system is configured to use FEFD in Normal mode and neighboring echoes are not received after three intervals, (each interval can be set between 3 and 300 seconds by the user) the state changes to unknown.

5. If the FEFD system has been set to Aggressive mode and neighboring echoes are not received after three intervals, the state changes to Err-disabled. All interfaces in the Err-disabled state must be manually reset using the **fefd reset** [*interface*] command in EXEC privilege mode (it can be done globally or one interface at a time) before the FEFD enabled system can become operational again.

Table 25-1. State Changes When Configuring FEFD

Local Event	Mode	Local State	Remote State	Local Admin Status	Local Protocol Status	Remote Admin Status	Remote Protocol Status
Shutdown	Normal	Admin Shutdown	Unknown	Down	Down	Up	Down
Shutdown	Aggressive	Admin Shutdown	Err-disabled	Up	Down	Up	Down
FEFD enable	Normal	Bi-directional	Bi-directional	Up	Up	Up	Up
FEFD enable	Aggressive	Bi-directional	Bi-directional	Up	Up	Up	Up
FEFD + FEFD disable	Normal	Locally disabled	Unknown	Up	Down	Up	Down
FEFD + FEFD disable	Aggressive	Locally disabled	Err-disabled	Up	Down	Up	Down
Link Failure	Normal	Unknown	Unknown	Up	Down	Up	Down
Link Failure	Aggressive	Err-disabled	Err-disabled	Up	Down	Up	Down

Important Points to Remember

- FEFD enabled ports are subject to an 8 to 10 second delay during an RPM failover before becoming operational.
- FEFD can be enabled globally or on a per interface basis, interface FEFD configurations override global FEFD configurations.
- FTOS supports FEFD on physical Ethernet interfaces only, excluding the management interface.

Configuring FEFD

You can configure FEFD for all interfaces from CONFIGURATION mode, or on individual interfaces from INTERFACE mode.

Enable FEFD Globally

To enable FEFD globally on all interfaces enter the command **fefd-global** in CONFIGURATION mode.

Report interval frequency and mode adjustments can be made by supplementing this command as well.

Step	Task	Command Syntax	Command Mode
1	Setup two or more connected interfaces for Layer 2 or Layer 3 use	**ip address** *ip address*, **switchport**	INTERFACE

Step	Task	Command Syntax	Command Mode
2	Activate the necessary ports administratively	**no shutdown**	INTERFACE
3	Enable fefd globally	**fefd** {**interval** | **mode**}	CONFIGURATION

Entering the **show fefd** command in EXEC privilege mode displays information about the state of each interface.

Figure 25-11. Show FEFD global outputs

```
FTOS#show fefd
FEFD is globally 'ON', interval is 3 seconds, mode is 'Normal'.

INTERFACE      MODE           INTERVAL       STATE
                              (second)
Gi 1/0         Normal         3              Bi-directional
Gi 1/1         Normal         3              Admin Shutdown
Gi 1/2         Normal         3              Admin Shutdown
Gi 1/3         Normal         3              Admin Shutdown

FTOS#show run fefd
!
fefd-global mode normal
fefd-global interval 3
```

Enable FEFD on an Interface

Entering the command **fefd** in INTERFACE mode enables FEFD on a per interface basis. To change the FEFD mode, supplement the **fefd** command in INTERFACE mode by entering the command **fefd** [**mode** {**aggressive** | **normal**}].

To disable FEFD protocol on one interface, enter the command **fefd disable** in INTERFACE mode. Disabling an interface will shut down all protocols working on that interface's connected line, and will not delete your previous FEFD configuration which can be enabled again at any time.

Step	Task	Command Syntax	Command Mode
1	Setup two or more connected interfaces for Layer 2 or Layer 3 use	**ip address** *ip address*, **switchport**	INTERFACE
2	Activate the necessary ports administratively	**no shutdown**	INTERFACE
3	Enable FEFD on each interface	**fefd** {**disable | interval** | **mode**}	INTERFACE

Figure 25-12. FEFD enabled interface configuration

```
FTOS(conf-if-gi-1/0)#show config
!
interface GigabitEthernet 1/0
 no ip address
 switchport
 fefd mode normal
 no shutdown

FTOS(conf-if-gi-1/0)#do show fefd | grep 1/0
Gi 1/0         Normal         3              Unknown
```

Debugging FEFD

By entering the command **debug fefd events** in EXEC privilege mode, output is displayed whenever events occur that initiate or disrupt an FEFD enabled connection.

Figure 25-13. Debug FEFD events display

```
FTOS#debug fefd events
FTOS#config
FTOS(conf)#int gi 1/0
FTOS(conf-if-gi-1/0)#shutdown
2w1d22h: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Gi 1/0
FTOS(conf-if-gi-1/0)#2w1d22h : FEFD state on Gi 1/0 changed from ANY to Unknown
2w1d22h: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Gi 1/0
2w1d22h: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Gi 4/0
2w1d22h: %RPM0-P:CP %IFMGR-5-INACTIVE: Changed Vlan interface state to inactive: Vl 1
2w1d22h : FEFD state on Gi 4/0 changed from Bi-directional to Unknown
```

Entering the command **debug fefd packets** in EXEC privilege mode will provide output for each packet transmission over the FEFD enabled connection.

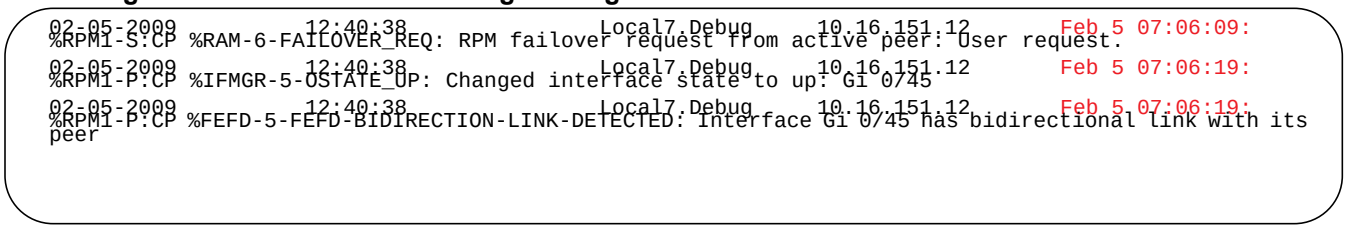
Figure 25-14. Debug FEFD packets display

```
FTOS#debug fefd packets
FTOS#2w1d22h : FEFD packet sent via interface Gi 1/0
    Sender state -- Bi-directional
    Sender info -- Mgmt Mac(00:01:e8:14:89:25), Slot-Port(Gi 1/0)
    Peer info -- Mgmt Mac (00:01:e8:14:89:25), Slot-Port(Gi 4/0)
    Sender hold time -- 3 (second)

2w1d22h : FEFD packet received on interface Gi 4/0
    Sender state -- Bi-directional
    Sender info -- Mgmt Mac(00:01:e8:14:89:25), Slot-Port(Gi 1/0)
    Peer info -- Mgmt Mac (00:01:e8:14:89:25), Slot-Port(Gi 4/0)
    Sender hold time -- 3 (second)
```

During an RPM Failover

In the event that an RPM failover occurs, FEFD will become operationally down on all enabled ports for approximately 8-10 seconds before automatically becoming operational again.

Figure 25-15. FEFD state change during an RPM failover



26

Link Layer Discovery Protocol

Link Layer Discovery Protocol is supported only on platforms: C E S

LLDP is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

This chapter contains the following sections:

802.1AB (LLDP) Overview

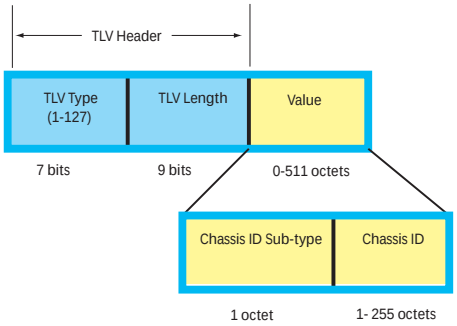
Link Layer Discovery Protocol (LLDP)—defined by IEEE 802.1AB—is a protocol that enables a LAN device to advertise its configuration and receive configuration information from adjacent LLDP-enabled LAN infrastructure devices. The collected information is stored in a management information base (MIB) on each device, and is accessible via SNMP.

Protocol Data Units

Configuration information is exchanged in the form of Type, Length, Value (TLV) segments. Figure 26-1 shows the Chassis ID TLV.

- **Type—**The kind of information included in the TLV
- **Length—**The value, in octets, of the TLV after the Length field
- **Value—**The configuration information that the agent is advertising

Figure 26-1. Type, Length, Value (TLV) Segment



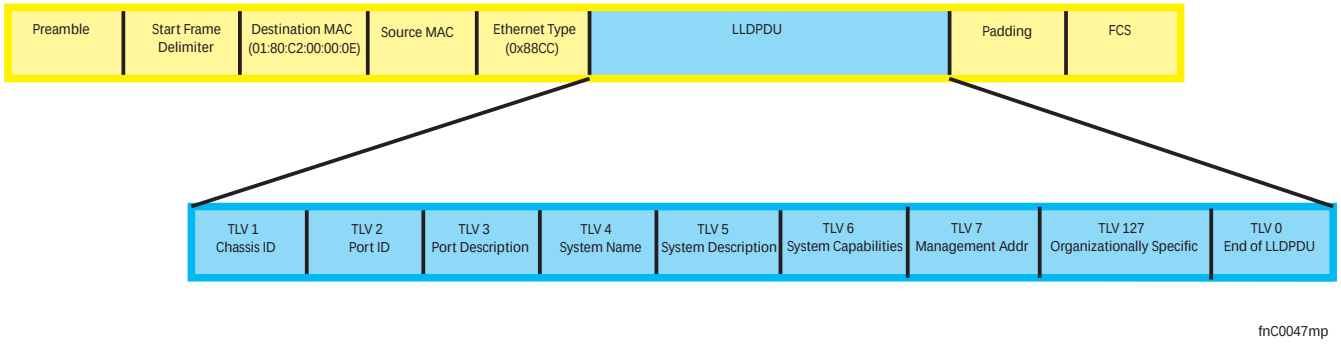
TLVs are encapsulated in a frame called an LLDP Data Unit (LLDPDU) (Figure 26-2), which is transmitted from one LLDP-enabled device to its LLDP-enabled neighbors. LLDP is a one-way protocol. LLDP-enabled devices (LLDP agents) can transmit and/or receive advertisements, but they cannot solicit and do not respond to advertisements.

There are five types of TLVs. All types are mandatory in the construction of an LLDPDU except Optional TLVs. The inclusion of individual Optional TLVs is user configurable.

Table 26-1. Type, Length, Value (TLV) Types

Type	TLV	Description
0	End of LLDPDU	Marks the end of an LLDPDU
1	Chassis ID	An administratively assigned name that identifies the LLDP agent
2	Port ID	An administratively assigned name that identifies a port through which TLVs are sent and received
3	Time to Live	A value that tells the receiving agent how long the information contained in the TLV Value field is valid
—	Optional	Includes sub-types of TLVs that advertise specific configuration information. These sub-types are Management TLVs, IEEE 802.1, IEEE 802.3, and TIA-1057 Organizationally Specific TLVs.

Figure 26-2. LLDPDU Frame



Optional TLVs

FTOS supports the following optional TLVs:

- Management TLVs
- IEEE 802.1 and 802.3 Organizationally Specific TLVs
- TIA-1057 Organizationally Specific TLVs

Management TLVs

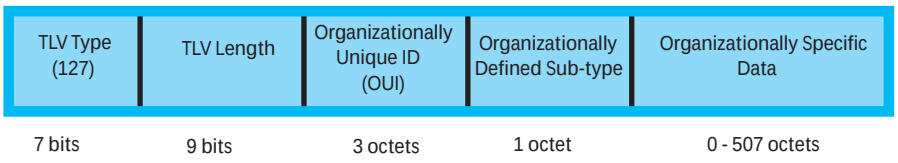
A Management TLV is an Optional TLVs sub-type. This kind of TLV contains essential management information about the sender. The five types are described in Table 26-2.

Organizationally Specific TLVs

Organizationally specific TLVs can be defined by a professional organization or a vendor. They have two mandatory fields (Figure 26-3) in addition to the basic TLV fields (Figure 26-1):

- Organizationally Unique Identifier (OUI)—a unique number assigned by the IEEE to an organization or vendor.
- OUI Sub-type—These sub-types indicate the kind of information in the following data field. The sub-types are determined by the owner of the OUI.

Figure 26-3. Organizationally Specific TLV



IEEE Organizationally Specific TLVs

Eight TLV types have been defined by the IEEE 802.1 and 802.3 working groups (Table 26-2) as a basic part of LLDP; the IEEE OUI is 00-80-C2. You can configure the Dell Force10 system to advertise any or all of these TLVs.

Table 26-2. Optional TLV Types

Type	TLV	Description
Optional TLVs		
4	Port description	A user-defined alphanumeric string that describes the port. FTOS does not currently support this TLV.
5	System name	A user-defined alphanumeric string that identifies the system.
6	System description	A user-defined alphanumeric string that describes the system

Table 26-2. Optional TLV Types

Type	TLV	Description
7	System capabilities	Identifies the chassis as one or more of the following: repeater, bridge, WLAN Access Point, Router, Telephone, DOCSIS cable device, end station only, or other
8	Management address	Indicates the network address of the management interface. FTOS does not currently support this TLV.
IEEE 802.1 Organizationally Specific TLVs		
127	Port-VLAN ID	On Dell Force10 systems, indicates the untagged VLAN to which a port belongs
127	Port and Protocol VLAN ID	On Dell Force10 systems, indicates the tagged VLAN to which a port belongs (and the untagged VLAN to which a port belongs if the port is in hybrid mode)
127	VLAN Name	Indicates the user-defined alphanumeric string that identifies the VLAN. This TLV is supported on C-Series only.
127	Protocol Identity	Indicates the protocols that the port can process. FTOS does not currently support this TLV.
IEEE 802.3 Organizationally Specific TLVs		
127	MAC/PHY Configuration/Status	Indicates the capability and current setting of the duplex status and bit rate, and whether the current settings are the result of auto-negotiation. This TLV is not available in the FTOS implementation of LLDP, but is available and mandatory (non-configurable) in the LLDP-MED implementation.
127	Power via MDI	Dell Force10 supports LLDP-MED protocol, which recommends that Power via MDI TLV be not implemented, and therefore Dell Force10 implements Extended Power via MDI TLV only.
127	Link Aggregation	Indicates whether the link is capable of being aggregated, whether it is currently in a LAG, and the port identification of the LAG. FTOS does not currently support this TLV.
127	Maximum Frame Size	Indicates the maximum frame size capability of the MAC and PHY

TIA-1057 (LLDP-MED) Overview

Link Layer Discovery Protocol—Media Endpoint Discovery (LLDP-MED)—as defined by ANSI/ TIA-1057— provides additional organizationally specific TLVs so that endpoint devices and network connectivity devices can advertise their characteristics and configuration information; the OUI for the Telecommunications Industry Association (TIA) is 00-12-BB.

- **LLDP-MED Endpoint Device**—any device that is on an IEEE 802 LAN network edge can communicate using IP and uses the LLDP-MED framework.
- **LLDP-MED Network Connectivity Device**—any device that provides access to an IEEE 802 LAN to an LLDP-MED endpoint device and supports IEEE 802.1AB (LLDP) and TIA-1057 (LLDP-MED). The Dell Force10 system is an LLDP-MED network connectivity device.

With regard to connected endpoint devices, LLDP-MED provides network connectivity devices with the ability to:

- manage inventory
- manage Power over Ethernet (PoE)
- identify physical location
- identify network policy

LLDP-MED is designed for, but not limited to, VoIP endpoints.

TIA Organizationally Specific TLVs

The Dell Force10 system is an LLDP-MED Network Connectivity Device (Device Type 4). Network connectivity devices are responsible for:

- transmitting an LLDP-MED capabilities TLV to endpoint devices
- storing the information that endpoint devices advertise

Table describes the five types of TIA-1057 Organizationally Specific TLVs.

Table 26-3. TIA-1057 (LLDP-MED) Organizationally Specific TLVs

Type	Sub-type	TLV	Description
127	1	LLDP-MED Capabilities	Indicates: • whether the transmitting device supports LLDP-MED • what LLDP-MED TLVs it supports • LLDP device class
127	2	Network Policy	Indicates the application type, VLAN ID, Layer 2 Priority, and DSCP value
127	3	Location Identification	Indicates the physical location of the device expressed in one of three possible formats: • Coordinate Based LCI • Civic Address LCI • Emergency Call Services ELIN
127	4	Extended Power via MDI	Indicates power requirements, priority, and power status
Inventory Management TLVs			Implementation of this set of TLVs is optional in LLDP-MED devices. None or all TLVs must be supported. FTOS does not currently support these TLVs.
127	5	Inventory - Hardware Revision	Indicates the hardware revision of the LLDP-MED device
127	6	Inventory - Firmware Revision	Indicates the firmware revision of the LLDP-MED device
127	7	Inventory - Software Revision	Indicates the software revision of the LLDP-MED device
127	8	Inventory - Serial Number	Indicates the device serial number of the LLDP-MED device
127	9	Inventory - Manufacturer Name	Indicates the manufacturer of the LLDP-MED device
127	10	Inventory - Model Name	Indicates the model of the LLDP-MED device

Table 26-3. TIA-1057 (LLDP-MED) Organizationally Specific TLVs (continued)

Type	Sub-type	TLV	Description
127	11	Inventory - Asset ID	Indicates a user specified device number to manage inventory
127	12-255	Reserved	—

LLDP-MED Capabilities TLV

The LLDP-MED Capabilities TLV communicates the types of TLVs that the endpoint device and the network connectivity device support. LLDP-MED network connectivity devices must transmit the Network Policies TLV.

- The value of the LLDP-MED Capabilities field in the TLV is a 2 octet bitmap (Figure 26-4), each bit represents an LLDP-MED capability (Table 26-4).
- The possible values of the LLDP-MED Device Type is listed in Table 26-5. The Dell Force10 system is a Network Connectivity device, which is Type 4.

When you enable LLDP-MED in FTOS (using the command **advertise med**) the system begins transmitting this TLV.

Figure 26-4. LLDP-MED Capabilities TLV

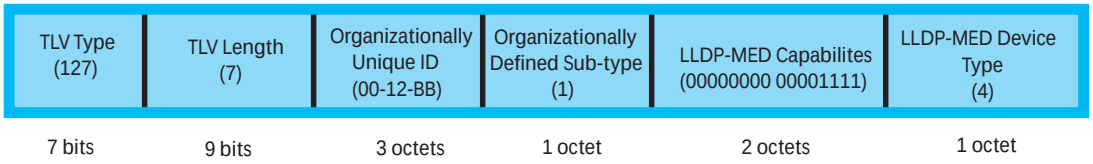


Table 26-4. FTOS LLDP-MED Capabilities

Bit Position	TLV	FTOS Support
0	LLDP-MED Capabilities	Yes
1	Network Policy	Yes
2	Location Identification	Yes
3	Extended Power via MDI-PSE	Yes
4	Extended Power via MDI-PD	No
5	Inventory	No
6-15	reserved	No

Table 26-5. LLDP-MED Device Types

Value	Device Type
0	Type Not Defined
1	Endpoint Class 1
2	Endpoint Class 2

Table 26-5. LLDP-MED Device Types

Value	Device Type
3	Endpoint Class 3
4	Network Connectivity
5-255	Reserved

LLDP-MED Network Policies TLV

A network policy in the context of LLDP-MED is a device's VLAN configuration and associated Layer 2 and Layer 3 configurations, specifically:

- VLAN ID
- VLAN tagged or untagged status
- Layer 2 priority
- DSCP value

The application type is a represented by an integer (the Type integer in Table 26-6), which indicates a device function for which a unique network policy is defined. An individual LLDP-MED Network Policy TLV is generated for each application type that you specify with the FTOS CLI (Advertising TLVs on page 592).

Note: With regard to Table 26-6, signaling is a series of control packets that are exchanged between an endpoint device and a network connectivity device to establish and maintain a connection. These signal packets might require a different network policy than the media packets for which a connection is made. In this case, configure the signaling application.

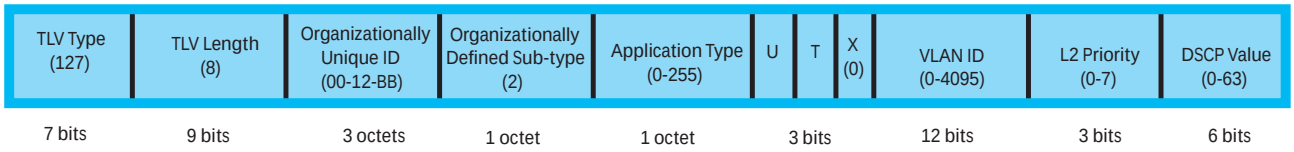
Table 26-6. Network Policy Applications

Type	Application	Description
0	Reserved	—
1	Voice	Specify this application type for dedicated IP telephony handsets and other appliances supporting interactive voice services.
2	Voice Signaling	Specify this application type only if voice control packets use a separate network policy than voice data.
3	Guest Voice	Specify this application type for a separate limited voice service for guest users with their own IP telephony handsets and other appliances supporting interactive voice services.
4	Guest Voice Signaling	Specify this application type only if guest voice control packets use a separate network policy than voice data.
5	Softphone Voice	Softphone is a computer program that enables IP telephony on a computer, rather than using a phone. Specify this application type for this type of endpoint device.
6	Video Conferencing	Specify this application type for dedicated video conferencing and other similar appliances supporting real-time interactive video.

Table 26-6. Network Policy Applications (continued)

Type	Application	Description
7	Streaming Video	Specify this application type for broadcast or multicast based video content distribution and other similar applications supporting streaming video services. This does not include video applications relying on TCP with buffering.
8	Video Signaling	Specify this application type only if video control packets use a separate network policy than video data.
9-255	Reserved	—

Figure 26-5. LLDP-MED Policies TLV

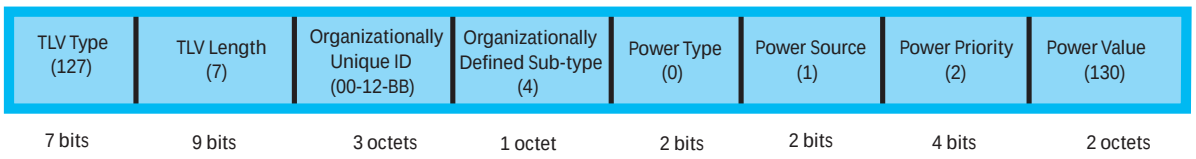


Extended Power via MDI TLV

The Extended Power via MDI TLV enables advanced PoE management between LLDP-MED endpoints and network connectivity devices. Advertise the Extended Power via MDI on all ports that are connected to an 802.3af powered, LLDP-MED endpoint device.

- **Power Type—**there are two possible power types: Power Sourcing Entity (PSE) or Power Device (PD). The Dell Force10 system is a PSE, which corresponds to a value of 0, based on the TIA-1057 specification.
- **Power Source**—there are two possible power sources: Primary and Backup. The Dell Force10 system is a Primary Power Source, which corresponds to a value of 1, based on the TIA-1057 specification.
- **Power Priority**—there are three possible priorities: Low, High, and Critical. On Dell Force10 systems, the default power priority is “High,” which corresponds to a value of 2 based on the TIA-1057 specification. You can configure a different power priority through the CLI, Dell Force10 also honors the power priority value sent by the powered device. However, the CLI configuration takes precedence.
- **Power Value**—Dell Force10 advertises the maximum amount of power that can be supplied on the port. By default it is 15.4W, which corresponds to a Power Value of 130, based on the TIA-1057 specification. You can advertise a different Power Value using the **max-milliwatts** option with the **power inline auto** | **static** command. Dell Force10 also honors the power value (power requirement) sent by the powered device when the port is configured for **power inline auto**.

Figure 26-6. Extended Power via MDI TLV



Configuring LLDP

Configuring LLDP is a two-step process:

1. Enable LLDP globally. See page 592.
2. Advertise TLVs out of an interface. See page 592.

Related Configuration Tasks

- Viewing the LLDP Configuration on page 594
- Viewing Information Advertised by Adjacent LLDP Agents on page 594
- Configuring LLDPDU Intervals on page 595
- Configuring Transmit and Receive Mode on page 596
- Configuring a Time to Live on page 597
- Debugging LLDP on page 598

Important Points to Remember

- LLDP is disabled by default.
- Dell Force10 systems support up to 8 neighbors per interface.
- Dell Force10 systems support a maximum of 8000 total neighbors per system. If the number of interfaces multiplied by 8 exceeds the maximum, the system will not configure more than 8000.
- INTERFACE level configurations override all CONFIGURATION level configurations.
- LLDP is not hitless.

LLDP Compatibility

- Spanning Tree and Force10 Ring Protocol “blocked” ports allow LLDPDUs.
- 802.1X controlled ports do not allow LLDPDUs until the connected device is authenticated.

CONFIGURATION versus INTERFACE Configurations

All LLDP configuration commands are available in PROTOCOL LLDP mode, which is a sub-mode of CONFIGURATION mode and INTERFACE mode.

- Configurations made at CONFIGURATION level are global, that is, they affect all interfaces on the system.
- Configurations made at INTERFACE level affect only the specific interface, and they override CONFIGURATION level configurations.

Figure 26-7. Configuration and Interface mode LLDP Commands

```
R1(conf)#protocol lldp
R1(conf-lldp)#?
advertise               Advertise TLVs
disable                 Disable LLDP protocol globally
end                     Exit from configuration mode
exit                    Exit from LLDP configuration mode
hello                   LLDP hello configuration
mode                    LLDP mode configuration (default = rx and tx)
multiplier              LLDP multiplier configuration
no                      Negate a command or set its defaults
show                    Show LLDP configuration
R1(conf-lldp)#exit
R1(conf)#interface gigabitethernet 1/31
R1(conf-if-gi-1/31)#protocol lldp
R1(conf-if-gi-1/31-lldp)#?
advertise               Advertise TLVs
disable                 Disable LLDP protocol on this interface
end                     Exit from configuration mode
exit                    Exit from LLDP configuration mode
hello                   LLDP hello configuration
mode                    LLDP mode configuration (default = rx and tx)
multiplier              LLDP multiplier configuration
no                      Negate a command or set its defaults
show                    Show LLDP configuration
R1(conf-if-gi-1/31-lldp)#
```

Enabling LLDP

LLDP is disabled by default. LLDP can be enabled and disabled globally or per interface. If LLDP is enabled globally, all up interfaces send periodic LLDPDUs. To enable LLDP:

Step	Task	Command	Command Mode
1	Enter Protocol LLDP mode.	**protocol lldp**	CONFIGURATION or INTERFACE
2	Enable LLDP.	**no disable**	PROTOCOL LLDP

Disabling and Undoing LLDP

- Disable LLDP globally or for an interface using the command **disable**.
- Undo an LLDP configuration by preceding the relevant command with the keyword **no**.

Advertising TLVs

You can configure the system to advertise TLVs out of all interfaces or out of specific interfaces.

- If you configure the system globally, all interfaces will send LLDPDUs with the specified TLVs.

- If you configure an interface, only the interface will send LLDPDUs with the specified TLVs.

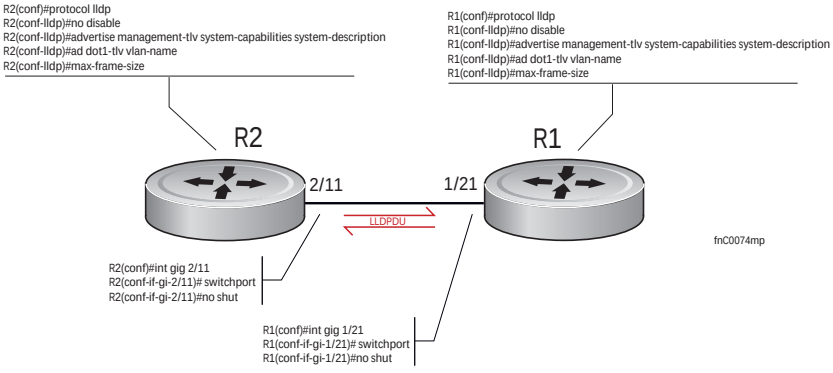
If LLDP is configured both globally and at interface level, the interface level configuration overrides the global configuration. To advertise TLVs:

Step	Task	Command	Command Mode
1	Enter LLDP mode.	**protocol lldp**	CONFIGURATION or INTERFACE
2	Advertise one or more TLVs. Include the keyword for each TLV you want to advertise. • For management TLVs: **system-capabilities**, **system-description** • For 802.1 TLVs: **port-protocol-vlan-id**, **port-vlan-id**, **vlan-name** • For 802.3 TLVs: **max-frame-size** • For TIA-1057 TLVs: **•guest-voice** **•guest-voice-signaling** **•location-identification** **•power-via-mdi** **•softphone-voice** **•streaming-video** **•video-conferencing** **•video-signaling** **•voice** **•voice-signaling**	**advertise {management-tlv | dot1-tlv | dot3-tlv | med}**	PROTOCOL LLDP

Note: **vlan-name** is supported on C-Series and S-Series only.

In Figure 26-8, LLDP is enabled globally. R1 and R2 are transmitting periodic LLDPDUs that contain management, 802.1, and 802.3 TLVs.

Figure 26-8. Configuring LLDP



Viewing the LLDP Configuration

Display the LLDP configuration using the command **show config** in either CONFIGURATION or INTERFACE mode, as shown in Figure 26-9 and Figure 26-10, respectively

Figure 26-9. Viewing LLDP Global Configurations

```
R1(conf)#protocol lldp
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 hello 10
 no disable
R1(conf-lldp)#
```

Figure 26-10. Viewing LLDP Interface Configurations

```
R1(conf-lldp)#exit
R1(conf)#interface gigabitethernet 1/31
R1(conf-if-gi-1/31)#show config
!
interface GigabitEthernet 1/31
 no ip address
 switchport
 no shutdown
R1(conf-if-gi-1/31)#protocol lldp
R1(conf-if-gi-1/31-lldp)#show config
!
 protocol lldp
R1(conf-if-gi-1/31-lldp)#
```

Viewing Information Advertised by Adjacent LLDP Agents

Display brief information about adjacent devices using the command **show lldp neighbors**, as shown in Figure 26-11. Display all of the information that neighbors are advertising using the command **show lldp neighbors detail**, as shown in Figure 26-12.

Figure 26-11. Viewing Brief Information Advertised by Adjacent LLDP Agents

```
R1(conf-if-gi-1/31-lldp)#end
R1(conf-if-gi-1/31)#do show lldp neighbors
 Loc PortID   Rem Host Name      Rem Port Id           Rem Chassis Id
 ---------------------------------------------------------------------------

 Gi 1/21      -                 GigabitEthernet 2/11  00:01:e8:06:95:3e
 Gi 1/31      -                 GigabitEthernet 3/11  00:01:e8:09:c2:4a
```

Figure 26-12. Viewing All Information Advertised by Adjacent LLDP Agent

```
R1#show lldp neighbors detail
============================================================================
 Local Interface Gi 1/21 has 1 neighbor
  Total Frames Out: 6547
  Total Frames In: 4136
  Total Neighbor information Age outs: 0
  Total Frames Discarded: 0
  Total In Error Frames: 0
  Total Unrecognized TLVs: 0
  Total TLVs Discarded: 0
  Next packet will be sent after 7 seconds
  The neighbors are given below:
  ---------------------------------------------------------------------------

    Remote Chassis ID Subtype: Mac address (4)
    Remote Chassis ID:  00:01:e8:06:95:3e
    Remote Port Subtype:  Interface name (5)
    Remote Port ID:  GigabitEthernet 2/11
    Local Port ID: GigabitEthernet 1/21
    Locally assigned remote Neighbor Index: 4
    Remote TTL:  120
    Information valid for next 120 seconds
    Time since last information change of this neighbor:  01:50:16
    Remote MTU:  1554
    Remote System Desc:  Force10 Networks Real Time Operating System Software
     . Force10 Operating System Version: 1.0. Force10 App
     lication Software Version: 7.5.1.0. Copyright (c) 19
     99-Build Time: Thu Aug 9 01:05:51 PDT 2007
    Existing System Capabilities:  Repeater Bridge Router
    Enabled System Capabilities:  Repeater Bridge Router
    Remote Port Vlan ID:  1
    Port and Protocol Vlan ID: 1, Capability: Supported, Status: Enabled
   ---------------------------------------------------------------------------

============================================================================
```

Configuring LLDPDU Intervals

LLDPDUs are transmitted periodically; the default interval is 30 seconds. You can configure a non-default transmit interval—at CONFIGURATION level or INTERFACE level—using the command **hello** (Figure 26-13).

Figure 26-13. Configuring LLDPDU Transmit and Receive Mode

```
R1(conf)#protocol lldp
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 no disable
R1(conf-lldp)#mode ?
rx                     Rx only
tx                     Tx only
R1(conf-lldp)#mode tx
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 mode tx
 no disable
R1(conf-lldp)#no mode
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 no disable
R1(conf-lldp)#
```

Configuring Transmit and Receive Mode

Once LLDP is enabled, Dell Force10 systems transmit *and* receive LLDPDUs by default. You can configure the system—at CONFIGURATION level or INTERFACE level—to transmit only by executing the command **mode tx**, or receive only by executing the command **mode rx**. Return to the default with the **no mode** command (Figure 26-14).

Figure 26-14. Configuring LLDPDU Transmit and Receive Mode

```
R1(conf)#protocol lldp
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 no disable
R1(conf-lldp)#mode ?
rx                      Rx only
tx                      Tx only
R1(conf-lldp)#mode tx
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 mode tx
 no disable
R1(conf-lldp)#no mode
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 no disable
R1(conf-lldp)#
```

Configuring a Time to Live

The information received from a neighbor expires after a specific amount of time (measured in seconds) called a Time to Live (TTL). The TTL is the product of the LLDPDU transmit interval (hello) and an integer called a *multiplier*. The default multiplier is 4, which results in a default TTL of 120 seconds. Adjust the TTL value—at CONFIGURATION level or INTERFACE level—using the command **multiplier**. Return to the default multiplier value using the command **no multiplier** (Figure 26-15).

Figure 26-15. Configuring LLDPDU Time to Live

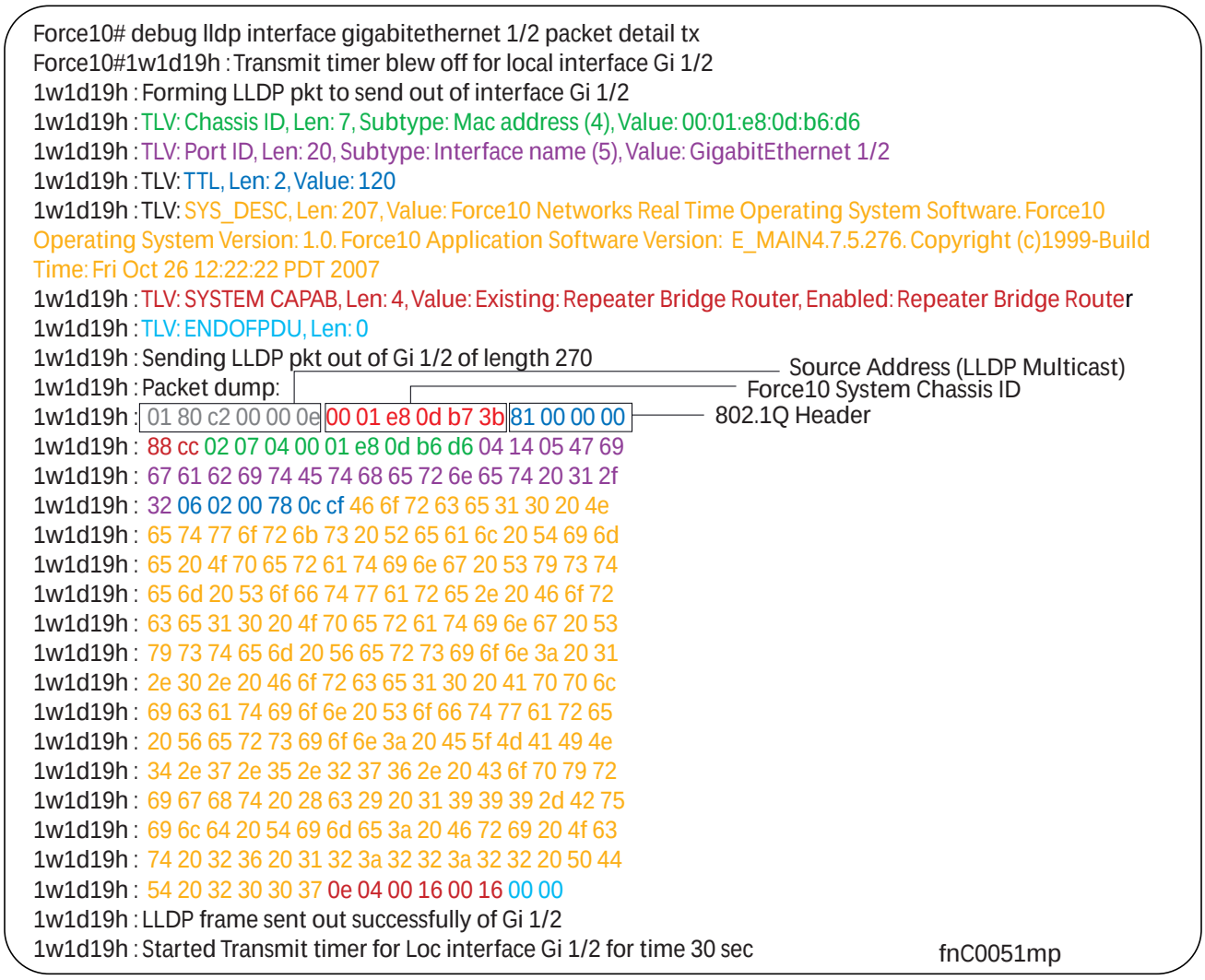
```
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 no disable
R1(conf-lldp)#multiplier ?
<2-10>                  Multiplier (default=4)
R1(conf-lldp)#multiplier 5
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 multiplier 5
 no disable
R1(conf-lldp)#no multiplier
R1(conf-lldp)#show config
!
protocol lldp
 advertise dot1-tlv port-protocol-vlan-id port-vlan-id
 advertise dot3-tlv max-frame-size
 advertise management-tlv system-capabilities system-description
 no disable
R1(conf-lldp)#
```

Debugging LLDP

The command **debug lldp** enables you to view the TLVs that your system is sending and receiving.

- Use the **debug lldp brief** command to view a readable version of the TLVs.
- Use the **debug lldp detail** command to view a readable version of the TLVs plus a hexadecimal version of the entire LLDPDU.

Figure 26-16. debug lldp detail—LLDPDU Packet Dissection



Relevant Management Objects

FTOS supports all IEEE 802.1AB MIB objects.

- Table lists the objects associated with received and transmitted TLVs.
- Table 26-8 lists the objects associated with the LLDP configuration on the local agent.
- Table 26-9 lists the objects associated with IEEE 802.1AB Organizationally Specific TLVs.
- Table 26-10 lists the objects associated with received and transmitted LLDP-MED TLVs.

Table 26-7. LLDP Configuration MIB Objects

MIB Object Category	LLDP Variable	LLDP MIB Object	Description
LLDP Configuration	adminStatus	lldpPortConfigAdminStatus	Whether the local LLDP agent is enabled for transmit, receive, or both
	msgTxHold	lldpMessageTxHoldMultiplier	Multiplier value
	msgTxInterval	lldpMessageTxInterval	Transmit Interval value
	rxInfoTTL	lldpRxInfoTTL	Time to Live for received TLVs
	txInfoTTL	lldpTxInfoTTL	Time to Live for transmitted TLVs
Basic TLV Selection	mibBasicTLVsTxEnable	lldpPortConfigTLVsTxEnable	Indicates which management TLVs are enabled for system ports
	mibMgmtAddrInstanceTxEnable	lldpManAddrPortsTxEnable	The management addresses defined for the system and and the ports through which they are enabled for transmission
LLDP Statistics	statsAgeoutsTotal	lldpStatsRxPortAgeoutsTotal	Total number of times that a neighbors information is deleted on the local system due to an rxInfoTTL timer expiration
	statsFramesDiscardedTotal	lldpStatsRxPortFramesDiscardedTotal	Total number of LLDP frames received then discarded
	statsFramesInErrorsTotal	lldpStatsRxPortFramesErrors	Total number of LLDP frames received on a port with errors
	statsFramesInTotal	lldpStatsRxPortFramesTotal	Total number of LLDP frames received through the port
	statsFramesOutTotal	lldpStatsTxPortFramesTotal	Total number of LLDP frames transmitted through the port
	statsTLVsDiscardedTotal	lldpStatsRxPortTLVsDiscardedTotal	Total number of TLVs received then discarded
	statsTLVsUnrecognizedTotal	lldpStatsRxPortTLVsUnrecognizedTotal	Total number of all TLVs the local agent does not recognize

Table 26-8. LLDP System MIB Objects

TLV Type	TLV Name	TLV Variable	System	LLDP MIB Object
1	Chassis ID	chassis ID subtype	Local	lldpLocChassisIdSubtype
			Remote	lldpRemChassisIdSubtype
		chassid ID	Local	lldpLocChassisId
			Remote	lldpRemChassisId
2	Port ID	port subtype	Local	lldpLocPortIdSubtype
			Remote	lldpRemPortIdSubtype
		port ID	Local	lldpLocPortId
			Remote	lldpRemPortId
4	Port Description	port description	Local	lldpLocPortDesc
			Remote	lldpRemPortDesc
5	System Name	system name	Local	lldpLocSysName
			Remote	lldpRemSysName
6	System Description	system description	Local	lldpLocSysDesc
			Remote	lldpRemSysDesc
7	System Capabilities	system capabilities	Local	lldpLocSysCapSupported
			Remote	lldpRemSysCapSupported
8	Management Address	enabled capabilities	Local	lldpLocSysCapEnabled
			Remote	lldpRemSysCapEnabled
		management address length	Local	lldpLocManAddrLen
			Remote	lldpRemManAddrLen
		management address subtype	Local	lldpLocManAddrSubtype
			Remote	lldpRemManAddrSubtype
		management address	Local	lldpLocManAddr
			Remote	lldpRemManAddr
		interface numbering subtype	Local	lldpLocManAddrIfSubtype
			Remote	lldpRemManAddrIfSubtype
		interface number	Local	lldpLocManAddrIfId
			Remote	lldpRemManAddrIfId
		OID	Local	lldpLocManAddrOID
			Remote	lldpRemManAddrOID

Table 26-9. LLDP 802.1 Organizationally Specific TLV MIB Objects

TLV Type	TLV Name	TLV Variable	System	LLDP MIB Object
127	Port-VLAN ID	PVID	Local	lldpXdot1LocPortVlanId
			Remote	lldpXdot1RemPortVlanId
127	Port and Protocol VLAN ID	port and protocol VLAN supported	Local	lldpXdot1LocProtoVlanSupported
			Remote	lldpXdot1RemProtoVlanSupported
		port and protocol VLAN enabled	Local	lldpXdot1LocProtoVlanEnabled
			Remote	lldpXdot1RemProtoVlanEnabled
		PPVID	Local	lldpXdot1LocProtoVlanId
			Remote	lldpXdot1RemProtoVlanId
127	VLAN Name	VID	Local	lldpXdot1LocVlanId
			Remote	lldpXdot1RemVlanId
		VLAN name length	Local	lldpXdot1LocVlanName
			Remote	lldpXdot1RemVlanName
		VLAN name	Local	lldpXdot1LocVlanName
			Remote	lldpXdot1RemVlanName

Table 26-10. LLDP-MED System MIB Objects

TLV Sub-Type	TLV Name	TLV Variable	System	LLDP-MED MIB Object
1	LLDP-MED Capabilities	LLDP-MED Capabilities	Local	lldpXMedPortCapSupported lldpXMedPortConfigTLVsTx Enable
			Remote	lldpXMedRemCapSupported, lldpXMedRemConfigTLVsTx Enable
		LLDP-MED Class Type	Local	lldpXMedLocDeviceClass
			Remote	lldpXMedRemDeviceClass
2	Network Policy	Application Type	Local	lldpXMedLocMediaPolicyApp Type
			Remote	lldpXMedRemMediaPolicyAp pType
		Unknown Policy Flag	Local	lldpXMedLocMediaPolicyUnk nown
			Remote	lldpXMedLocMediaPolicyUnk nown
		Tagged Flag	Local	lldpXMedLocMediaPolicyTag ged
			Remote	lldpXMedLocMediaPolicyTag ged
		VLAN ID	Local	lldpXMedLocMediaPolicyVla nID
			Remote	lldpXMedRemMediaPolicyVl anID
		L2 Priority	Local	lldpXMedLocMediaPolicyPrio rity
			Remote	lldpXMedRemMediaPolicyPri ority
		DSCP Value	Local	lldpXMedLocMediaPolicyDsc p
			Remote	lldpXMedRemMediaPolicyDs cp
3	Location Identifier	Location Data Format	Local	lldpXMedLocLocationSubtype
			Remote	lldpXMedRemLocationSubtyp e
		Location ID Data	Local	lldpXMedLocLocationInfo
			Remote	lldpXMedRemLocationInfo

Table 26-10. LLDP-MED System MIB Objects (continued)

TLV Sub-Type	TLV Name	TLV Variable	System	LLDP-MED MIB Object
4	Extended Power via MDI	Power Device Type	Local	lldpXMedLocXPoEDeviceType
			Remote	lldpXMedRemXPoEDeviceType
		Power Source	Local	lldpXMedLocXPoEPSEPowerSource, lldpXMedLocXPoEPDPowerSource
			Remote	lldpXMedRemXPoEPSEPowerSource, lldpXMedRemXPoEPDPowerSource
		Power Priority	Local	lldpXMedLocXPoEPDPowerPriority, lldpXMedLocXPoEPSEPortPDPriority
			Remote	lldpXMedRemXPoEPSEPowerPriority, lldpXMedRemXPoEPDPowerPriority
		Power Value	Local	lldpXMedLocXPoEPSEPortPowerAv, lldpXMedLocXPoEPDPowerReq
			Remote	lldpXMedRemXPoEPSEPowerAv, lldpXMedRemXPoEPDPowerReq

27

Multicast Listener Discovery

Multicast Listener Discovery is supported only on platform: E

MLD Snooping is supported only on platform: E

Multicast Listener Discovery (MLD) is a Layer 3 protocol that IPv6 routers use to learn of the multicast receivers that are directly connected to them and the groups in which the receivers are interested. Multicast routing protocols (like PIM) use the information learned from MLD to route multicast traffic to all interested receivers. MLD is analogous to IGMP, which tracks IPv4 multicast receivers.

Protocol Overview

MLD version 1 is analogous to IGMP version 2. MLD version 3 adds the ability to include and exclude sources and is analogous to IGMP version 3.

MLD Version 1

Routers use MLD to learn which multicast addresses have listeners on each of their attached links. For each link, the router keeps a list of which multicast addresses have listeners and a timer associated with each of those addresses.

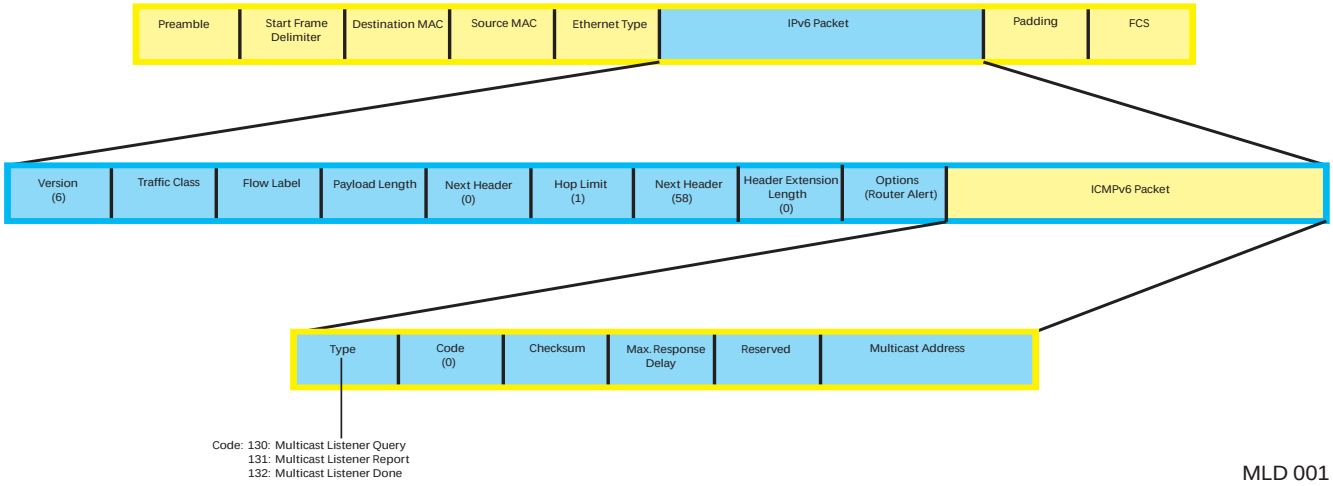
There are three types of MLD messages:

- **Multicast Listener Query** — a message sent by the Queerer to learn which multicast groups have listeners.
  - **General Query** — a query to which all listeners should respond.
  - **Multicast-Address-Specific Query** — a query to which only listeners for the specified group should respond.
- **Multicast Listener Report** — a message sent by listeners declaring their multicast group subscriptions.
- **Multicast Listener Done** — a message sent by a listener declaring that it is leaving a multicast group.

Figure 27-1 shows the packet structure of MLD version 1 packets.

- **Maximum Response Delay**—the maximum amount of time that the Querier waits to receive a response to a General or Multicast-Address-Specific Query. The value is zero in reports and Done messages.
- **Multicast Address** — set to zero in General Queries, and set to the relevant multicast address in multicast-address-specific queries and done messages.

Figure 27-1. MLD version 1 Packet Structure



MLD Querier Router

MLD routers periodically ask connected hosts in which, if any, multicasts groups they are interested. For any subnet, only on router solicit hosts for this information; this router is called the Querier, and all the other routers on the subnet are non-queriers. Initially, each router assumes that it is the Querier, and transmits queries. If a router receives a query with a source IP address lower than its own, it stops transmitting queries, and so the router with the lowest IP address is ultimately elected the Querier for the subnet.

Joining a Multicast Group

The Querier periodically sends a General Query to the all-nodes multicast address FF02::1. A host that wants to join a multicast group responds to the general query with a report that contains (in the MLD Multicast Address field, Figure 27-1) the group address; the report is also addressed to the group (in the IPv6 Destination Address field). To avoid duplicate reporting, any host that hears a report from another host for the same group in which it itself is interested cancels its report for that group.

A host does not have to wait for a General Query to join a group. If a host wants to become a member of a group for which the router is not currently forwarding traffic, it should send an unsolicited report.

When a router receives a report for a group, it either creates a new entry in the group membership table, or it updates an existing entry by adding the interface on which the report arrived to the outgoing interface list for the group.

Leaving a Multicast Group

A receiver that is no longer interested in traffic for a particular group should leave the group by sending a Done message to the link-scope all-routers multicast address, FF02::02.

When a Querier receives a Done message, it sends a Multicast-Address-Specific Query addressed to the relevant multicast group. Hosts still interested in receiving traffic for that group (according to the suppression mechanism) so that the group table entry is maintained. If no reports are received in response to the query, the group membership entry is cleared and the router stops forwarding traffic for that group.

MLD version 2

MLD version 2 (MLDv2) adds source-filtering capability. A node can report interest in multicast traffic only from specific source addresses or from all sources except for specific source addresses. MLDv2 is backwards compatible with MLD version 1.

There are two types of MLDv2 messages

- **Multicast Listener Query** — a message sent by the Querier to discover multicast listeners (Figure 27-2).
  - **General Query** — a query to which all listeners should respond.
  - **Multicast-Address-Specific Query** — a query to which listeners for the specified group should respond to affirm their membership.
  - **Multicast-Address-and-Source-Specific Query** — a query to determine if there are any listeners interested in a group and source pair.
- **Version 2 Multicast Listener Report** — a response to a query indicating listening state or state changes (Figure 27-3).

Figure 27-2. MLDv2 Multicast Listener Query

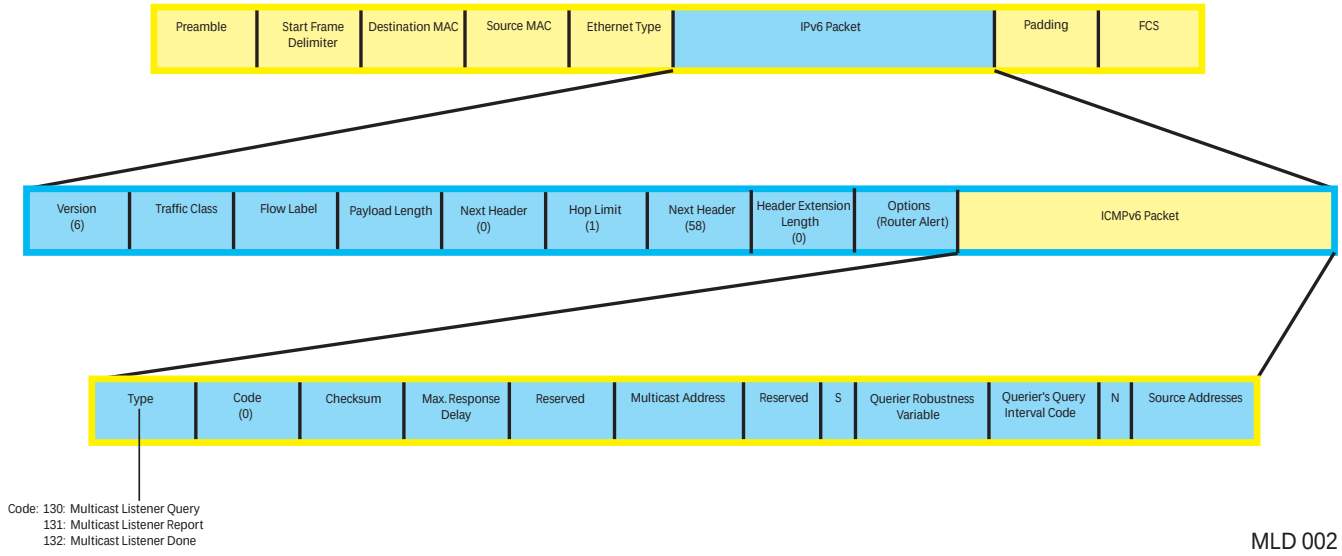
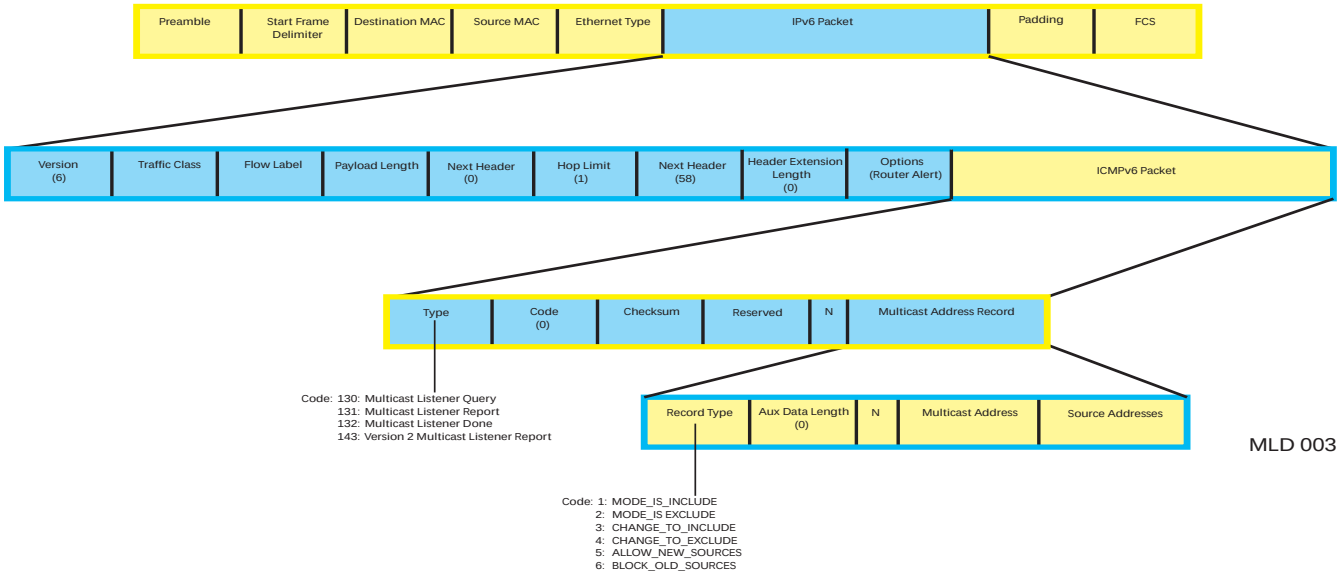


Figure 27-3. MLDv2 Multicast Listener Report



Implementation Information

- In FTOS versions prior to 8.3.1.0: when a switch on which MLD snooping is enabled acts as Querier, queries sent to a specific port (in the event of a port enable, MLD leave, or MLD join), were sent with an all-zero (::) IPv6 source address. This might unintentionally indicate to another MLD switch that it is elected the Querier. Beginning with FTOS version 8.3.1.0, all queries are flooded on the VLAN so that the IPv6 source address is correctly included in all queries.
- MLDv2 Snooping ignores sources specified in exclude reports; all exclude (S,G) reports are treated as exclude (*,G).
- The following querier commands are available for Layer 3 MLD and MLD Snooping: ipv6 mld **last-member-query-interval**, **ipv6 mld query-interval**, **ipv6 mld query-max-resp-time**.

Enabling MLD

MLD is enabled automatically when IPv6 PIM is enabled.

Related MLD Configuration Tasks

- Change MLD Timer Values on page 609
- Reduce Host Response Burstiness on page 609
- Reduce Leave Latency on page 609
- Configure a Static Group on page 610
- Clear MLD Groups on page 611
- Change the MLD Version on page 611

Change MLD Timer Values

All non-queriers have a timer that is refreshed when it hears a General Query. If the timer expires, then the router can assume that the Querier is not present, and so it assumes the role of Querier. The Other Querier Present Interval, or Querier Timeout Interval, is the amount of time that passes before a non-querier router assumes that there is no longer a Querier on the link.

Task	Command Syntax	Command Mode
Adjust the querier-timeout value.	**ipv6 mld querier-timeout** Default: 255 seconds	INTERFACE

The Query Interval is the amount of time between General Queries sent by the Querier.

Task	Command Syntax	Command Mode
Adjust the query interval.	**ipv6 mld query-interval** Default: 125 seconds	INTERFACE

Reduce Host Response Burstiness

General Queries contain a Query Response Interval value, which is the amount of time the host has to respond to a general query. Hosts set a timer to a random number less than the Query Response Interval upon receiving a general query, and send a report when the timer expires. Increasing this value spreads host responses over a greater period of time, and so reduces response burstiness.

Task	Command Syntax	Command Mode
Adjust the Query Response Interval.	**ipv6 mld query-max-resp-time** Default: 10 seconds	INTERFACE

Reduce Leave Latency

Leave Latency is the amount of time after the last host leaves the MLD group that the router stops forwarding traffic for that group. Latency is introduced because the router attempts several times to determine if there are any remaining members before stopping traffic for the group. There are two parameters you can configure to reduce leave latency.

Last Member Query Interval

The Querier sends a Multicast-Address-Specific Query upon receiving a Done message to ascertain whether there are any remain receivers for a group. The Last Listener Query Interval is the Maximum Response Delay for a Multicast-Address-Specific Query, and also the amount of time between Multicast-Address-Specific Query retransmissions. Lowering the Last Listener Query Interval reduces the time to detect that there are no remaining receivers for a group, and so can reduce the amount of unnecessarily forwarded traffic.

Task	Command Syntax	Command Mode
Adjust the last-member query interval.	**ipv6 mld last-member-query-interval** Default: 1000 milliseconds	INTERFACE

Explicit Tracking

If the Querier does not receive a response to a Multicast-Address-Specific Query, it sends another. Then, after no response, it removes the group entry from the group membership table. You can configure the system to remove specified groups immediately after receiving a Leave message to reduce leave latency.

Task	Command Syntax	Command Mode
Configure the system to remove a group after receiving immediately after receiving a Leave message. **Note:** If snooping is enabled on the VLAN, this command has no effect. In this case, enable **ipv6 mld snooping explicit tracking**.	**ipv6 mld explicit-tracking**	INTERFACE

Configure a Static Group

A group is entered into the group membership table if it has at least one member. Host memberships expire. When all memberships for a group expire, the group is removed from the group membership table. Hosts keep their memberships active by responding to queries. You can configure a group entry to never be removed regardless of membership by creating a static entry in the table.

Task	Command Syntax	Command Mode
Create a static entry in the group membership table.	**ipv6 mld static-group**	INTERFACE

Display the MLD Group Table

Task	Command Syntax	Command Mode
Display MLD groups. Group information can be filtered, see the *FTOS Command Line Reference* for the options available with this command.	**show ipv6 mld** {**groups** | **interface**}	EXEC Privilege

Clear MLD Groups

Clear a specific group or all groups on an interface from the multicast routing table using the command **clear ipv6 mld groups** from EXEC Privilege mode.

Change the MLD Version

Task	Command Syntax	Command Mode
Change the MLD version.	**ipv6 mld version 1** Default: MLD version 2	INTERFACE

Debug MLD

Task	Command Syntax	Command Mode
Display FTOS messages about the MLD process.	**debug ipv6 mld**	EXEC Privilege

MLD Snooping

Multicast packets are addressed with multicast MAC addresses, which represent a group of devices, rather than one unique device. Switches forward multicast frames out of all ports in a VLAN by default, even though there may be only some interested hosts, which is a waste of bandwidth. MLD Snooping enables switches to use information in MLD packets to generate a forwarding table that associates ports with multicast groups so that when they receive multicast frames, they can forward them only to interested receivers.

Enable MLD Snooping

MLD is automatically enabled when you enable IPv6 PIM, but MLD Snooping must be explicitly enabled.

Task	Command Syntax	Command Mode
Enable MLD Snooping	**ipv6 mld snooping enable**	CONFIGURATION

Disable MLD Snooping on a VLAN

When MLD is enabled globally, it is by default enabled on all VLANs. Disable snooping on a VLAN, using the command **no ipv6 mld snooping** from INTERFACE VLAN mode. Note that under the default configuration there is no need to configure **ipv6 mld snooping** for any VLAN.

Configure the Switch as a Querier

Hosts that do not support unsolicited reporting wait for a general query before sending a membership report. When the multicast source and receivers are in the same VLAN, multicast traffic is not routed, and so there is no querier. You must configure the switch to be the querier for a VLAN so that hosts send membership reports, and the switch can generate a forwarding table by snooping.

Configure the switch to be the querier for a Layer 2 VLAN using the command **ipv6 mld snooping querier** from INTERFACE VLAN mode. You must configure an IP address for the VLAN.

The source address of the queries is 0 to distinguish these queries from router queries. If the system receives a query with a non-zero address any VLAN interface, it stops sending queries. When a VLAN configured with snooping querier comes up, the VLAN interface waits for querier timeout to expire before becoming querier.

Disable Multicast Flooding

If the switch receives a multicast packet that has an IP address of a group it has not learned (unregistered frame), the switch floods that packet out of all ports on the VLAN.

You can configure the switch to only forward unregistered packets to ports on a VLAN that are connected to a multicast routers using the command **no ipv6 mld snooping flood** from CONFIGURATION mode. When flooding is disabled, if there are no such ports in the VLAN connected to a multicast router, the switch drops the packets.

Specify a Port as Connected to a Multicast Router

All MLD control packets and IP multicast data traffic originating from hosts are forwarded out all interfaces connected to multicast routers. These interfaces are called multicast router interfaces, or *mrouter* interfaces. You can statically specify a port in a VLAN as connected to a multicast router using the command **ipv6 mld snooping mrouter interface** from INTERFACE VLAN mode.

View the ports that are connected to multicast routers using the command **show ipv6 mld snooping mrouter** from EXEC Privilege mode.

Enable Snooping Explicit Tracking

The switch can be a querier, and therefore also has the option of updating the group table through explicit-tracking (see Explicit Tracking on page 610). Whether the switch is the Querier or not, if snooping is enabled, the switch tracks all MLD joins. It has separate explicit tracking table which contains group, source, interface, VLAN and reporter details.

Task	Command Syntax	Command Mode
Configure the system to remove a group immediately after receiving a Leave message.	**ipv6 mld snooping explicit-tracking**	VLAN INTERFACE
Display the MLD explicit-tracking table.	**show ipv6 mld snooping groups explicit**	EXEC Privilege

Display the MLD Snooping Table

Task	Command Syntax	Command Mode
Display the MLD Snooping table.	**show ipv6 mroute mld**	EXEC Privilege
Display group information in the table. Group information can be filtered, see the *FTOS Command Line Reference* for the options available with this command.	**show ipv6 mld snooping groups**	EXEC Privilege

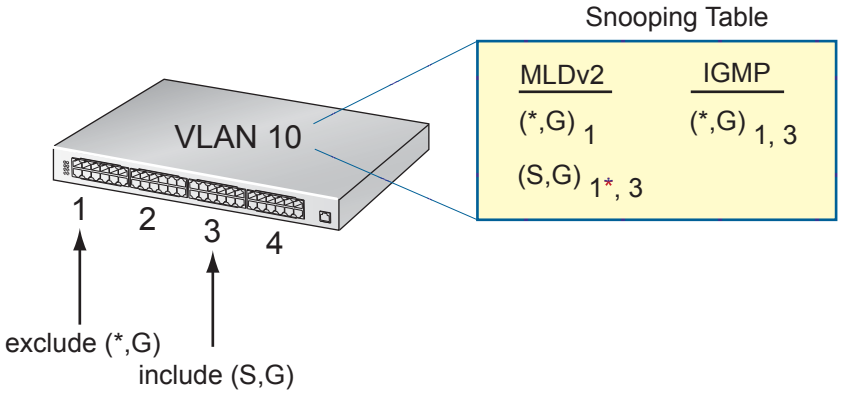
MLDv2 Snooping

With MLDv1 Snooping, multicast forwarding tables are formed for a group, G, based on received MLDv1 reports. With MLDv2 Snooping, multicast forwarding tables are formed for a source and group pair, (S,G), based on received MLDv2 include reports. MLDv2 Snooping is compatible with MLDv1 hosts and selects a port as dynamic mrouter port when it receives Membership Query on that port.

Port Inheritance on Mixed MLD Mode VLANs

A mixed MLD mode VLAN is one which has multiple hosts belonging to the same group, but some hosts exclude a source (S,G), and others include the same source (S,G).

Figure 27-4. Port Inheritance on Mixed-mode VLANs



In Figure 27-4, the host on Port 1 sends an exclude—that is, exclude nothing—report to join group G and receive traffic from all transmitting sources for the group. FTOS creates a (*,G) entry and lists Port 1 in the outgoing interface list. The host on Port 3 sends an include report to join the same group G, but receive traffic from only source S. FTOS creates a (S,G) entry and *could* list Port 3 as the outgoing interface. However, inbound traffic matches against the most specific entry, in this case, traffic from source S for group G matches the (S,G) entry. So, this traffic is forwarded out of only Port 3, which means that Port 1, which requested traffic from all sources, would be denied (S,G) traffic.

To reconcile this behavior, FTOS adds (*,G) ports to (S, G) entries. These inherited ports are marked with an asterisk to differentiate them from ports that have been snooped. In Figure 27-4, the (S,G) entry inherits Port 1 from the (*,G) entry. Now, (S,G) traffic is forwarded out Ports 1 and 3, so that Port 1 receives traffic from all sources, as requested.

Note: IGMPv3 does not inherit ports like MLDv2. Instead, when a VLAN has hosts that want to include and exclude the same source, S, the group defaults to exclude mode. That is, no (S,G) entry installed, and the excluding host receives all traffic. Notice, that in Figure 27-4, the MLD snooping table has an (S,G) entry, while the IGMP snooping does not.

28

Multicast Source Discovery Protocol

Multicast Source Discovery Protocol is supported only on platform E

MSDP addressing is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

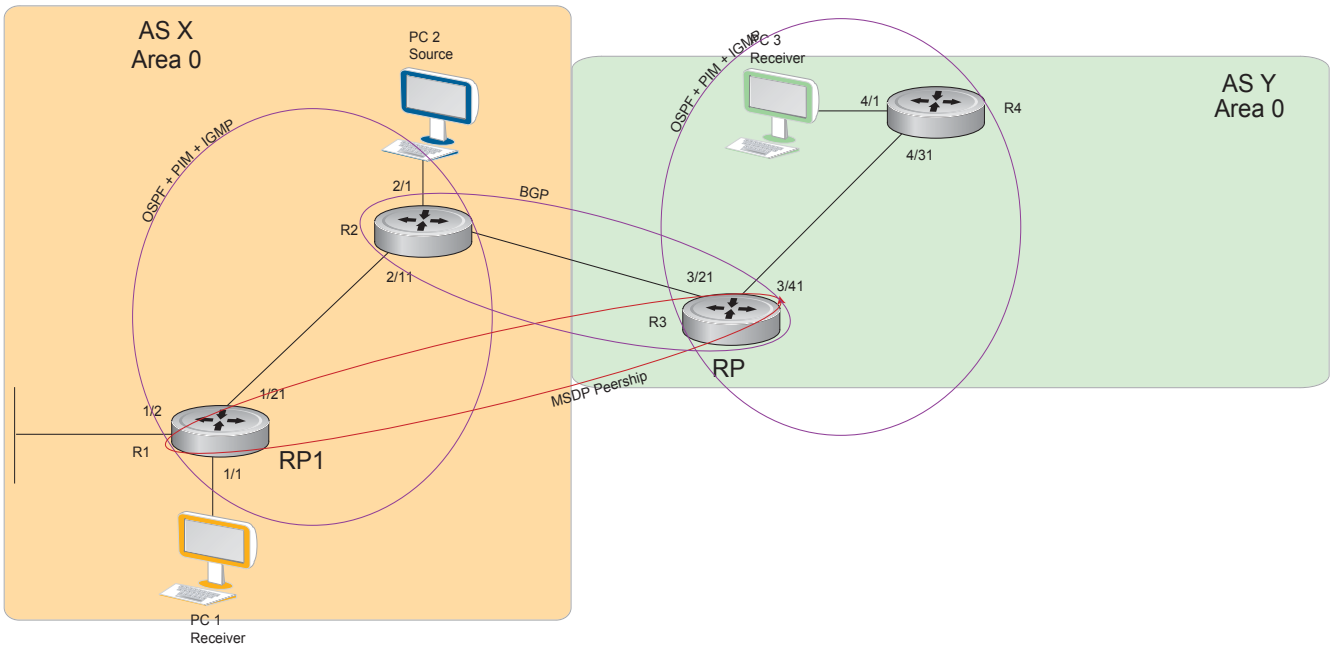
Protocol Overview

Multicast Source Discovery Protocol (MSDP) is a Layer 3 protocol that connects IPv4 PIM-SM domains. A domain in the context of MSDP is contiguous set of routers operating PIM within a common boundary defined by an exterior gateway protocol, such as BGP.

Each RP peers with every other RP via TCP. Through this connection, peers advertise the sources in their domain.

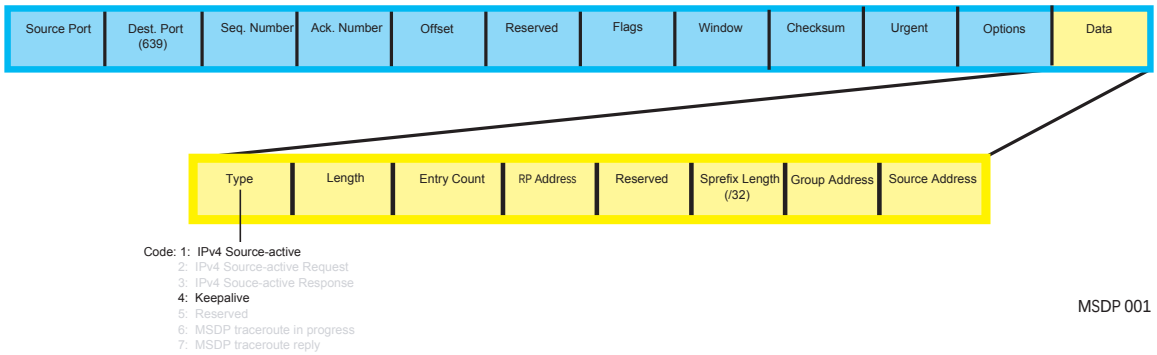
1. When an RP in a PIM-SM domain receives a PIM register message from a source it sends a Source-Active (SA) message (Figure 28-1) to MSDP peers.
2. Each MSDP peer receives and forwards the message to its peers away from the originating RP.
3. When an MSDP peer receives an SA message, it determines if there are any group members within the domain interested in any of the advertised sources. If there are, the receiving RP sends a join message to the originating RP, creating an SPT to the source.

Figure 28-1. Multicast Source Discovery Protocol



RPs advertise each (S,G) in its domain in Type, Length, Value (TLV) format. The total number of TLVs contained in the SA is indicated in the “Entry Count” field. SA messages are transmitted every 60 seconds, and immediately when a new source is detected.

Figure 28-2. MSDP SA Message Format



Implementation Information

- The FTOS implementation of MSDP is in accordance with RFC 3618 and Anycast RP is in accordance with RFC 3446.

Configuring Multicast Source Discovery Protocol

Configuring MSDP is a three-step process:

1. Enable an exterior gateway protocol (EGP) with at least two routing domains.

   Figure 28-5 and MSDP Sample Configurations on page 638 show the OSPF-BGP configuration used in this chapter for MSDP. Otherwise, see Chapter 32, Open Shortest Path First (OSPFv2 and OSPFv3), on page 691 and Chapter 10, Border Gateway Protocol IPv4 (BGPv4), on page 205.

2. Configure PIM-SM within each EGP routing domain.

   Figure 28-5 and MSDP Sample Configurations on page 638 show the PIM-SM configuration in this chapter for MSDP. Otherwise, see Chapter 34, PIM Sparse-Mode, on page 755.

3. Enable MSDP. See page 622.

4. Peer the RPs in each routing domain with each other. See page 622.

Related Configuration Tasks

- Enable MSDP on page 622
- Manage the Source-active Cache on page 622
- Accept Source-active Messages that fail the RFP Check on page 624
- Limit the Source-active Messages from a Peer on page 626
- Prevent MSDP from Caching a Local Source on page 627
- Prevent MSDP from Caching a Remote Source on page 628
- Prevent MSDP from Advertising a Local Source on page 629
- Terminate a Peership on page 630
- Clear Peer Statistics on page 631
- Clear Peer Statistics on page 631
- Debug MSDP on page 632
- MSDP with Anycast RP on page 632
- MSDP Sample Configurations on page 638

Figure 28-3. Configuring Interfaces for MSDP

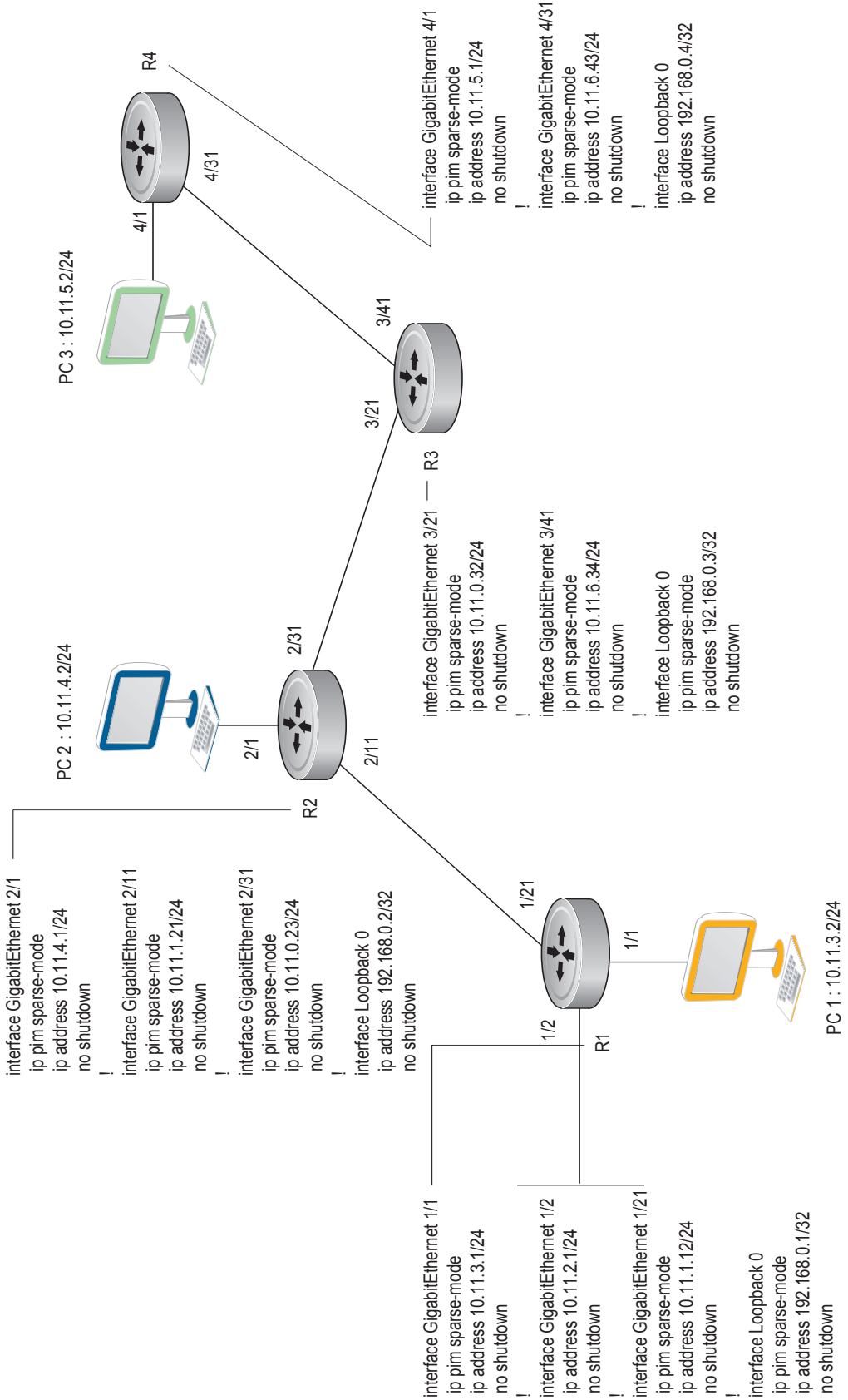


Figure 28-4. Configuring OSPF and BGP for MSDP

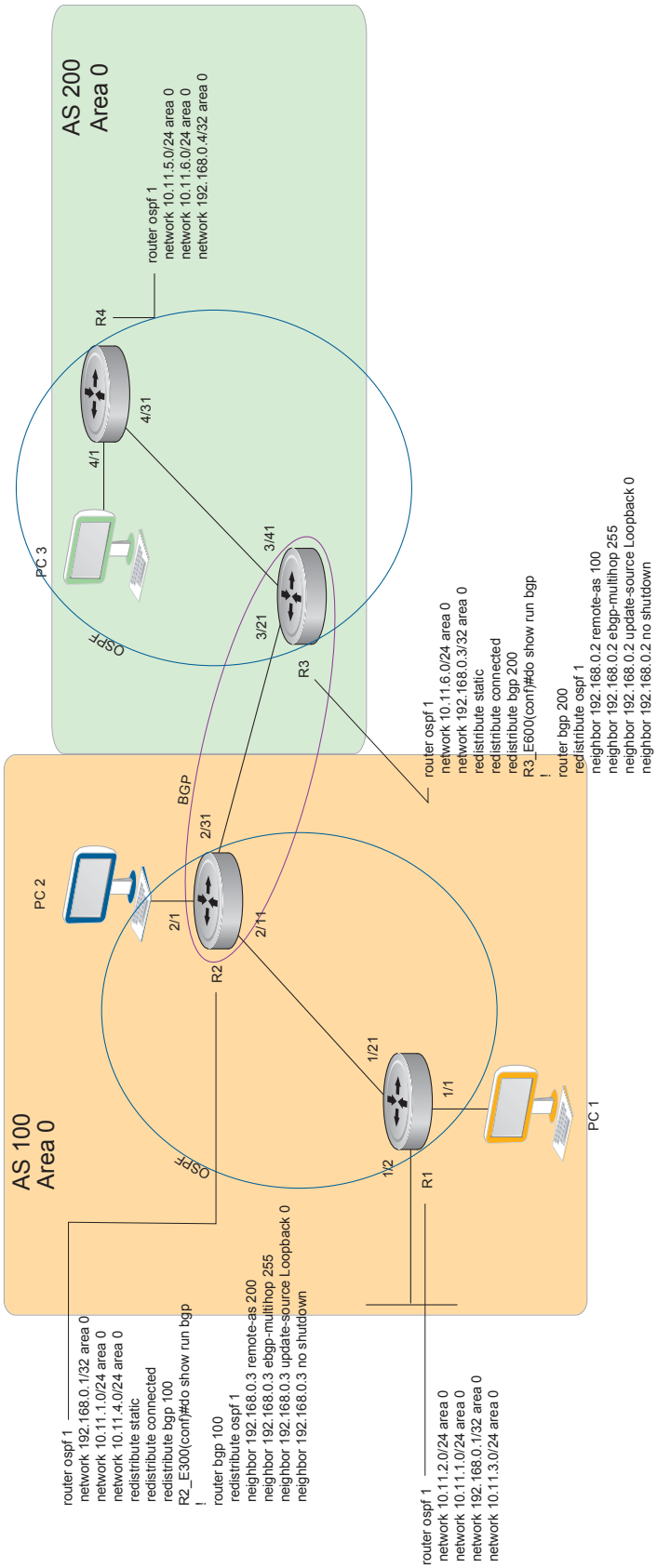


Figure 28-5. Configuring PIM in Multiple Routing Domains

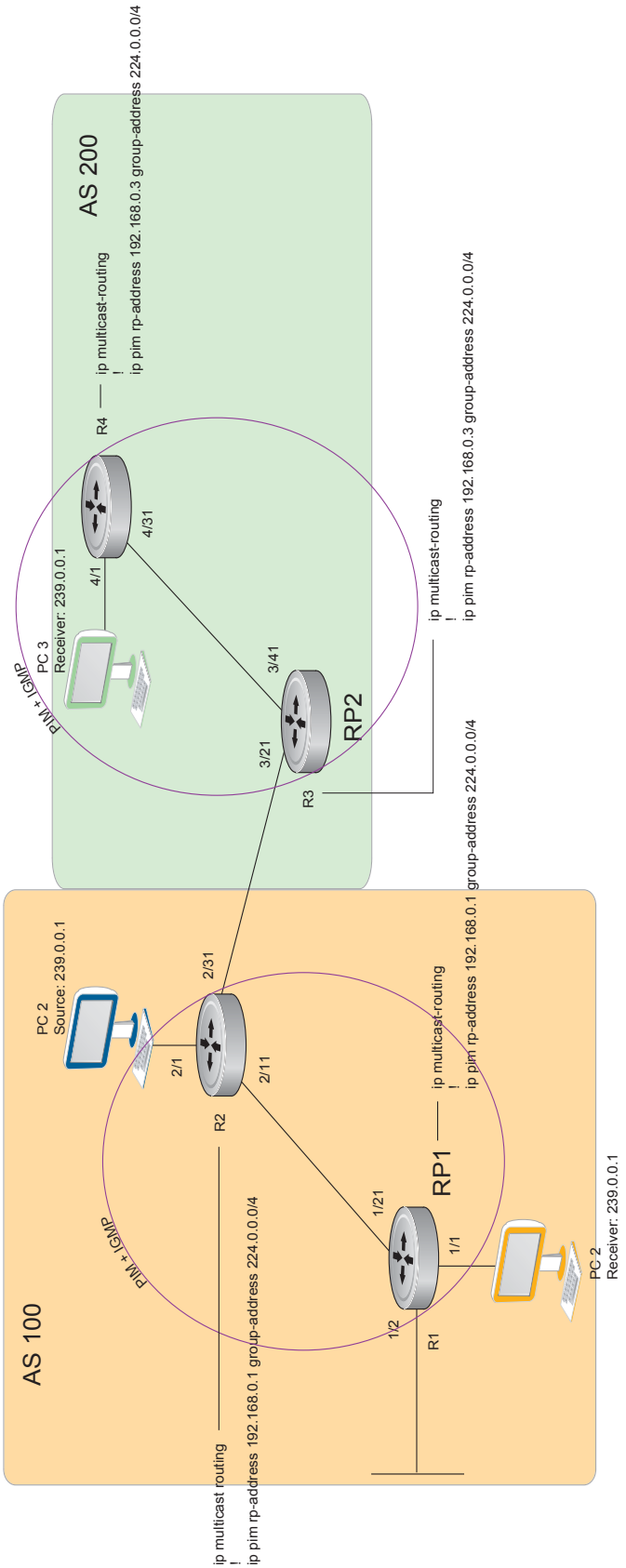
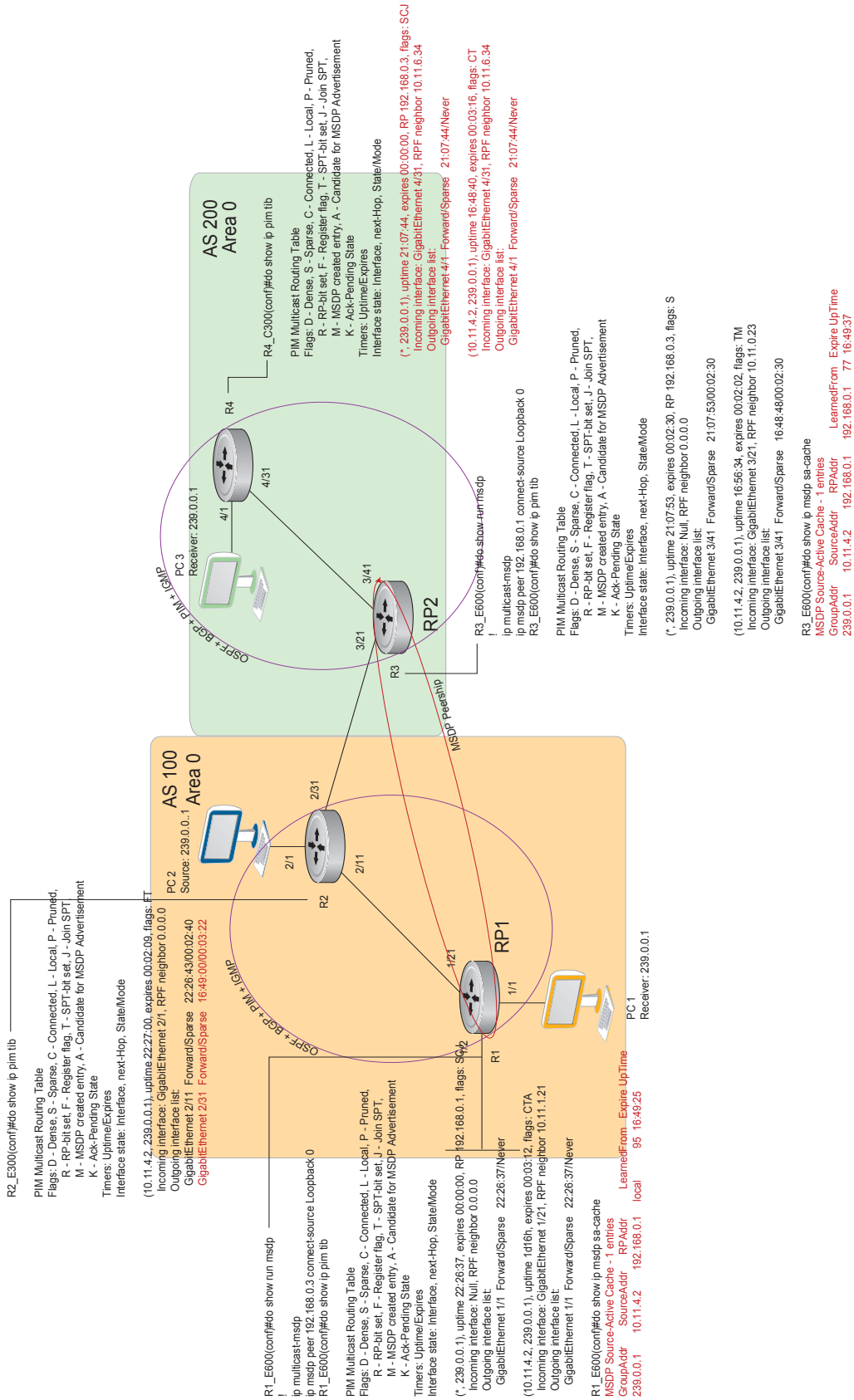


Figure 28-6. Configuring MSDP



Enable MSDP

Enable MSDP by peering RPs in different administrative domains.

Step	Task	Command Syntax	Command Mode
1	Enable MSDP.	**ip multicast-msdp**	CONFIGURATION
2	PeerPIM systems in different administrative domains.	**ip msdp peer connect-source**	CONFIGURATION

Figure 28-7. Configuring an MSDP Peer

```
R3_E600(conf)#ip multicast-msdp
R3_E600(conf)#ip msdp peer 192.168.0.1 connect-source Loopback 0
R3_E600(conf)#do show ip msdp summary

Peer Addr       Local Addr      State       Source    SA       Up/Down     Description
192.168.0.1     192.168.0.3     Established  Lo 0      1       00:05:29
```

Task	Command Syntax	Command Mode
View details about about a peer.	**show ip msdp peer**	EXEC Privilege

Figure 28-8. Displaying Details about a Peer

```
R3_E600#show ip msdp peer

Peer Addr: 192.168.0.1
    Local Addr: 192.168.0.3(639)  Connect Source: Lo 0
    State: Established  Up/Down Time: 00:15:20
    Timers: KeepAlive  30 sec, Hold time  75 sec
    SourceActive packet count (in/out): 8/0
    SAs learned from this peer: 1
    SA Filtering:
    Input (S,G) filter: none
    Output (S,G) filter: none
```

Multicast sources in remote domains are stored on the RP in the Source-active cache (SA cache). The system does not create entries in the multicast routing table until there is a local receiver for the corresponding multicast group.

Manage the Source-active Cache

Each SA-originating RP caches the sources inside its domain (domain-local), and the sources which it has learned from its peers (domain-remote). By caching sources:

- domain-local receivers experience a lower join latency,

- RPs can transmit SA messages periodically to prevent SA storms, and
- only sources that are in the cache are advertised in the SA to prevent transmitting multiple copies of the same source information.

View the Source-active Cache

Task	Command Syntax	Command Mode
View the SA cache.	**show ip msdp sa-cache**	EXEC Privilege

Figure 28-9. Displaying the MSDP Source-active Cache

```
R3_E600#show ip msdp sa-cache
MSDP Source-Active Cache - 1 entries
GroupAddr       SourceAddr      RPAddr          LearnedFrom   Expire UpTime
239.0.0.1       10.11.4.2       192.168.0.1     192.168.0.1     76  00:10:44
```

Limit the Source-active Cache

Set the upper limit of the number of active sources that FTOS caches. The default active source limit is 500K messages. When the total number of active sources reaches the specified limit, subsequent active sources are dropped even if they pass the RPF and policy check.

Task	Command Syntax	Command Mode
Limit the number of sources that can be stored in the SA cache.	**show ip msdp sa-limit**	EXEC Privilege

If the total number of active sources is already larger than the limit when limiting is applied, the sources that are already in FTOS are not discarded. To enforce the limit in such a situation, use the command **clear ip msdp sa-cache** to clear all existing entries.

Clear the Source-active Cache

Task	Command Syntax	Command Mode
Clear the SA cache of all, local, or rejected entries, or entries for a specific group.	**clear ip msdp sa-cache** [**group-address** | **local** | **rejected-sa**]	CONFIGURATION

Enable the Rejected Source-active Cache

Active sources can be rejected because

- the RPF check failed,
- the SA limit is reached,

- the peer RP is unreachable,
- or because of an SA message format error.

Task	Command Syntax	Command Mode
Cache rejected sources.	**ip msdp cache-rejected-sa**	CONFIGURATION

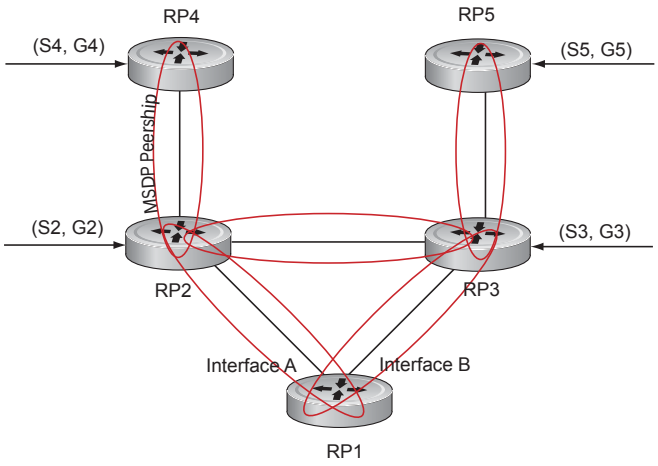
Accept Source-active Messages that fail the RFP Check

A default peer is a peer from which active sources are accepted even though they fail the RFP check.

- In Scenario 1 of Figure 28-10, all MSPD peers are up.
- In Scenario 2, the peership between RP1 and RP2 is down, but the link (and routing protocols) between them is still up. In this case, RP1 learns all active sources from RP3, the but the sources from RP2 and RP4 are rejected because the reverse path to these routers is through Interface A.
- In Scenario 3, RP3 is configured as a default MSDP peer for RP1 and so the RPF check is disregarded for RP3.
- In Scenario 4, RP1 has a default peer plus an access list. The list permits RP4 so the RPF check is disregarded for active sources from it, but RP5 (and all others because of the implicit deny all) are subject to the RPF check and fail, so those active sources are rejected.

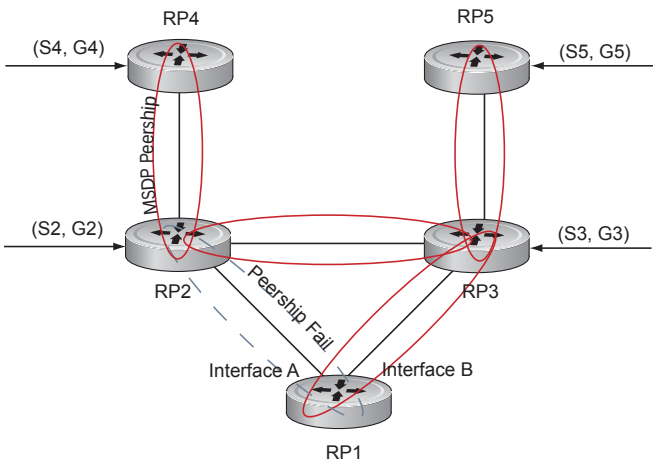
Figure 28-10. MSDP Default Peer

Scenario 1



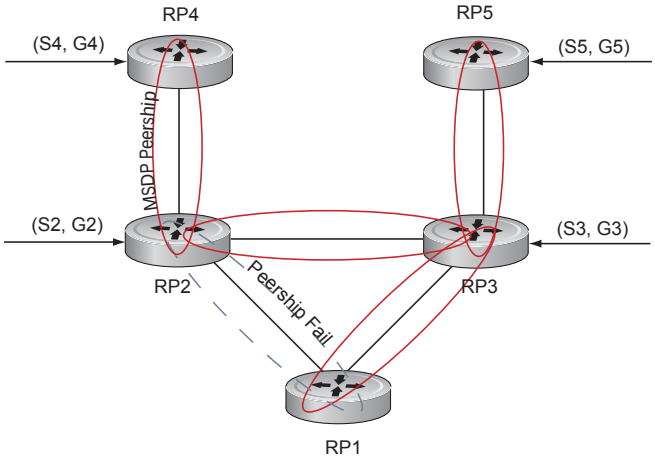
Group	Source	RP	Peer
G2	S2	RP2	R2
G3	S3	RP3	R3
G4	S4	RP4	R4
G5	S5	RP5	R5

Scenario 2



Group	Source	RP	Peer
G2	S2	RP2	R3 RPF-Fail
G3	S3	RP3	R3
G4	S4	RP4	R3 RPF-Fail
G5	S5	RP5	R3

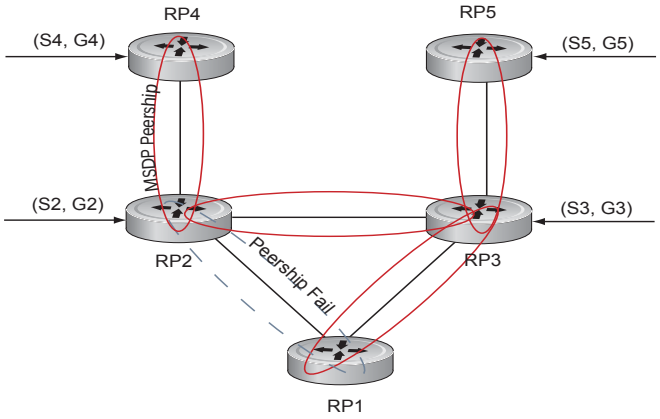
Scenario 3



ip msdp default-peer Router 3

Group	Source	RP	Peer
G2	S2	RP2	R3
G3	S3	RP3	R3
G4	S4	RP4	R3
G5	S5	RP5	R3

Scenario 4



ip msdp default-peer Router 3 access-list list123
ip access-list list123
permit RP4
Deny RP5

Group	Source	RP	Peer
G2	S2	RP2	R3 RPF-Fail
G3	S3	RP3	R3
G4	S4	RP4	R3
G5	S5	RP5	R3

Task	Command Syntax	Command Mode
Specify the forwarding-peer and originating-RP from which all active sources are accepted without regard for the the RPF check. If you do not specify an access list, the peer accepts all sources advertised by that peer. All sources from RPs denied by the ACL are subjected to the normal RPF check.	**ip msdp default-peer** *ip-address* **list**	CONFIGURATION

Figure 28-11. Accepting Source-active Messages with

```
FTOS(conf)#ip msdp peer 10.0.50.2 connect-source Vlan 50
FTOS(conf)#ip msdp default-peer 10.0.50.2 list fifty

FTOS(conf)#ip access-list standard fifty
FTOS(conf)#seq 5 permit host 200.0.0.50

FTOS#sh ip msdp sa-cache
MSDP Source-Active Cache - 3 entries
GroupAddr       SourceAddr      RPAddr          LearnedFrom   Expire UpTime
229.0.50.2      24.0.50.2       200.0.0.50      10.0.50.2       73  00:13:49
229.0.50.3      24.0.50.3       200.0.0.50      10.0.50.2       73  00:13:49
229.0.50.4      24.0.50.4       200.0.0.50      10.0.50.2       73  00:13:49

FTOS#sh ip msdp sa-cache rejected-sa
MSDP Rejected SA Cache
 3 rejected SAs received, cache-size 32766
UpTime      GroupAddr       SourceAddr      RPAddr          LearnedFrom     Reason
00:33:18    229.0.50.64     24.0.50.64      200.0.1.50      10.0.50.2       Rpf-Fail
00:33:18    229.0.50.65     24.0.50.65      200.0.1.50      10.0.50.2       Rpf-Fail
00:33:18    229.0.50.66     24.0.50.66      200.0.1.50      10.0.50.2       Rpf-Fail
```

Limit the Source-active Messages from a Peer

Task	Command Syntax	Command Mode
OPTIONAL: Store sources that are received after the limit is reached in the rejected SA cache.	**ip msdp cache-rejected-sa**	CONFIGURATION
Set the upper limit for the number of sources allowed from an MSDP peer. The default limit is 100K.	**ip msdp peer** *peer-address* **sa-limit**	CONFIGURATION

If the total number of sources received from the peer is already larger than the limit when this configuration is applied, those sources are not discarded. To enforce the limit in such a situation, first clear the SA cache.

Prevent MSDP from Caching a Local Source

You can prevent MSDP from caching an active source based on source and/or group. Since the source is not cached, it is not advertised to remote RPs.

Task	Command Syntax	Command Mode
OPTIONAL: Cache sources that are denied by the redistribute list in the rejected SA cache.	**ip msdp cache-rejected-sa**	CONFIGURATION
Prevent the system from caching local SA entries based on source and group using an extended ACL.	**ip msdp redistribute list**	CONFIGURATION

When you apply this filter, the SA cache is not affected immediately. When sources which are denied by the ACL time out, they are not refreshed. Until they time out, they continue to reside in the cache. To apply the redistribute filter to entries already present in the SA cache, first clear the SA cache. You may optionally store denied sources in the rejected SA cache.

Figure 28-12. Preventing MSDP from Caching a Local Source

```
R1_E600(conf)#do show run msdp
!
ip multicast-msdp
ip msdp peer 192.168.0.3 connect-source Loopback 0
ip msdp redistribute list mylocalfilter
ip msdp cache-rejected-sa 1000
R1_E600(conf)#do show run acl
!
ip access-list extended mylocalfilter
 seq 5 deny ip host 239.0.0.1 host 10.11.4.2
 seq 10 deny ip any any
R1_E600(conf)#do show ip msdp sa-cache
R1_E600(conf)#do show ip msdp sa-cache rejected-sa
MSDP Rejected SA Cache
 1 rejected SAs received, cache-size 1000
UpTime      GroupAddr       SourceAddr      RPAddr          LearnedFrom     Reason
00:02:20    239.0.0.1       10.11.4.2       192.168.0.1     local           Redistribute
```

Prevent MSDP from Caching a Remote Source

Task	Command Syntax	Command Mode
OPTIONAL: Cache sources that are denied by the SA filter in the rejected SA cache.	**ip msdp cache-rejected-sa**	CONFIGURATION
Prevent the system from caching remote sources learned from a specific peer based on source and group.	**ip msdp sa-filter list out** *peer* **list** *ext-acl*	CONFIGURATION

In Figure 28-14, R1 is advertising source 10.11.4.2. It is already in the SA cache of R3 when an ingress SA filter is applied to R3. The entry remains in the SA cache until it expires; it is not stored in the rejected SA cache.

Figure 28-13. Preventing MSDP from Advertising a Local Source

```
[Router 3]
R3_E600(conf)#do show run msdp
!
ip multicast-msdp
ip msdp peer 192.168.0.1 connect-source Loopback 0
ip msdp sa-filter in 192.168.0.1 list myremotefilter
R3_E600(conf)#do show run acl
!
ip access-list extended myremotefilter
 seq 5 deny ip host 239.0.0.1 host 10.11.4.2
R3_E600(conf)#do show ip msdp sa-cache
MSDP Source-Active Cache - 1 entries
GroupAddr       SourceAddr      RPAddr          LearnedFrom   Expire UpTime
239.0.0.1       10.11.4.2       192.168.0.1     192.168.0.1     1   00:03:59
R3_E600(conf)#do show ip msdp sa-cache
R3_E600(conf)#
R3_E600(conf)#do show ip msdp peer

Peer Addr: 192.168.0.1
    Local Addr: 0.0.0.0(639)  Connect Source: Lo 0
    State: Listening    Up/Down Time: 00:01:19
    Timers: KeepAlive  30 sec, Hold time  75 sec
    SourceActive packet count (in/out): 0/0
    SAs learned from this peer: 0
    SA Filtering:
    Input (S,G) filter: myremotefilter
    Output (S,G) filter: none
```

Prevent MSDP from Advertising a Local Source

Task	Command Syntax	Command Mode
Prevent an RP from advertising a source in the SA cache.	**ip msdp sa-filter list in** *peer* **list** *ext-acl*	CONFIGURATION

In Figure 28-14, R1 stops advertising source 10.11.4.2. Since it is already in the SA cache of R3, the entry remains there until it expires.

Figure 28-14. Preventing MSDP from Advertising a Local Source

```
[Router 1]
R1_E600(conf)#do show run msdp
!
ip multicast-msdp
ip msdp peer 192.168.0.3 connect-source Loopback 0
ip msdp sa-filter out 192.168.0.3 list mylocalfilter
R1_E600(conf)#do show run acl
!
ip access-list extended mylocalfilter
 seq 5 deny ip host 239.0.0.1 host 10.11.4.2
 seq 10 deny ip any any
R1_E600(conf)#do show ip msdp sa-cache
MSDP Source-Active Cache - 1 entries
GroupAddr       SourceAddr      RPAddr          LearnedFrom   Expire UpTime
239.0.0.1       10.11.4.2       192.168.0.1     local           70  00:27:20
R3_E600(conf)#do show ip msdp sa-cache
MSDP Source-Active Cache - 1 entries
GroupAddr       SourceAddr      RPAddr          LearnedFrom   Expire UpTime
239.0.0.1       10.11.4.2       192.168.0.1     192.168.0.1     1   00:10:29

[Router 3]
R3_E600(conf)#do show ip msdp sa-cache
R3_E600(conf)#
```

Display the configured SA filters for a peer using the command **show ip msdp peer** from EXEC Privilege mode (see Figure 28-14).

Log Changes in Peership States

Task	Command Syntax	Command Mode
Log peership state changes.	**ip msdp log-adjacency-changes**	CONFIGURATION

Terminate a Peership

MSDP uses TCP as its transport protocol. In a peering relationship, the peer with the lower IP address initiates the TCP session, while the peer with the higher IP address listens on port 639.

Task	Command Syntax	Command Mode
Terminate the TCP connection with a peer.	**ip msdp shutdown**	CONFIGURATION

Once the relationship is terminated, the peering state of the terminator is SHUTDOWN, while the peering state of the peer is INACTIVE.

Figure 28-15. Terminating a Peership

```
[Router 3]
R3_E600(conf)#ip msdp shutdown 192.168.0.1
R3_E600(conf)#do show ip msdp peer

Peer Addr: 192.168.0.1
    Local Addr: 0.0.0.0(0)  Connect Source: Lo 0
    State: Shutdown     Up/Down Time: 00:00:18
    Timers: KeepAlive  30 sec, Hold time  75 sec
    SourceActive packet count (in/out): 0/0
    SAs learned from this peer: 0
    SA Filtering:
    Input (S,G) filter: myremotefilter
    Output (S,G) filter: none
[Router 1]
R1_E600(conf)#do show ip msdp peer

Peer Addr: 192.168.0.3
    Local Addr: 0.0.0.0(0)  Connect Source: Lo 0
    State: Inactive     Up/Down Time: 00:00:03
    Timers: KeepAlive  30 sec, Hold time  75 sec
    SourceActive packet count (in/out): 0/0
    SAs learned from this peer: 0
    SA Filtering:
```

Clear Peer Statistics

Task	Command Syntax	Command Mode
Reset the TCP connection to the peer and clear all peer statistics.	**clear ip msdp peer peer-address**	CONFIGURATION

Figure 28-16. Clearing Peer Statistics

```
R3_E600(conf)#do show ip msdp peer

Peer Addr: 192.168.0.1
    Local Addr: 192.168.0.3(639)  Connect Source: Lo 0
    State: Established  Up/Down Time: 00:04:26
    Timers: KeepAlive  30 sec, Hold time  75 sec
    SourceActive packet count (in/out): 5/0
    SAs learned from this peer: 0
    SA Filtering:
    Input (S,G) filter: myremotefilter
    Output (S,G) filter: none
R3_E600(conf)#do clear ip msdp peer 192.168.0.1
R3_E600(conf)#do show ip msdp peer

Peer Addr: 192.168.0.1
    Local Addr: 0.0.0.0(0)  Connect Source: Lo 0
    State: Inactive     Up/Down Time: 00:00:04
    Timers: KeepAlive  30 sec, Hold time  75 sec
    SourceActive packet count (in/out): 0/0
    SAs learned from this peer: 0
    SA Filtering:
    Input (S,G) filter: myremotefilter
    Output (S,G) filter: none
```

Debug MSDP

Task	Command Syntax	Command Mode
Display the information exchanged between peers.	**debug ip msdp**	CONFIGURATION

Figure 28-17. Debugging MSDP

```
R1_E600(conf)#do debug ip msdp
All MSDP debugging has been turned on
R1_E600(conf)#03:16:08 : MSDP-0: Peer 192.168.0.3,  sent Keepalive msg
03:16:09 : MSDP-0: Peer 192.168.0.3,  rcvd Keepalive msg
03:16:27 : MSDP-0: Peer 192.168.0.3,  sent Source Active msg
03:16:38 : MSDP-0: Peer 192.168.0.3,  sent Keepalive msg
03:16:39 : MSDP-0: Peer 192.168.0.3,  rcvd Keepalive msg
03:17:09 : MSDP-0: Peer 192.168.0.3,  sent Keepalive msg
03:17:10 : MSDP-0: Peer 192.168.0.3,  rcvd Keepalive msg
03:17:27 : MSDP-0: Peer 192.168.0.3,  sent Source Active msg
Input (S,G) filter: none
    Output (S,G) filter: none
```

MSDP with Anycast RP

Anycast RP use MSDP with PIM-SM to allow more than one active group to RP mapping.

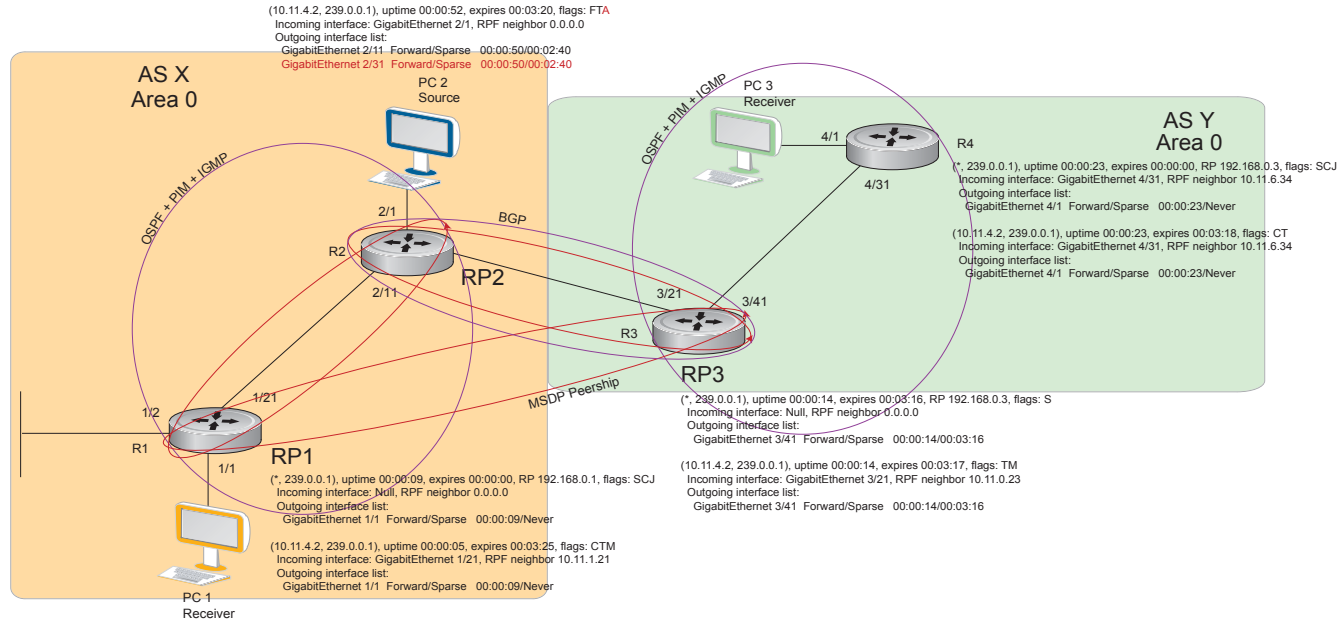
PIM-SM allows only active group to RP mapping, which has several implications:

- **traffic concentration**: PIM-SM allows only one active group to RP mapping which means that all traffic for the group must, at least initially, travel over the same part of the network. You can load balance source registration between multiple RPs by strategically mapping groups to RPs, but this technique is less effective as traffic increases because preemptive load balancing requires prior knowledge of traffic distributions.
- **lack of scalable register decasulation**: With only a single RP per group, all joins are sent to that RP regardless of the topological distance between the RP, sources, and receivers, and data is transmitted to the RP until the SPT switch threshold is reached.
- **slow convergence when an active RP fails**: When multiple RPs are configured, there can be considerable convergence delay involved in switching to the backup RP.

Anycast RP relieves these limitations by allowing multiple RPs per group, which can be distributed in a topologically significant manner according to the locations of the sources and receivers.

1. All the RPs serving a given group are configured with an identical anycast address.
2. Sources then register with the topologically closest RP.
3. RPs use MSDP to peer with each other using a unique address.

Figure 28-18. MSDP with Anycast RP



To configure Anycast RP:

Step	Task	Command Syntax	Command Mode
1	In each routing domain that will have multiple RPs serving a group, create a loopback interface on each RP serving the group with the same IP address.	**interface loopback**	CONFIGURATION
2	Make this address the RP for the group.	**ip pim rp-address**	CONFIGURATION
3	In each routing domain that will have multiple RPs serving a group, create another loopback interface on each RP serving the group with a unique IP address.	**interface loopback**	CONFIGURATION
4	Peer each RP with every other RP using MSDP, specifying the unique loopback address as the connect-source.	**ip msdp peer**	CONFIGURATION
5	Advertise the network of each of the unique loopback addresses throughout the network.	**network**	ROUTER OSPF

Reducing Source-active Message Flooding

RPs flood source-active messages to all of their peers away from the RP. When multiple RPs exist within a domain, the RPs forward received active source information back to the originating RP, which violates the RFP rule. You can prevent this unnecessary flooding by creating a mesh-group. A mesh in this context is a topology in which each RP in a set of RPs has a peership with all other RPs in the set. When an RP is a member of the mesh group, it forwards active source information only to its peers outside of the group.

Task	Command Syntax	Command Mode
Create a mesh group.	**ip msdp mesh-group**	CONFIGURATION

Specify the RP Address Used in SA Messages

The default originator-id is the address of the RP that created the message. In the case of Anycast RP, there are multiple RPs all with the same address. You can use the (unique) address of another interface as the originator-id.

Task	Command Syntax	Command Mode
Use the address of another interface as the originator-id instead of the RP address.	**ip msdp originator-id**	CONFIGURATION

Figure 28-19. R1 Configuration for MSDP with Anycast RP

```
ip multicast-routing
!
interface GigabitEthernet 1/1
 ip pim sparse-mode
 ip address 10.11.3.1/24
 no shutdown
!
interface GigabitEthernet 1/2
 ip address 10.11.2.1/24
 no shutdown
!
interface GigabitEthernet 1/21
 ip pim sparse-mode
 ip address 10.11.1.12/24
 no shutdown
!
interface Loopback 0
 ip pim sparse-mode
 ip address 192.168.0.1/32
 no shutdown
!
interface Loopback 1
 ip address 192.168.0.11/32
 no shutdown
!
router ospf 1
 network 10.11.2.0/24 area 0
 network 10.11.1.0/24 area 0
 network 10.11.3.0/24 area 0
 network 192.168.0.11/32 area 0
!
ip multicast-msdp
ip msdp peer 192.168.0.3 connect-source Loopback 1
ip msdp peer 192.168.0.22 connect-source Loopback 1
ip msdp mesh-group AS100 192.168.0.22
ip msdp originator-id Loopback 1
!
ip pim rp-address 192.168.0.1 group-address 224.0.0.0/4
```

Figure 28-20. R2 Configuration for MSDP with Anycast RP

```
ip multicast-routing
!
interface GigabitEthernet 2/1
 ip pim sparse-mode
 ip address 10.11.4.1/24
 no shutdown
!
interface GigabitEthernet 2/11
 ip pim sparse-mode
 ip address 10.11.1.21/24
 no shutdown
!
interface GigabitEthernet 2/31
 ip pim sparse-mode
 ip address 10.11.0.23/24
 no shutdown
!
interface Loopback 0
 ip pim sparse-mode
 ip address 192.168.0.1/32
 no shutdown
!
interface Loopback 1
 ip address 192.168.0.22/32
 no shutdown
!
router ospf 1
 network 10.11.1.0/24 area 0
 network 10.11.4.0/24 area 0
 network 192.168.0.22/32 area 0
 redistribute static
 redistribute connected
 redistribute bgp 100
!
router bgp 100
 redistribute ospf 1
 neighbor 192.168.0.3 remote-as 200
 neighbor 192.168.0.3 ebgp-multihop 255
 neighbor 192.168.0.3 no shutdown
!
ip multicast-msdp
ip msdp peer 192.168.0.3 connect-source Loopback 1
ip msdp peer 192.168.0.11 connect-source Loopback 1
ip msdp mesh-group AS100 192.168.0.11
ip msdp originator-id Loopback 1
!
ip route 192.168.0.3/32 10.11.0.32
!
ip pim rp-address 192.168.0.1 group-address 224.0.0.0/4
```

Figure 28-21. R3 Configuration for MSDP with Anycast RP

```
ip multicast-routing
!
interface GigabitEthernet 3/21
 ip pim sparse-mode
 ip address 10.11.0.32/24
 no shutdown

interface GigabitEthernet 3/41
 ip pim sparse-mode
 ip address 10.11.6.34/24
 no shutdown
!
interface Loopback 0
 ip pim sparse-mode
 ip address 192.168.0.3/32
 no shutdown
!
router ospf 1
 network 10.11.6.0/24 area 0
 network 192.168.0.3/32 area 0
 redistribute static
 redistribute connected
 redistribute bgp 200
!
router bgp 200
 redistribute ospf 1
 neighbor 192.168.0.22 remote-as 100
 neighbor 192.168.0.22 ebgp-multihop 255
 neighbor 192.168.0.22 update-source Loopback 0
 neighbor 192.168.0.22 no shutdown
!
ip multicast-msdp
ip msdp peer 192.168.0.11 connect-source Loopback 0
ip msdp peer 192.168.0.22 connect-source Loopback 0
ip msdp sa-filter out 192.168.0.22
!
ip route 192.168.0.1/32 10.11.0.23
ip route 192.168.0.22/32 10.11.0.23
!
ip pim rp-address 192.168.0.3 group-address 224.0.0.0/4
```

MSDP Sample Configurations

The following figures show the running-configurations for the routers shown in figures Figure 28-5, Figure 28-4, Figure 28-5, Figure 28-6.

Figure 28-22. MSDP Sample Configuration: R1 Running-config

```
ip multicast-routing
!
interface GigabitEthernet 1/1
 ip pim sparse-mode
 ip address 10.11.3.1/24
 no shutdown
!
interface GigabitEthernet 1/2
 ip address 10.11.2.1/24
 no shutdown
!
interface GigabitEthernet 1/21
 ip pim sparse-mode
 ip address 10.11.1.12/24
 no shutdown
!
interface Loopback 0
 ip pim sparse-mode
 ip address 192.168.0.1/32
 no shutdown
!
router ospf 1
 network 10.11.2.0/24 area 0
 network 10.11.1.0/24 area 0
 network 192.168.0.1/32 area 0
 network 10.11.3.0/24 area 0
!
ip multicast-msdp
ip msdp peer 192.168.0.3 connect-source Loopback 0
!
ip pim rp-address 192.168.0.1 group-address 224.0.0.0/4
```

Figure 28-23. MSDP Sample Configuration: R2 Running-config

```
ip multicast-routing
!
interface GigabitEthernet 2/1
 ip pim sparse-mode
 ip address 10.11.4.1/24
 no shutdown
!
interface GigabitEthernet 2/11
 ip pim sparse-mode
 ip address 10.11.1.21/24
 no shutdown
!
interface GigabitEthernet 2/31
 ip pim sparse-mode
 ip address 10.11.0.23/24
 no shutdown
!
interface Loopback 0
 ip address 192.168.0.2/32
 no shutdown
!
router ospf 1
 network 10.11.1.0/24 area 0
 network 10.11.4.0/24 area 0
 network 192.168.0.2/32 area 0
 redistribute static
 redistribute connected
 redistribute bgp 100
!
router bgp 100
 redistribute ospf 1
 neighbor 192.168.0.3 remote-as 200
 neighbor 192.168.0.3 ebgp-multihop 255
 neighbor 192.168.0.3 update-source Loopback 0
 neighbor 192.168.0.3 no shutdown
!
ip route 192.168.0.3/32 10.11.0.32
!
ip pim rp-address 192.168.0.1 group-address 224.0.0.0/4
```

Figure 28-24. MSDP Sample Configuration: R3 Running-config

```
ip multicast-routing
!
interface GigabitEthernet 3/21
 ip pim sparse-mode
 ip address 10.11.0.32/24
 no shutdown
!
interface GigabitEthernet 3/41
 ip pim sparse-mode
 ip address 10.11.6.34/24
 no shutdown
!
interface ManagementEthernet 0/0
 ip address 10.11.80.3/24
 no shutdown
!
interface Loopback 0
 ip pim sparse-mode
 ip address 192.168.0.3/32
 no shutdown
!
router ospf 1
 network 10.11.6.0/24 area 0
 network 192.168.0.3/32 area 0
 redistribute static
 redistribute connected
 redistribute bgp 200
!
router bgp 200
 redistribute ospf 1
 neighbor 192.168.0.2 remote-as 100
 neighbor 192.168.0.2 ebgp-multihop 255
 neighbor 192.168.0.2 update-source Loopback 0
 neighbor 192.168.0.2 no shutdown
!
ip multicast-msdp
ip msdp peer 192.168.0.1 connect-source Loopback 0
!
ip route 192.168.0.2/32 10.11.0.23
```

Figure 28-25. MSDP Sample Configuration: R4 Running-config

```
ip multicast-routing
!
interface GigabitEthernet 4/1
 ip pim sparse-mode
 ip address 10.11.5.1/24
 no shutdown
!
interface GigabitEthernet 4/22
 ip address 10.10.42.1/24
 no shutdown
!
interface GigabitEthernet 4/31
 ip pim sparse-mode
 ip address 10.11.6.43/24
 no shutdown
!
interface Loopback 0
 ip address 192.168.0.4/32
 no shutdown
!
router ospf 1
 network 10.11.5.0/24 area 0
 network 10.11.6.0/24 area 0
 network 192.168.0.4/32 area 0
!
ip pim rp-address 192.168.0.3 group-address 224.0.0.0/4
```

29

Multiple Spanning Tree Protocol

Multiple Spanning Tree Protocol is supported on platforms: C E S

MSTP addressing is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

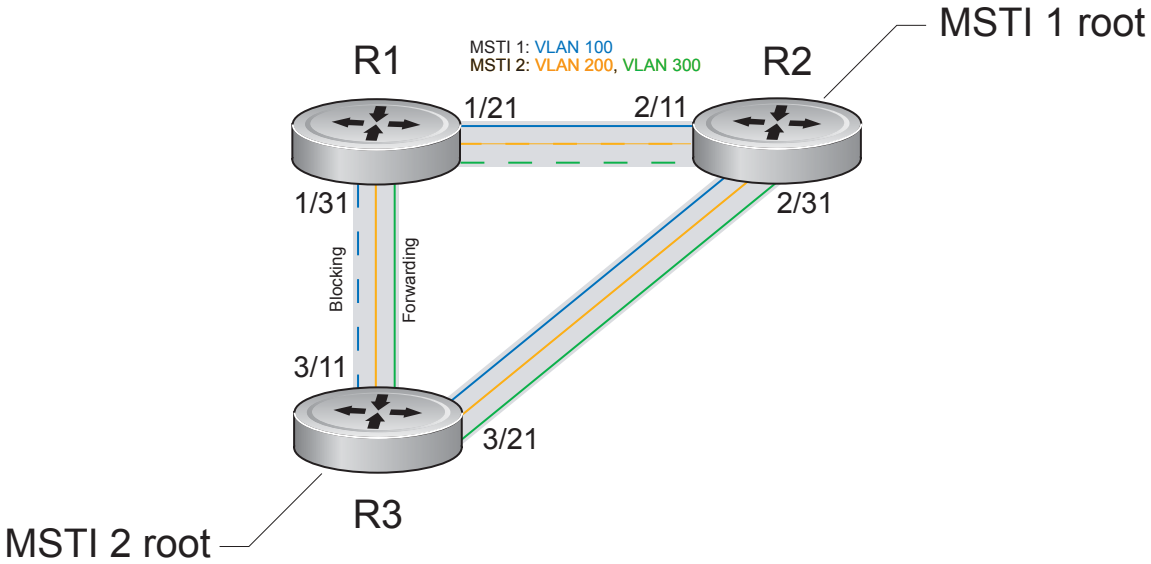
Protocol Overview

Multiple Spanning Tree Protocol (MSTP)—specified in IEEE 802.1Q-2003—is an RSTP-based spanning tree variation that improves on PVST+. MSTP allows multiple spanning tree instances and allows you to map many VLANs to one spanning tree instance to reduce the total number of required instances.

In contrast, PVST+ allows a spanning tree instance for each VLAN. This 1:1 approach is not suitable if you have many VLANs, because each spanning tree instance costs bandwidth and processing resources.

In Figure 29-1, three VLANs are mapped to two Multiple Spanning Tree instances (MSTI). VLAN 100 traffic takes a different path than VLAN 200 and 300 traffic. The behavior in Figure 29-1 demonstrates how you can use MSTP to achieve load balancing.

Figure 29-1. MSTP with Three VLANs Mapped to Two Spanning Tree Instances



FTOS supports three other variations of Spanning Tree, as shown in Table 44.

Table 29-1. FTOS Supported Spanning Tree Protocols

Dell Force10 Term	IEEE Specification
Spanning Tree Protocol	802.1d
Rapid Spanning Tree Protocol	802.1w
Multiple Spanning Tree Protocol	802.1s
Per-VLAN Spanning Tree Plus	Third Party

Implementation Information

- The FTOS MSTP implementation is based on IEEE 802.1Q-2003, and interoperates only with bridges that also use this standard implementation.
- MSTP is compatible with STP and RSTP.
- FTOS supports only one MSTP region.
- When you enable MSTP, all ports in Layer 2 mode participate in MSTP.
- On the C-Series and S-Series, you can configure 64 MSTIs including the default instance 0 (CIST).

Configure Multiple Spanning Tree Protocol

Configuring Multiple Spanning Tree is a four-step process:

1. Configure interfaces for Layer 2. See page 1051.
2. Place the interfaces in VLANs.
3. Enable Multiple Spanning Tree Protocol. See page 645.
4. Create Multiple Spanning Tree Instances, and map VLANs to them. See page 645.

Related Configuration Tasks

- Create Multiple Spanning Tree Instances on page 645
- Add and Remove Interfaces on page 645
- Influence MSTP Root Selection on page 647
- Interoperate with Non-FTOS Bridges on page 647
- Modify Global Parameters on page 648
- Modify Interface Parameters on page 650
- Configure an EdgePort on page 651
- Flush MAC Addresses after a Topology Change on page 654
- Debugging and Verifying MSTP Configuration on page 660
- Preventing Network Disruptions with BPDU Guard on page 1057
- SNMP Traps for Root Elections and Topology Changes on page 908
- Configuring Spanning Trees as Hitless on page 1064

Enable Multiple Spanning Tree Globally

MSTP is not enabled by default. To enable MSTP:

Step	Task	Command Syntax	Command Mode
1	Enter PROTOCOL MSTP mode.	**protocol spanning-tree mstp**	CONFIGURATION
2	Enable MSTP.	**no disable**	PROTOCOL MSTP

Verify that MSTP is enabled using the **show config** command from PROTOCOL MSTP mode, as shown in Figure 29-2.

Figure 29-2. Verifying MSTP is Enabled

```
FTOS(conf)#protocol spanning-tree mstp
FTOS(config-mstp)#show config
!
protocol spanning-tree mstp
 no disable
FTOS#
```

When you enable MSTP, all physical, VLAN, and port-channel interfaces that are enabled and in Layer 2 mode are automatically part of the MSTI 0.

- Within an MSTI, only one path from any bridge to any other bridge is enabled.
- Bridges block a redundant path by disabling one of the link ports.

Add and Remove Interfaces

- To add an interface to the MSTP topology, configure it for Layer 2 and add it to a VLAN. If you previously disabled MSTP on the interface using the command **no spanning-tree 0**, re-enable it using the command **spanning-tree 0**.
- Remove an interface from the MSTP topology using the command **no spanning-tree 0** command. See also Removing an Interface from the Spanning Tree Group on page 1054 for BPDU Filtering behavior.

Create Multiple Spanning Tree Instances

A single MSTI provides no more benefit than RSTP. To take full advantage of MSTP you must create multiple MSTIs and map VLANs to them.

Create an MSTI using the command **msti** from PROTOCOL MSTP mode. Specify the keyword **vlan** followed by the VLANs that you want to participate in the MSTI, as shown in Figure 29-3.

Figure 29-3. Mapping VLANs to MSTI Instances

```
FTOS(conf)#protocol spanning-tree mstp
FTOS(conf-mstp)#msti 1 vlan 100
FTOS(conf-mstp)#msti 2 vlan 200-300
FTOS(conf-mstp)#show config
!
protocol spanning-tree mstp
 no disable
 MSTI 1 VLAN 100
 MSTI 2 VLAN 200-300
```

All bridges in the MSTP region must have the same VLAN-to-instance mapping. View to which instance a VLAN is mapped using the command **show spanning-tree mst vlan** from EXEC Privilege mode, as shown in Figure 29-6.

View the forwarding/discarding state of the ports participating in an MSTI using the command **show spanning-tree msti** from EXEC Privilege mode, as shown in Figure 29-4.

Figure 29-4. Viewing MSTP Port States

```
FTOS#show spanning-tree msti 1
MSTI 1 VLANs mapped  100

Root Identifier has priority 32768, Address 0001.e806.953e
Root Bridge hello time 2, max age 20, forward delay 15, max hops 19
Bridge Identifier has priority 32768, Address 0001.e80d.b6d6
Configured hello time 2, max age 20, forward delay 15, max hops 20
Current root has priority 32768, Address 0001.e806.953e
Number of topology changes 2, last change occured 1d2h ago on Gi 1/21

Port 374 (GigabitEthernet 1/21) is root Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.374
Designated root has priority 32768, address 0001.e806.953e
Designated bridge has priority 32768, address 0001.e806.953e
Designated port id is 128.374, designated path cost 20000
Number of transitions to forwarding state 1
BPDU (MRecords): sent 93671, received 46843
The port is not in the Edge port mode

Port 384 (GigabitEthernet 1/31) is alternate Discarding
Port path cost 20000, Port priority 128, Port Identifier 128.384
Designated root has priority 32768, address 0001.e806.953e
Designated bridge has priority 32768, address 0001.e809.c24a
Designated port id is 128.384, designated path cost 20000
Number of transitions to forwarding state 1
BPDU (MRecords): sent 39291, received 7547
The port is not in the Edge port mode
```

Influence MSTP Root Selection

MSTP determines the root bridge, but you can assign one bridge a lower priority to increase the probability that it will become the root bridge.

To change the bridge priority:

Task	Command Syntax	Command Mode
Assign a number as the bridge priority. A lower number increases the probability that the bridge becomes the root bridge. **Range**: 0-61440, in increments of 4096 **Default**: 32768	**msti** *instance* **bridge-priority** *priority*	PROTOCOL MSTP

The simple configuration Figure 29-1 by default yields the same forwarding path for both MSTIs. Figure 29-5, shows how R3 is assigned bridge priority 0 for MSTI 2, which elects a different root bridge than MSTI 2. View the bridge priority using the command **show config** from PROTOCOL MSTP mode, also shown in Figure 29-5.

Figure 29-5. Changing the Bridge Priority

```
R3(conf-mstp)#msti 2 bridge-priority 0
1d2h51m: %RPM0-P:RP2 %SPANMGR-5-STP_ROOT_CHANGE: MSTP root changed for instance 2. My
Bridge ID: 0:0001.e809.c24a Old Root: 32768:0001.e806.953e New Root: 0:0001.e809.c24a

R3(conf-mstp)#show config
!
protocol spanning-tree mstp
 no disable
 MSTI 1 VLAN 100
 MSTI 2 VLAN 200,300
 MSTI 2 bridge-priority 0
```

Interoperate with Non-FTOS Bridges

FTOS supports only one MSTP region. A region is a combination of three unique qualities:

- **Name** is a mnemonic string you assign to the region. The default region name on FTOS is null.
- **Revision** is a two-byte number. The default revision number on FTOS is 0.
- VLAN-to-instance mapping is the placement of a VLAN in an MSTI.

For a bridge to be in the same MSTP region as another, all three of these qualities must match exactly. The default values for name and revision will match on all Dell Force10 FTOS equipment. If you have non-FTOS equipment that will participate in MSTP, ensure these values to match on all the equipment.

Note: Some non-FTOS equipment may implement a non-null default region name. SFTOS, for example, uses the Bridge ID, while others may use a MAC address.

To change the region name or revision:

Task	Command Syntax	Command Mode
Change the region name.	**name** *name*	PROTOCOL MSTP
Change the region revision number. • Range: 0 to 65535 • Default: 0	**revision** *number*	PROTOCOL MSTP

View the current region name and revision using the command **show spanning-tree mst configuration** from EXEC Privilege mode, as shown in Figure 29-6.

Figure 29-6. Viewing the MSTP Region Name and Revision

```
FTOS(conf-mstp)#name my-mstp-region
FTOS(conf-mstp)#exit
FTOS(conf)#do show spanning-tree mst config
MST region name: my-mstp-region
Revision: 0
MSTI    VID
   1    100
   2    200-300
```

Modify Global Parameters

The root bridge sets the values for forward-delay, hello-time, max-age, and max-hops and overwrites the values set on other MSTP bridges.

- **Forward-delay** is the amount of time an interface waits in the Listening State and the Learning State before it transitions to the Forwarding State.
- **Hello-time** is the time interval in which the bridge sends MSTP Bridge Protocol Data Units (BPDUs).
- **Max-age** is the length of time the bridge maintains configuration information before it refreshes that information by recomputing the MST topology.
- **Max-hops** is the maximum number of hops a BPDU can travel before a receiving switch discards it.

Note: Dell Force10 recommends that only experienced network administrators change MSTP parameters. Poorly planned modification of MSTP parameters can negatively impact network performance.

To change MSTP parameters, use the following commands on the root bridge:

Task	Command Syntax	Command Mode
Change the forward-delay parameter. • Range: 4 to 30 • Default: 15 seconds	**forward-delay** *seconds*	PROTOCOL MSTP

Task	Command Syntax	Command Mode
Change the hello-time parameter. **Note:** With large configurations (especially those with more ports) Dell Force10 recommends that you increase the hello-time. Range: 1 to 10 Default: 2 seconds	**hello-time** *seconds*	PROTOCOL MSTP
Change the max-age parameter. Range: 6 to 40 Default: 20 seconds	**max-age** *seconds*	PROTOCOL MSTP
Change the max-hops parameter. Range: 1 to 40 Default: 20	**max-hops** *number*	PROTOCOL MSTP

View the current values for MSTP parameters using the **show running-config spanning-tree mstp** command from EXEC privilege mode.

Figure 29-7. Viewing the Current Values for MSTP Parameters

```
FTOS(conf-mstp)#forward-delay 16
FTOS(conf-mstp)#exit
FTOS(conf)#do show running-config spanning-tree mstp
!
protocol spanning-tree mstp
 no disable
 name my-mstp-region
 MSTI 1 VLAN 100
 MSTI 2 VLAN 200-300
 forward-delay 16
 MSTI 2 bridge-priority 4096
FTOS(conf)#
```

Modify Interface Parameters

You can adjust two interface parameters to increase or decrease the probability that a port becomes a forwarding port:

- **Port cost** is a value that is based on the interface type. The greater the port cost, the less likely the port will be selected to be a forwarding port.
- **Port priority** influences the likelihood that a port will be selected to be a forwarding port in case that several ports have the same port cost.

Table 29-2 lists the default values for port cost by interface.

Table 29-2. MSTP Default Port Cost Values

Port Cost	Default Value
100-Mb/s Ethernet interfaces	200000
1-Gigabit Ethernet interfaces	20000
10-Gigabit Ethernet interfaces	2000
Port Channel with 100 Mb/s Ethernet interfaces	180000
Port Channel with 1-Gigabit Ethernet interfaces	18000
Port Channel with 10-Gigabit Ethernet interfaces	1800

To change the port cost or priority of an interface:

Task	Command Syntax	Command Mode
Change the port cost of an interface. Range: 0 to 200000 Default: see Table 29-2.	**spanning-tree msti** *number* **cost** *cost*	INTERFACE
Change the port priority of an interface. Range: 0 to 240, in increments of 16 Default: 128	**spanning-tree msti** *number* **priority** *priority*	INTERFACE

View the current values for these interface parameters using the command **show config** from INTERFACE mode. See Figure 29-8.

Configure an EdgePort

The EdgePort feature enables interfaces to begin forwarding traffic approximately 30 seconds sooner. In this mode an interface forwards frames by default until it receives a BPDU that indicates that it should behave otherwise; it does not go through the Learning and Listening states. The **bpduguard shutdown-on-violation** option causes the interface hardware to be shutdown when it receives a BPDU. When only **bpduguard** is implemented, although the interface is placed in an Error Disabled state when receiving the BPDU, the physical interface remains up and spanning-tree will drop packets in the hardware after a BPDU violation. BPDUs are dropped in the software after receiving the BPDU violation. This feature is the same as PortFast mode in Spanning Tree.

△ **Caution:** Configure EdgePort only on links connecting to an end station. EdgePort can cause loops if it is enabled on an interface connected to a network.

To enable EdgePort on an MSTP-enabled interface, use the following command:

Task	Command Syntax	Command Mode
Enable EdgePort on an interface.	**spanning-tree mstp edge-port** [**bpduguard | shutdown-on-violation**]	INTERFACE

Verify that EdgePort is enabled on a port using the command **show config** from the INTERFACE mode, as shown in Figure 29-8.

FTOS Behavior: Regarding **bpduguard shutdown-on-violation** behavior:

1. If the interface to be shutdown is a port channel then all the member ports are disabled in the hardware.
2. When a physical port is added to a port channel already in error disable state, the new member port will also be disabled in the hardware.
3. When a physical port is removed from a port channel in error disable state, the error disabled state is cleared on this physical port (the physical port will be enabled in the hardware).
4. The **reset linecard** command does not clear the error disabled state of the port or the hardware disabled state. The interface continues to be disables in the hardware.

   The error disabled state can be cleared with any of the following methods:

   - Perform an **shutdown** command on the interface.
   - Disable the **shutdown-on-violation** command on the interface (**no spanning-tree** *stp-id* **portfast** [**bpduguard | [shutdown-on-violation**]]).
   - Disable spanning tree on the interface (**no spanning-tree** in INTERFACE mode).
   - Disabling global spanning tree (**no spanning-tree** in CONFIGURATION mode).

Figure 29-8. Configuring EdgePort

```
FTOS(conf-if-gi-3/41)#spanning-tree mstp edge-port
FTOS(conf-if-gi-3/41)#show config
!
interface GigabitEthernet 3/41
 no ip address
 switchport
 spanning-tree mstp edge-port
 spanning-tree MSTI 1 priority 144
 no shutdown
```

Configure a Root Guard

Use the Root Guard feature in a Layer 2 MSTP network to avoid bridging loops.

You enable root guard on a per-port or per-port-channel basis.

FTOS Behavior: The following conditions apply to a port enabled with root guard:

- Root guard is supported on any MSTP-enabled port or port-channel interface except when used as a stacking port.
- Root guard is supported on a port in any Spanning Tree mode:
  - Spanning Tree Protocol (STP)
  - Rapid Spanning Tree Protocol (RSTP)
  - Multiple Spanning Tree Protocol (MSTP)
  - Per-VLAN Spanning Tree Plus (PVST+)
- When enabled on a port, root guard applies to all VLANs configured on the port.
- Root guard and loop guard cannot be enabled at the same time on an MSTP port. For example, if you configure loop guard on a port on which root guard is already configured, the following error message is displayed:
  `% Error: RootGuard is configured. Cannot configure LoopGuard.`
- When used in an MSTP network, if root guard blocks a boundary port in the CIST, the port is also blocked in all other MST instances.

To enable a root guard on an MSTP-enabled port or port-channel interface, enter the **spanning-tree mstp rootguard** command. Refer to STP Root Guard on page 1060 for more information on how to use the root guard feature.

Task	Command Syntax	Command Mode
Enable root guard on a port or port-channel interface.	**spanning-tree mstp rootguard**	INTERFACE INTERFACE PORT-CHANNEL

To disable MSTP root guard on a port or port-channel interface, enter the **no spanning-tree mstp rootguard** command in an interface configuration mode.

To verify the MSTP root guard configuration on a port or port-channel interface, enter the **show spanning-tree msti** [*instance-number*] **guard** command in global configuration mode.

Configure a Loop Guard

The Loop Guard feature provides protection against Layer 2 forwarding loops (STP loops) caused by a hardware failure, such as a cable failure or an interface fault. When a cable or interface fails, a participating STP link may become unidirectional (STP requires links to be bidirectional) and an STP port does not receive BPDUs. When an STP blocking port does not receive BPDUs, it transitions to a forwarding state. This condition can create a loop in the network.

You enable loop guard on a per-port or per-port channel basis.

FTOS Behavior: The following conditions apply to a port enabled with loop guard:

- Loop guard is supported on any MSTP-enabled port or port-channel interface.
- Loop guard is supported on a port or port-channel in any Spanning Tree mode:
  - Spanning Tree Protocol (STP)
  - Rapid Spanning Tree Protocol (RSTP)
  - Multiple Spanning Tree Protocol (MSTP)
  - Per-VLAN Spanning Tree Plus (PVST+)
- Root guard and loop guard cannot be enabled at the same time on an MSTP port. For example, if you configure root guard on a port on which loop guard is already configured, the following error message is displayed:
  `% Error: LoopGuard is configured. Cannot configure RootGuard.`
- Enabling Portfast BPDU guard and loop guard at the same time on a port results in a port that remains in a blocking state and prevents traffic from flowing through it. For example, when Portfast BPDU guard and loop guard are both configured:

  - If a BPDU is received from a remote device, BPDU guard places the port in an err-disabled blocking state and no traffic is forwarded on the port.

  - If no BPDU is received from a remote device, loop guard places the port in a loop-inconsistent blocking state and no traffic is forwarded on the port.

To enable a loop guard on an MSTP-enabled port or port-channel interface, enter the **spanning-tree mstp loopguard** command. Refer to STP Loop Guard on page 1064 for more information on how to use the loop guard feature.

Task	Command Syntax	Command Mode
Enable loop guard on an MSTP-enabled port or port-channel interface.	**spanning-tree mstp loopguard**	INTERFACE INTERFACE PORT-CHANNEL

To disable MSTP loop guard on a port or port-channel interface, enter the **no spanning-tree mstp loopguard** command in an INTERFACE configuration mode.

To verify the MSTP loop guard configuration on a port or port-channel interface, enter the **show spanning-tree msti** [*instance-number*] **guard** command in global configuration mode.

Flush MAC Addresses after a Topology Change

FTOS has an optimized MAC address flush mechanism for RSTP, MSTP, and PVST+ that flushes addresses only when necessary, which allows for faster convergence during topology changes. However, you may activate the flushing mechanism defined by 802.1Q-2003 using the command **tc-flush-standard**, which flushes MAC addresses upon every topology change notification. View the enable status of this feature using the command **show running-config spanning-tree mstp** from EXEC Privilege mode.

Displaying STP Guard Configuration

To verify the STP guard configured on MSTP interfaces, enter the **show spanning-tree msti** [*instance-number*] **guard** command. Refer to Chapter 52, “Spanning Tree Protocol,” on page 1049 for information on how to configure and use the STP root guard, loop guard, and BPDU guard features.

Figure 29-9 shows an example for instance 5 in an MSTP network in which:

- Root guard is enabled on a port that is in a root-inconsistent state.
- Loop guard is enabled on a port that is in a forwarding state.
- BPDU guard is enabled on a port that is shut down.

Figure 29-9. Displaying STP Guard Configuration

```
FTOS#show spanning-tree msti 5 guard
Interface
Name      Instance    Sts          Guard type
---------  --------   ---------    ----------
Gi 0/1     5          INCON(Root)  Root
Gi 0/2     5          FWD          Loop
Gi 0/3     5          EDS(shut)    Bpdu
```

MSTP Sample Configurations

The running-configurations in Figure 29-11, Figure 29-12, and Figure 29-12 support the topology shown in Figure 29-10. The configurations are from FTOS systems. An S50 system using SFTOS, configured as shown Figure 29-14, could be substituted for an FTOS router in this sample following topology and MSTP would function as designed.

Figure 29-10. MSTP with Three VLANs Mapped to Two Spanning Tree Instances

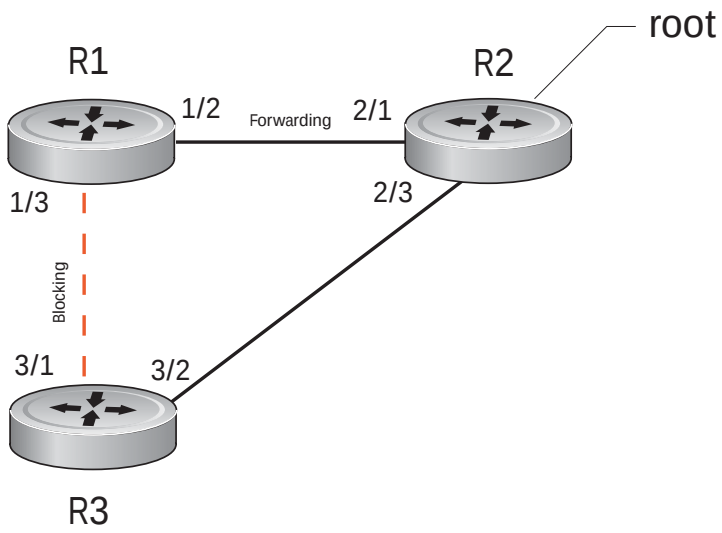


Figure 29-11. Router 1 Running-configuration

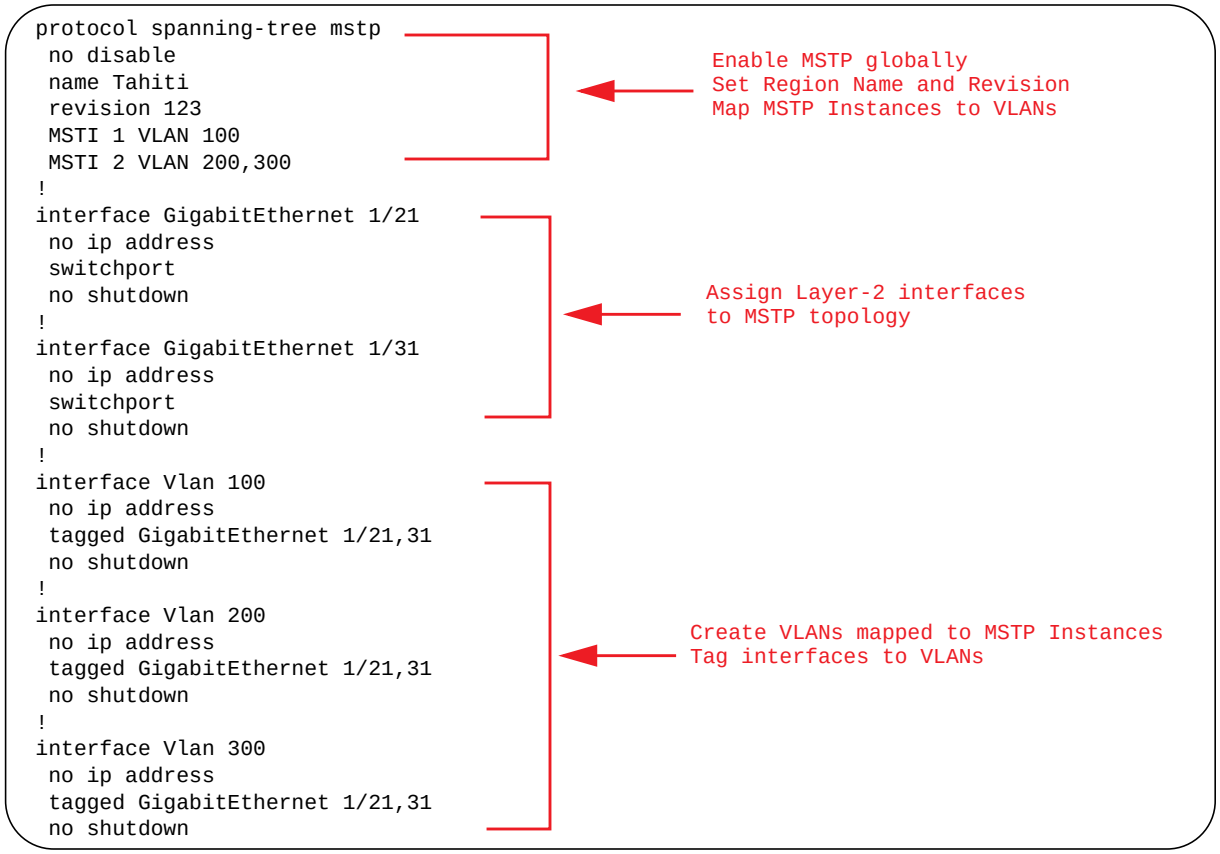


Figure 29-12. Router 2 Running-configuration

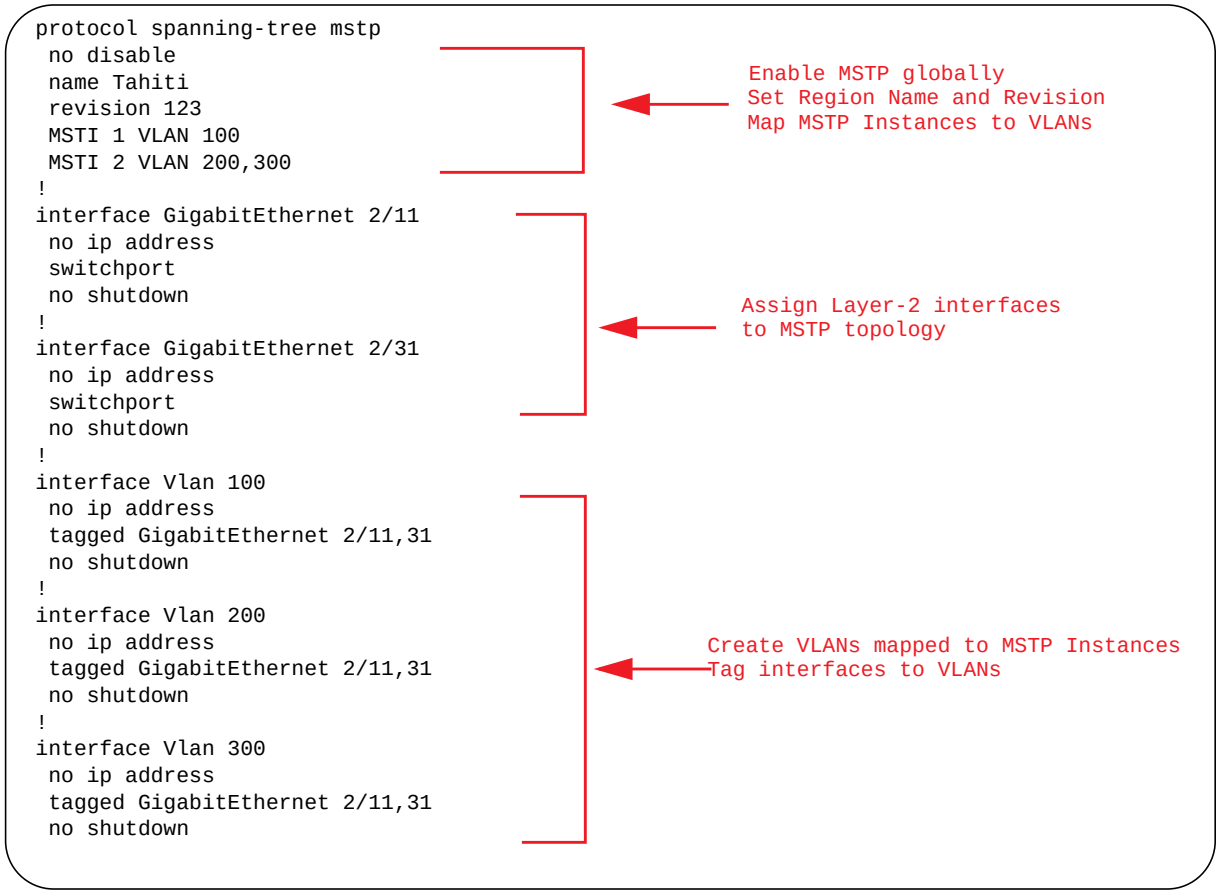


Figure 29-13. Router 3 Running-configuration

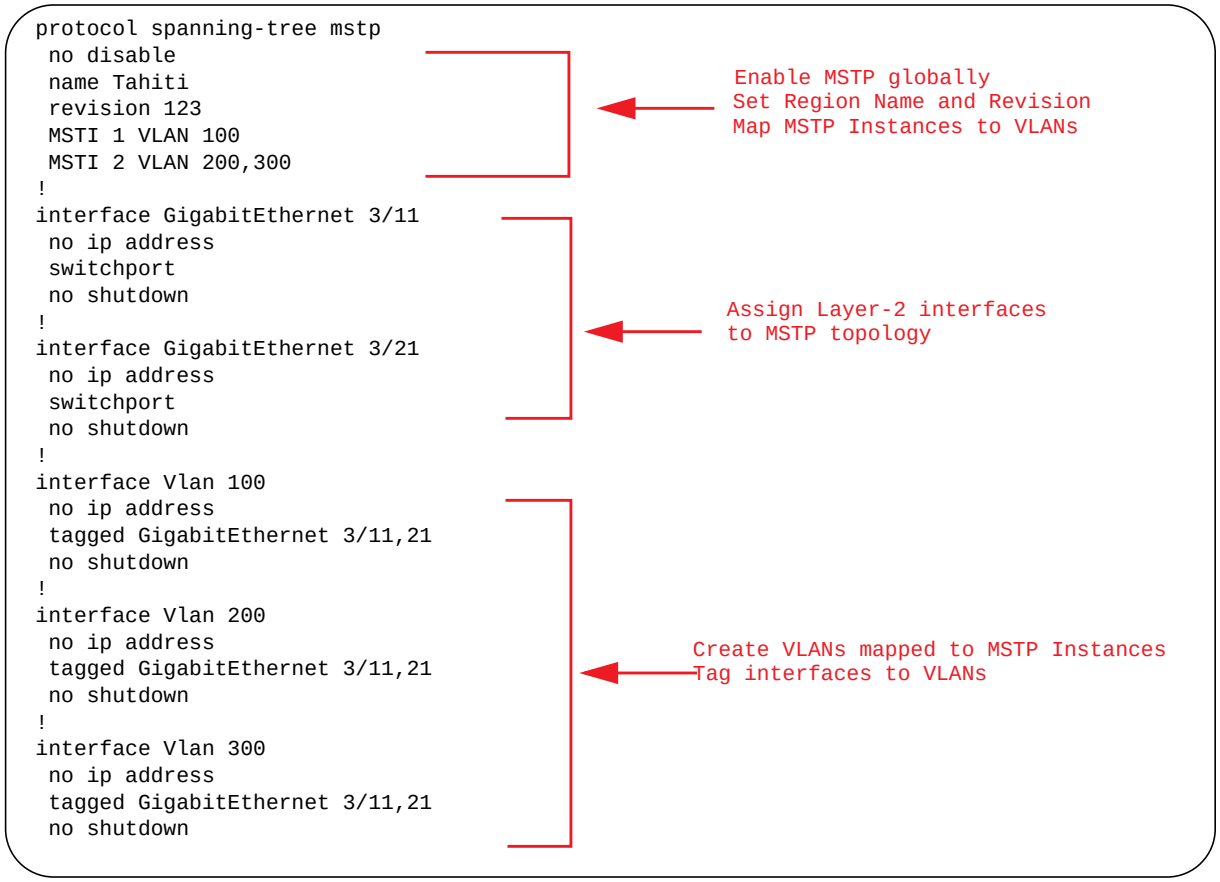
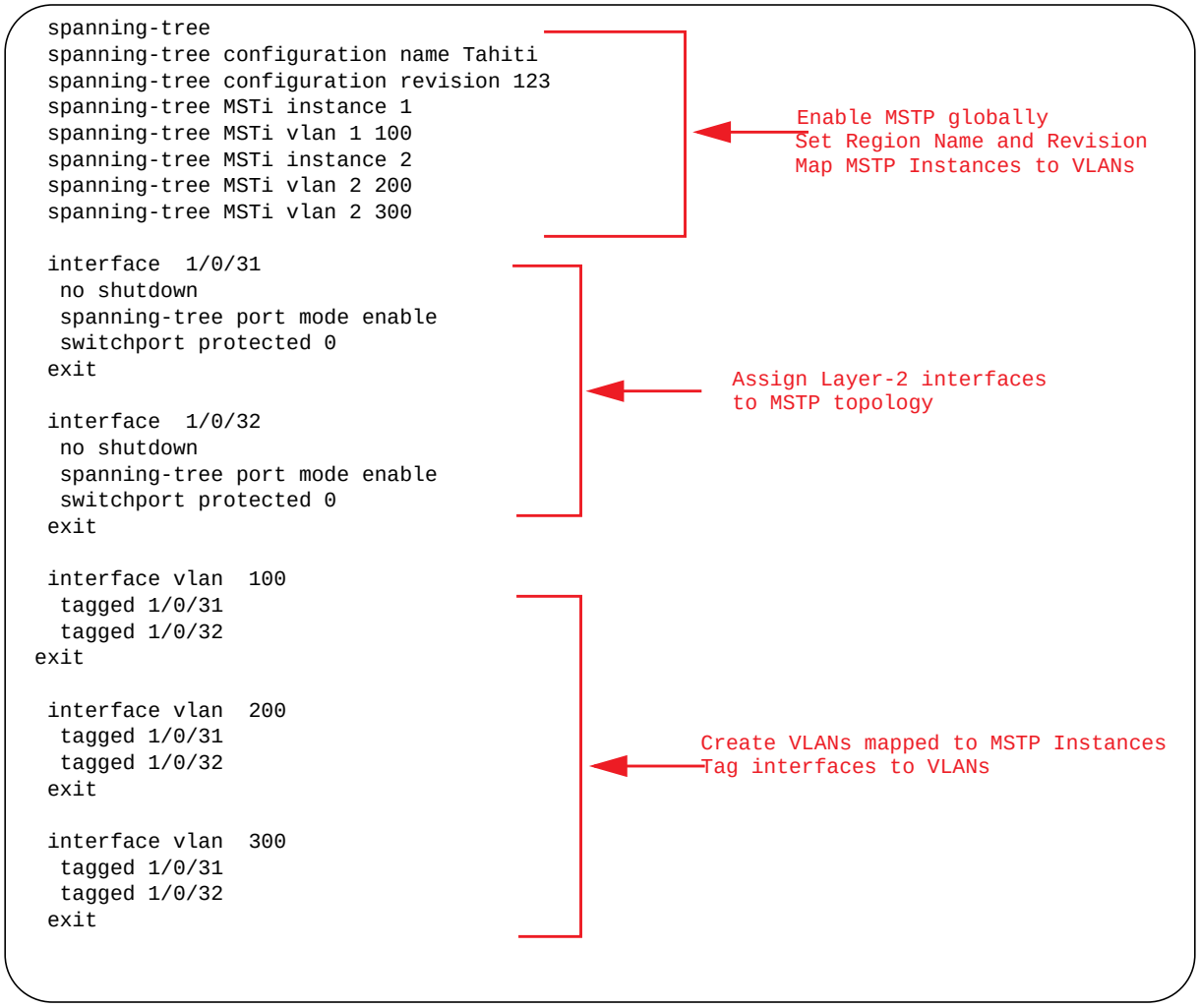


Figure 29-14. SFTOS Example Running-Configuration



Debugging and Verifying MSTP Configuration

Display BPDUs using the command **debug spanning-tree mstp bpdu** from EXEC Privilege mode. Display MSTP-triggered topology change messages **debug spanning-tree mstp events**.

Figure 29-15. Displaying BPDUs and Events

```
FTOS#debug spanning-tree mstp bpdu
1w1d17h : MSTP: Sending BPDU on Gi 1/31 :
ProtId: 0, Ver: 3, Bpdu Type: MSTP, Flags 0x68
CIST Root Bridge Id: 32768:0001.e806.953e, Ext Path Cost: 20000
Regional Bridge Id: 32768:0001.e809.c24a, CIST Port Id: 128:384
Msg Age: 2, Max Age: 20, Hello: 2, Fwd Delay: 15, Ver1 Len: 0, Ver3 Len: 96
Name: my-mstp-region, Rev: 0, Int Root Path Cost: 20000
Rem Hops: 19, Bridge Id: 32768:0001.e80d.b6d6
E1200#1w1d17h : INST 1: Flags: 0x28, Reg Root: 32768:0001.e809.c24a, Int Root Co
         Brg/Port Prio: 32768/128, Rem Hops: 19
INST 2: Flags: 0x68, Reg Root: 4096:0001.e809.c24a, Int Root Cost: 20000
         Brg/Port Prio: 32768/128, Rem Hops: 19
[output omitted]
FTOS#debug spanning-tree mstp events
1w1d17h : MSTP: TC flag set in the incoming BPDU on port Gi 1/31 for instance 0

1w1d17h : MSTP: TC flag set in the incoming BPDU on port Gi 1/31 for instance 0

1w1d17h : MSTP: TC flag set in the incoming BPDU on port Gi 1/31 for instance 0
```

Examine your individual routers to ensure all the necessary parameters match.

1. Region Name
2. Region Version
3. VLAN to Instance mapping

The **show spanning-tree mst** commands will show various portions of the MSTP configuration. To view the overall MSTP configuration on the router, use the **show running-configuration spanning-tree mstp** in the EXEC Privilege mode (output sample shown in Figure 29-16).

Use the **debug spanning-tree mstp bpdu** command to monitor and verify that the MSTP configuration is connected and communicating as desired (output sample shown in Figure 29-17).

Key items to look for in the debug report:

- MSTP flags indicate communication received from the same region.
  - In Figure 29-17, the output shows that the MSTP routers are located in the same region.
  - Does the debug log indicate that packets are coming from a “Different Region” (Figure 29-18)? If so, one of the key parameters is not matching.
- MSTP Region Name and Revision
  - The configured name and revisions *must* be identical among all the routers.
  - Is the Region name blank? That may mean that a name was configured on one router and but was not configured or was configured differently on another router (spelling and capitalization counts).
- MSTP Instances.
  - Use the show commands to verify the VLAN to MSTP Instance mapping.
  - Are there “extra” MSTP Instances in the Sending or Received logs? That may mean that an additional MSTP Instance was configured on one router but not the others.

Figure 29-16. Sample Output for show running-configuration spanning-tree mstp command

```
FTOS#show run spanning-tree mstp
!
protocol spanning-tree mstp
 name Tahiti
 revision 123
 MSTI 1 VLAN 100
 MSTI 2 VLAN 200,300
```

Figure 29-17. Displaying BPDUs and Events - Debug Log of Successful MSTP Configuration

```
FTOS#debug spanning-tree mstp bpdu
MSTP debug bpdu is ON
FTOS#
4w0d4h : MSTP: Sending BPDU on Gi 2/21 :
ProtId: 0, Ver: 3, Bpdu Type: MSTP, Flags 0x6e
CIST Root Bridge Id: 32768:0001.e806.953e, Ext Path Cost: 0
Regional Bridge Id: 32768:0001.e806.953e, CIST Port Id: 128:470
Msg Age: 0, Max Age: 20, Hello: 2, Fwd Delay: 15, Ver1 Len: 0, Ver3 Len: 96
Name: Tahiti, Rev: 123, Int Root Path Cost: 0
Rem Hops: 20, Bridge Id: 32768:0001.e806.953e
4w0d4h : INST 1: Flags: 0x6e, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
         Brg/Port Prio: 32768/128, Rem Hops: 20
INST 2: Flags: 0x6e, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
         Brg/Port Prio: 32768/128, Rem Hops: 20

4w0d4h : MSTP: Received BPDU on Gi 2/21 :
ProtId: 0, Ver: 3, Bpdu Type: MSTP, Flags 0x78Same Region
CIST Root Bridge Id: 32768:0001.e806.953e, Ext Path Cost: 0
Regional Bridge Id: 32768:0001.e806.953e, CIST Port Id: 128:470
Msg Age: 0, Max Age: 20, Hello: 2, Fwd Delay: 15, Ver1 Len: 0, Ver3 Len: 96
Name: Tahiti, Rev: 123, Int Root Path Cost: 0
Rem Hops: 19, Bridge Id: 32768:0001.e8d5.cbbd
4w0d4h : INST 1: Flags: 0x78, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
         Brg/Port Prio: 32768/128, Rem Hops: 19
INST 2: Flags: 0x78, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
         Brg/Port Prio: 32768/128, Rem Hops: 19
```

Indicates MSTP routers are in the (single) region

MSTP Instance

MSTP Region name and revision

Figure 29-18. Displaying BPDUs and Events - Debug Log of Unsuccessful MSTP Configuration

```
4w0d4h : MSTP: Received BPDU on Gi 2/21 :
ProtId: 0, Ver: 3, Bpdu Type: MSTP, Flags 0x78Different Region
CIST Root Bridge Id: 32768:0001.e806.953e, Ext Path Cost: 0
Regional Bridge Id: 32768:0001.e806.953e, CIST Port Id: 128:470
Msg Age: 0, Max Age: 20, Hello: 2, Fwd Delay: 15, Ver1 Len: 0, Ver
Name: Tahiti, Rev: 123, Int Root Path Cost: 0
Rem Hops: 20, Bridge Id: 32768:0001.e8d5.cbbd
4w0d4h : INST 1: Flags: 0x70, Reg Root: 32768:0001.e8d5.cbbd, Int
         Brg/Port Prio: 32768/128, Rem Hops: 20
INST 2: Flags: 0x70, Reg Root: 32768:0001.e8d5.cbbd, Int Root Cost
         Brg/Port Prio: 32768/128, Rem Hops: 20
```

Indicates MSTP routers are in different regions and are not communicating with each other

30

Multicast Features

Multicast Features are supported on platforms: C E S

Multicast is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

This chapter contains the following sections:

FTOS supports the following multicast protocols:

Implementation Information

- Multicast is not supported on secondary IP addresses.

Enable IP Multicast

Enable IP Multicast is supported on platforms C E S

Prior to enabling any multicast protocols, you must enable multicast routing.

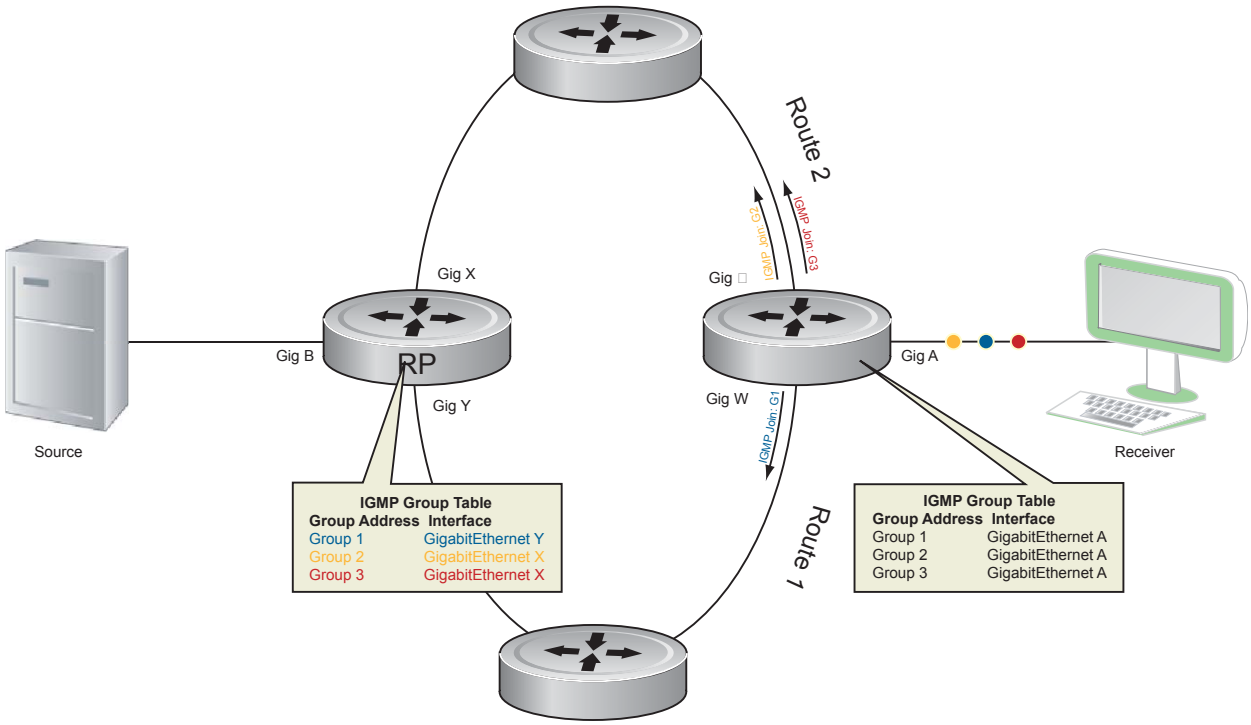
Task	Command Syntax	Command Mode
Enable multicast routing.	**ip multicast-routing**	CONFIGURATION

Multicast with ECMP

Dell Force10 multicast uses Equal-cost Multi-path (ECMP) routing to load-balance multiple streams across equal cost links. When creating the shared-tree Protocol Independent Multicast (PIM) uses routes from all configured routing protocols to select the best route to the rendezvous point (RP). If there are multiple, equal-cost paths, the PIM selects the route with the least number of currently running multicast streams. If multiple routes have the same number of streams, PIM selects the first equal-cost route returned by the Route Table Manager (RTM).

In Figure 30-1, the receiver joins three groups. The last-hop DR initially has two equal-cost routes to the RP with no streams, so it non-deterministically selects Route 1 for the Group 1 IGMP Join message. Route 1 then has one stream associated with it, so the last-hop DR sends the Group 2 Join by Route 2. It then non-deterministically selects Route 2 for the Group 3 Join since both routes already have one multicast stream.

Figure 30-1. Multicast with ECMP



Implementation Information

- Because protocol control traffic in FTOS is redirected using the MAC address, and multicast control traffic and multicast data traffic might map to the same MAC address, FTOS might forward data traffic with certain MAC addresses to the CPU in addition to control traffic.

  As the upper five bits of an IP Multicast address are dropped in the translation, 32 different multicast group IDs all map to the same Ethernet address. For example, 224.0.0.5 is a well known IP address for OSPF that maps to the multicast MAC address 01:00:5e:00:00:05. However, 225.0.0.5, 226.0.0.5, etc., map to the same multicast MAC address. The Layer 2 FIB alone cannot differentiate multicast control traffic multicast data traffic with the same address, so if you use IP address 225.0.0.5 for data traffic, both the multicast data and OSPF control traffic match the same entry and are forwarded to the CPU.

  Therefore, do not use well-known protocol multicast addresses for data transmission, such as the ones below.

Protocol	Ethernet Address
OSPF	01:00:5e:00:00:05 01:00:5e:00:00:06
RIP	01:00:5e:00:00:09
NTP	01:00:5e:00:01:01
VRRP	01:00:5e:00:00:12
PIM-SM	01:00:5e:00:00:0d

- The FTOS implementation of MTRACE is in accordance with IETF draft *draft-fenner-traceroute-ipm*.
- Multicast is not supported on secondary IP addresses.
- Egress L3 ACL is not applied to multicast data traffic if multicast routing is enabled.

Multicast Policies

FTOS offers parallel Multicast features for IPv4 and IPv6.

- IPv4 Multicast Policies on page 665
- IPv6 Multicast Policies on page 673

IPv4 Multicast Policies

- Limit the Number of Multicast Routes on page 666
- Prevent a Host from Joining a Group on page 667
- Rate Limit IGMP Join Requests on page 669
- Prevent a PIM Router from Forming an Adjacency on page 669
- Prevent a Source from Registering with the RP on page 669
- Prevent a PIM Router from Processing a Join on page 670
- Using a Static Multicast MAC Address on page 671

Limit the Number of Multicast Routes

Task	Command Syntax	Command Mode
Limit the total number of multicast routes on the system.	**ip multicast-limit** Range: 1-50000 Default: 15000	CONFIGURATION

When the limit is reached, FTOS does not process any IGMP or MLD joins to PIM—though it still processes leave messages—until the number of entries decreases below 95% of the limit. When the limit falls below 95% after hitting the maximum, the system begins relearning route entries through IGMP, MLD, and MSDP.

- If the limit is increased after it is reached, join subsequent join requests are accepted. In this case, you must increase the limit by at least 10% for IGMP and MLD to resume.
- If the limit is decreased after it is reached, FTOS does not clear the existing sessions. Entries are cleared upon a timeout (you may also clear entries using **clear ip mroute**).

Note: FTOS waits at least 30 seconds between stopping and starting IGMP join processing. You may experience this delay when manipulating the limit after it is reached.

When the multicast route limit is reached, FTOS displays Message 1.

Message 1 Multicast Route Limit Error

```
3w1d13h: %RPM0-P:RP2 %PIM-3-PIM_TIB_LIMIT: PIM TIB limit reached. No new routes will be
learnt until TIB level falls below low watermark.
3w1d13h: %RPM0-P:RP2 %PIM-3-PIM_TIB_LIMIT: PIM TIB below low watermark. Route learning will
begin.
```

Note: The IN-L3-McastFib CAM partition is used to store multicast routes and is a separate hardware limit that is exists per port-pipe. Any software-configured limit might be superseded by this hardware space limitation. The opposite is also true, the CAM partition might not be exhausted at the time the system-wide route limit set by the **ip multicast-limit** is reached.

Prevent a Host from Joining a Group

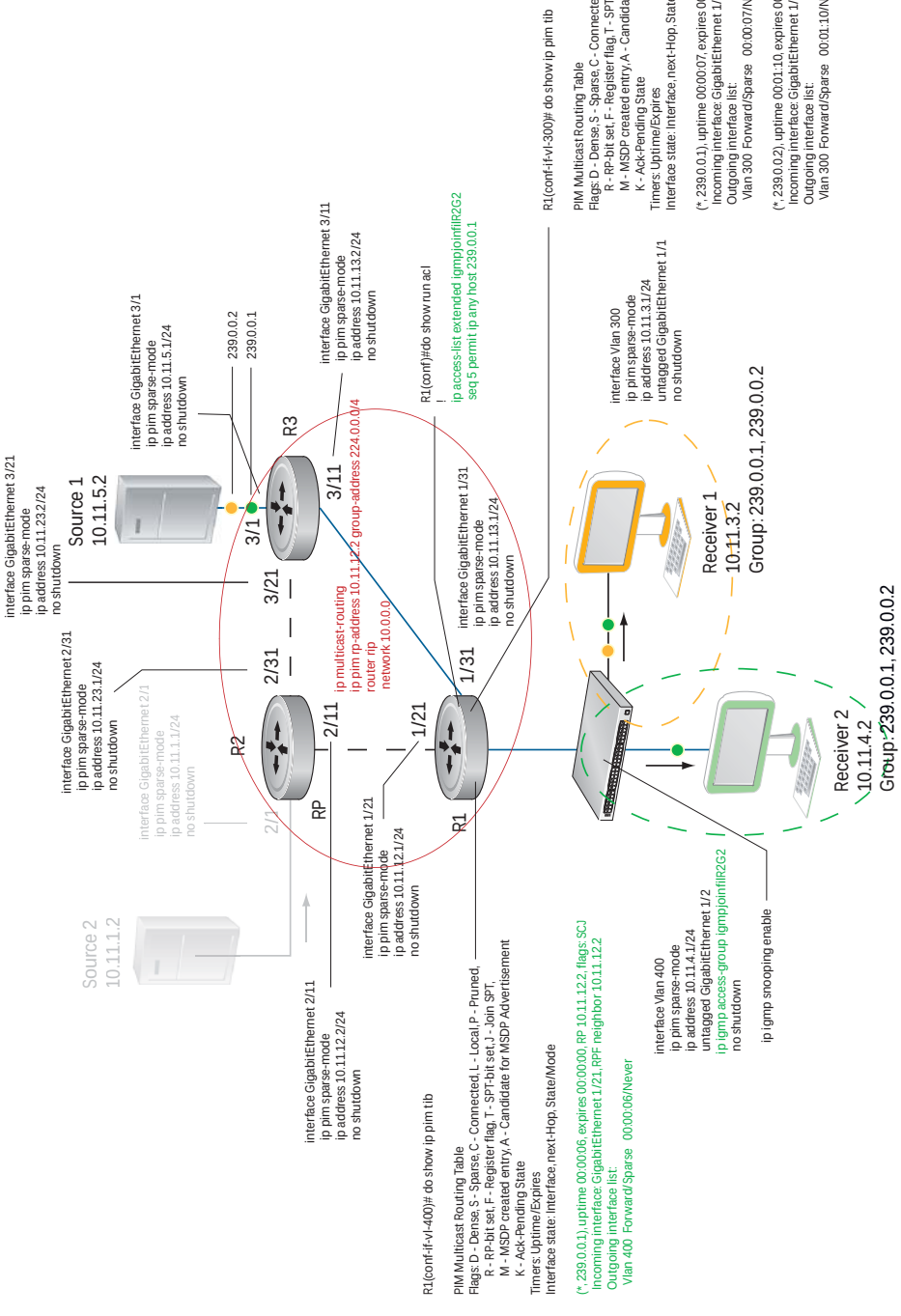
You can prevent a host from joining a particular group by blocking specific IGMP reports. Create an extended access list containing the permissible source-group pairs. Use the command **ip igmp access-group** *access-list-name* from INTERFACE mode to apply the access list.

Note: For rules in IGMP access lists, *source* is the multicast source, not the source of the IGMP packet. For IGMPv2, use the keyword **any** for *source* (as shown in Figure 30-2), since IGMPv2 hosts do not know in advance who the source is for the group in which they are interested.

FTOS Behavior: Do not enter the command **ip igmp access-group** before creating the access-list. If you do, upon entering your first *deny* rule, FTOS clears multicast routing table and re-learns all groups, even those not covered by the rules in the access-list, because there is an implicit *deny all* rule at the end of all access-lists. Therefore, configuring an IGMP join request filter in this order might result in data loss. If you must enter the command **ip igmp access-group** before creating the access-list, prevent FTOS from clearing the routing table by entering a *permit any* rule with high sequence number before you enter any other rules.

In Figure 30-2, VLAN 400 is configured with an access list to permit only IGMP reports for group 239.0.0.1. Though Receiver 2 sends a membership report for groups 239.0.0.1 and 239.0.0.2, a multicast routing table entry is created only for group 239.0.0.1. VLAN 300 has no access list limiting Receiver 1, so both IGMP reports are accepted, and two corresponding entries are created in the routing table.

Figure 30-2. Preventing a Host from Joining a Group



Rate Limit IGMP Join Requests

If you expect a burst of IGMP Joins, protect the IGMP process from overload by limiting that rate at which new groups can be joined using the command **ip igmp group-join-limit** from INTERFACE mode. Hosts whose IGMP requests are denied will use the retry mechanism built-in to IGMP so that they're membership is delayed rather than permanently denied.

View the enable status of this feature using the command **show ip igmp interface** from EXEC Privilege mode.

Prevent a PIM Router from Forming an Adjacency

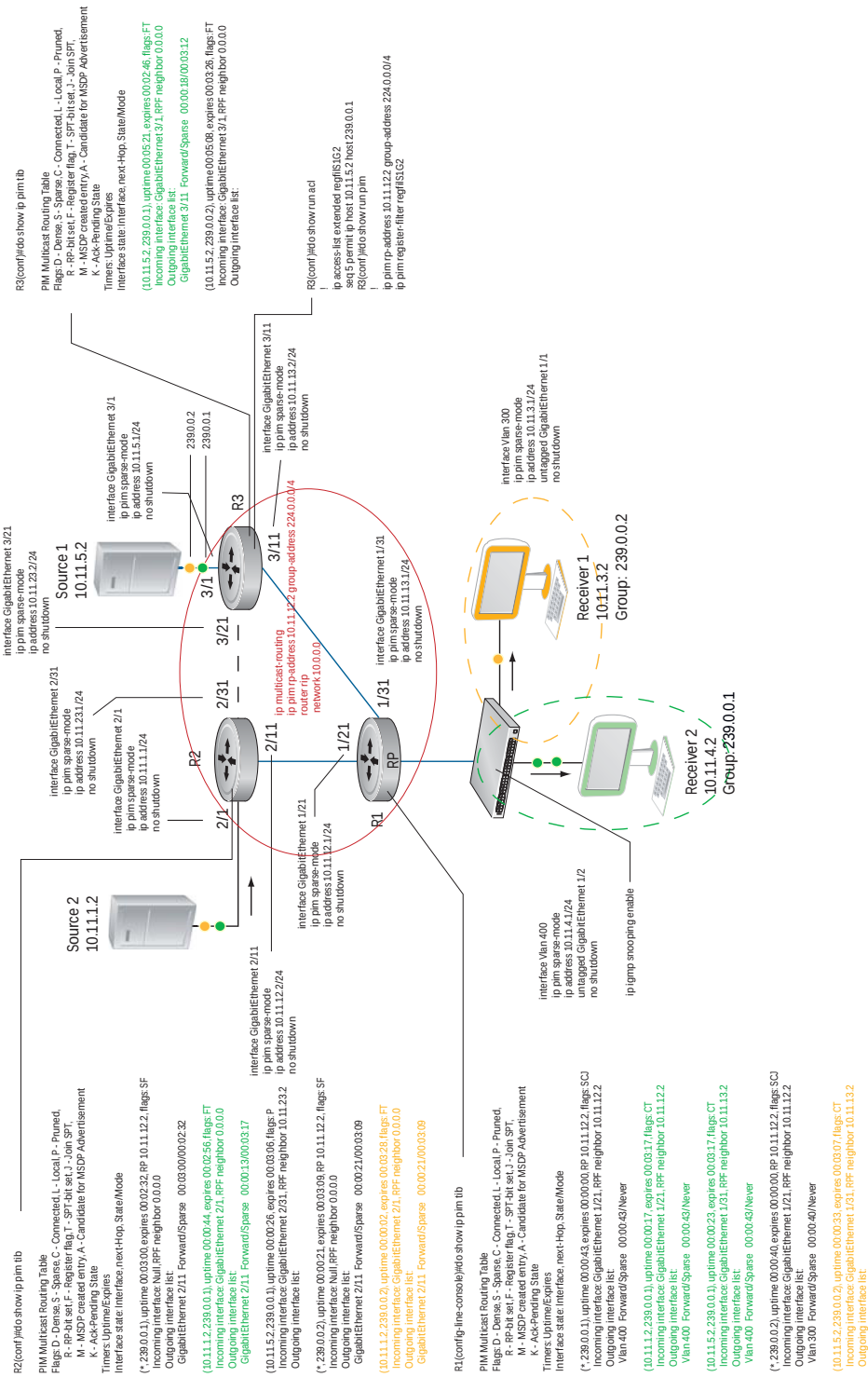
To prevent a router from participating in Protocol Independent Multicast (PIM) (for example, to configure stub multicast routing), use the **ip pim neighbor-filter** command from INTERFACE mode.

Prevent a Source from Registering with the RP

Use the command **ip pim register-filter** from CONFIGURATION mode to prevent a source from transmitting to a particular group. This command prevents the PIM source DR from sending register packets to RP for the specified multicast source and group; if the source DR never sends register packets to the RP, no hosts can ever discover the source and create an SPT to it.

In Figure 30-3, Source 1 and Source 2 are both transmitting packets for groups 239.0.0.1 and 239.0.0.2. R3 has a PIM register filter that only permits packets destined for group 239.0.0.2. An entry is created for group 239.0.0.1 in the routing table, but no outgoing interfaces are listed. R2 has no filter, so it is allowed to forward both groups. As a result, Receiver 1 receives only one transmission, while Receiver 2 receives duplicate transmissions.

Figure 30-3. Preventing a Source from Transmitting to a Group



Prevent a PIM Router from Processing a Join

Permit or deny PIM Join/Prune messages on an interface using an extended IP access list. Use the command **ip pim join-filter** to prevent the PIM SM router from creating state based on multicast source and/or group.

Using a Static Multicast MAC Address

Using a Static Multicast MAC Address is supported on platform C

When a multicast source and multicast receivers are in the same VLAN, you can configure a router so that multicast traffic is switched only to the ports assigned to a VLAN that is associated with a static multicast MAC address. This task improves router performance by limiting the output ports to which multicast data is sent.

To enable a router to switch multicast traffic in a Layer 2 VLAN only to the ports associated with a static multicast MAC address, you must:

1. Enable the router for Layer 2 multicast switching.
2. Configure a static MAC address and associate it with a Layer 2 VLAN used to switch multicast traffic on the router to output ports assigned to the VLAN.

To enable Layer 2 switching of multicast traffic, follow these steps:

Step	Task	Command Syntax	Command Mode
1	Enable Layer 2 multicast switching.	**ip multicast-mode l2**	CONFIGURATION
	Note: Enabling Layer 2 multicast switching automatically disables default Layer 3 multicast routing on the router.		
2	Configure a static multicast MAC address, associate the multicast MAC address with the VLAN used to switch Layer 2 multicast traffic, and add output ports that will receive multicast streams to the VLAN.	**mac-address-table static** *multicast-mac-address* **multicast vlan** *vlan-id* **range-output {***single-interface* | *interface-list* | *interface-range***}**	CONFIGURATION
	Note: You can add individual or a range of ports to the VLAN used for Layer 2 multicast forwarding as follows: **range-output** *single-interface* specifies one of the following port types: - 1-Gigabit Ethernet: Enter **gigabitethernet** *slot*/*port*. - 10-Gigabit Ethernet: Enter **tengigabitethernet** *slot*/*port*. - Port channel: Enter **port-channel** {1-128}. **range-output** *interface-list* specifies multiple ports separated by a space, comma and space; for example: **tengigabitethernet 0/1** , **gigabitethernet 0/3** , ... **range-output** *interface-range* specifies a port range in the format: *interface-type slot/first_port - last_port*; for example: **tengigabitethernet 0/1 - 3**		

To return to the default Layer 3 multicast forwarding on the router, enter the **no ip multicast-mode l2** command after you remove the static multicast MAC address with the **no mac-address-table static multicast vlan output-range** command.

To display the current configuration of Layer 2 multicast switching on a router, enter the **show mac-address-table static multicast [vlan** *vlan-id* **|** *multicast-mac-address* **[vlan** *vlan-id***]]** command in EXEC mode. Static MAC addresses configured for Layer 2 multicast forwarding with an associated VLAN and assigned output ports are displayed as shown in Figure 30-4.

Figure 30-4. show mac-address-table static multicast Command Output

```
FTOS# show mac-address-table static multicast

VlanId     Mac Address           Type   State  L2MCIndex  Interfaces
    10    01:00:5e:01:01:01     static Active          0  Gi 1/2,
                                                        Gi 2/47

    11    01:00:5e:01:01:02     static Active          1  Po 10

    12    01:00:5e:01:01:01     static Inactive        0
```

The state of a static multicast MAC address is inactive if the associated VLAN has not been configured.

Figure 30-5. show mac-address-table static multicast vlan Command Output

```
FTOS# show mac-address-table static multicast vlan 10

VlanId     Mac Address           Type   State  L2MCIndex   Interfaces
    10    01:00:5e:01:01:01     static Active          0    Gi 1/2,
                                                           Gi 2/47
```

To display the number of static multicast MAC addresses in use for all VLANs on a router, enter the **show mac-address-table static multicast count [vlan** *vlan-id***]** command in EXEC mode.

Figure 30-6. show mac-address-table static multicast count Command Example

```
FTOS#show mac-address-table static multicast count
```

Figure 30-7. show mac-address-table static multicast count vlan Command Example

```
FTOS#show mac-address-table static multicast count vlan 10
```

IPv6 Multicast Policies

IPv6 Multicast Policies is available only on platform: E

Limit the Number of IPv6 Multicast Routes

You can limit the total number of IPv6 multicast routes on the system. The maximum number of multicast entries allowed on each line card is determined by the CAM profile. Multicast routes are stored in the IN-V6-McastFib CAM region, which has a fixed number of entries. Any limit configured via the CLI is superseded by this hardware limit. The opposite is also true; the CAM might not be exhausted at the time the CLI-configured route limit is reached.

Task	Command Syntax	Command Mode
Limit the total number of IPv6 multicast routes on the system.	**ipv6 multicast-limit** Range: 1-50000 Default: 15000	CONFIGURATION

Prevent an IPv6 Neighbor from Forming an Adjacency

Task	Command Syntax	Command Mode
Prevent a router from participating in PIM.	**ipv6 pim neighbor-filter** *access-list*	CONFIGURATION

```
FTOS(conf)#ipv6 pim neighbor-filter NEIGH_ACL
FTOS(conf)#ipv6 access-list NEIGH_ACL
FTOS(conf-ipv6-acl)#show config
!
ipv6 access-list NEIGH_ACL
 seq 5 deny ipv6 host fe80::201:e8ff:fe0a:5ad any
 seq 10 permit ipv6 any any
FTOS(conf-ipv6-acl)#
```

Prevent an IPv6 Source from Registering with the RP

Task	Command Syntax	Command Mode
Configured on the source DR, prevent the source DR from sending register packets to the RP for specific sources and groups.	**ipv6 pim register-filter** *access-list*	CONFIGURATION

```
FTOS(conf)#ipv6 pim register-filter REG-FIL_ACL
FTOS(conf)#ipv6 access-list REG-FIL_ACL
FTOS(conf-ipv6-acl)#deny ipv6 165:87:34::10/128 ff0e::225:1:2:0/112
FTOS(conf-ipv6-acl)#permit ipv6 any any
FTOS(conf-ipv6-acl)#exit
```

Prevent an IPv6 PIM Router from Processing an IPv6 Join

Task	Command Syntax	Command Mode
Permit or deny PIM Join/Prune messages on an interface using an access list. This command prevents the PIM-SM router from creating state based on multicast source and/or group.	**ipv6 pim join-filter***access-list* [**in** | **out**]	INTERFACE

```
FTOS(conf)#ipv6 access-list JOIN-FIL_ACL
FTOS(conf-ipv6-acl)#permit ipv6 165:87:34::0/112 ff0e::225:1:2:0/112
FTOS(conf-ipv6-acl)#permit ipv6 any ff0e::230:1:2:0/112
FTOS(conf-ipv6-acl)#permit ipv6 165:87:32::0/112 any
FTOS(conf-ipv6-acl)#exit
FTOS(conf)#interface gigabitethernet 0/84
FTOS(conf-if-gi-0/84)#ipv6 pim join-filter JOIN-FIL_ACL in
FTOS(conf-if-gi-0/84)#ipv6 pim join-filter JOIN-FIL_ACL out
```

Multicast Traceroute

Multicast Traceroute is supported only on platform: E

MTRACE is an IGMP-based tool that prints that network path that a multicast packet takes from a source to a destination, for a particular group. FTOS has mtrace client and mtrace transmit functionality.

- **MTRACE Client**—an mtrace client transmits mtrace queries and prints out the details received responses.

- **MTRACE Transit**—when a Dell Force10 system is an intermediate router between the source and destination in an MTRACE query, FTOS computes the RPF neighbor for the source, fills in the request, and forwards the request to the RPF neighbor. While computing the RPF neighbor, static mroutes and mBGP routes are preferred over unicast routes. When a Dell Force10 system is the last hop to the destination, FTOS sends a response to the query.

Task	Command Syntax	Command Mode
Print the network path that a multicast packet takes from a multicast source to receiver, for a particular group.	**mtrace** *multicast-source-address multicast-receiver-address multicast-group-address*	EXEC Privilege

Figure 30-8. Tracing a Multicast Route

```
FTOS#mtrace 10.11.5.2 10.11.3.2 239.0.0.1
Type Ctrl-C to abort.
Mtrace from 10.11.5.2 to 10.11.3.2 via group 239.0.0.1
From source (?) to destination (?)
Querying full reverse path...
 0  10.11.3.2
-1  10.11.3.1  PIM  Reached RP/Core [default]
-2  10.11.5.2
```

Multicast Quality of Service

Multicast Quality of Service is supported only on platform: E

The Quality of Service (QoS) features available for unicast traffic can be applied to multicast flows. The following QoS features are available:

- Policy-based QoS—Classifying, rate policing, and marking ingress traffic
- WRED
  - See also Allocate More Buffer Memory for Multicast WRED on page 676.

Optimize the E-Series for Multicast Traffic

Optimize the E-Series for Multicast Traffic is supported only on platform: E

The default hardware settings for the E-series are for unicast applications like data centers and ISP networks. This means that the E-Series gives priority to unicast data forwarding rather than multicast data forwarding. For multicast intensive applications like trading, Dell Force10 recommends reconfiguring some default settings.

You may do one or more for the following to optimize the E-Series for your multicast application:

- Tune the Central Scheduler for Multicast on page 676

Allocate More Buffer Memory for Multicast WRED

Allocate more buffer memory to multicast WRED (Weighted Random Early Detection) for bursty multicast traffic that might temporarily become oversubscribed. For example, the example WRED profile in Figure 41-14 on page 872 allocates multicast traffic a minimum of 40 megabytes (out of 80 megabytes) of buffer memory and up to 60 megabytes.

Figure 30-9. Allocating More Bandwidth for Multicast WRED

```
FTOS(Conf)#queue egress multicast linecard all wred-profile Egress
FTOS(conf)#wred-profile Egress
FTOS(conf-wred)# threshold min 40960 max 61440
```

Allocate More Bandwidth to Multicast using Egress WFQ

Egress Weighted Fair Queuing (WFQ) determines per port the ratio of egress bandwidth allocated to multicast, replication, and unicast traffic. By default, FTOS provides 1/64 to multicast, 1/64 to replication, and 62/64 for unicast, which is shared between 8 unicast queues. Allocate more bandwidth for multicast using the command **queue egress multicast linecard** (from CONFIGURATION mode) with the keyword **bandwidth-percent**. For example, allocate 80% of egress bandwidth to multicast on all line cards using the command **queue egress multicast linecard all bandwidth-percent 80**.

Tune the Central Scheduler for Multicast

The Central Scheduler is responsible for scheduling unicast and multicast packets via the Terabit backplane. The default configuration of the Central Scheduler is optimized for network environments that forward primarily unicast traffic—80% of the scheduler weight is for unicast traffic and 20% is for multicast traffic.

FTOS provides the ability to adjust the scheduling weight for multicast traffic. For example, if the majority of your traffic is multicast, the default configuration might yield greater latency. In this case, allocate more backplane bandwidth for multicast using the command **queue multicast bandwidth-percent** from CONFIGURATION mode. View your configuration using the command **show queue backplane multicast bandwidth-percentage**.

Figure 30-10. Tuning the Central Scheduler for Multicast

```
FTOS#show queue backplane multicast bandwidth-percent
Configured multicast bandwidth percentage is 80
```

31

Object Tracking

IPv4/IPv6 Object Tracking is available on platforms: C E S

This chapter covers the following information:

- Object Tracking Overview
- Object Tracking Configuration
- Displaying Tracked Objects

Object tracking allows FTOS client processes, such as VRRP, to monitor tracked objects (for example, interface or link status) and take appropriate action when the state of an object changes.

Note: In release 8.4.1.0, object tracking is supported only on VRRP.

Object Tracking Overview

Object tracking allows you to define objects of interest, monitor their state, and report to a client when a change in an object's state occurs. The following tracked objects are supported:

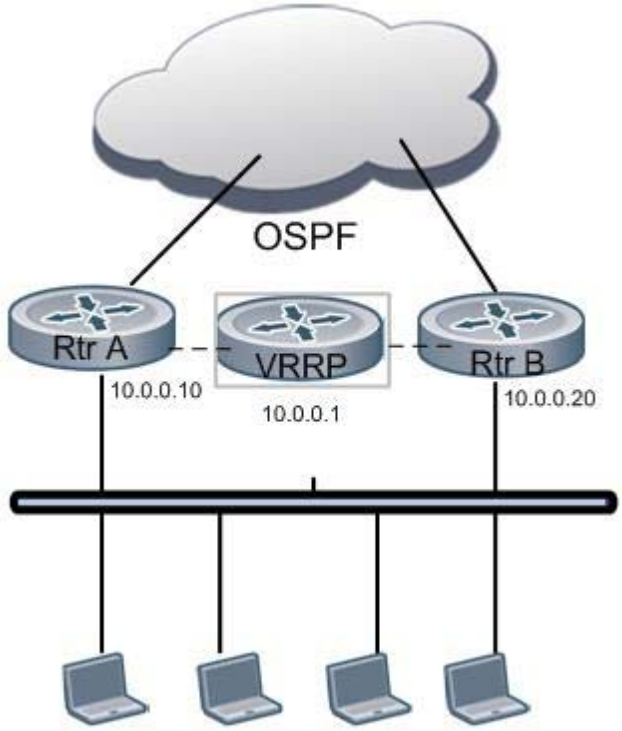
- Link status of Layer 2 interfaces
- Routing status of Layer 3 interfaces (IPv4 and IPv6)
- Reachability of IPv4 and IPv6 routes
- Metric thresholds of IPv4 and IPv6 routes

In future releases, environmental alarms and available free memory will be supported. You can configure client applications, such VRRP, to receive a notification when the state of a tracked object changes.

For example, Figure 31-1 shows how object tracking is performed. Router A and Router B are both connected to the Internet via interfaces running OSPF. Both routers belong to a VRRP group with a virtual router at 10.0.0.1 on the LAN side. Neither Router A nor Router B is the owner of the group. Although Router A and Router B use the same default VRRP priority (100), Router B would normally become the master for the VRRP group because it has a higher IP address.

You can create a tracked object to monitor the metric of the default route 0.0.0.0/0. After you configure the default route as a tracked object, you can configure the VRRP group to track the state of the route. In this way, the VRRP priority of the router with the better metric as determined by OSPF automatically becomes master of the VRRP group. Later, if network conditions change and the cost of the default route in each router changes, the mastership of the VRRP group is automatically reassigned to the router with the better metric.

Figure 31-1. Object Tracking Example



When you configure a tracked object, such as an IPv4/IPv6 a route or interface, you specify an object number to identify the object. Optionally, you can also specify:

- UP and DOWN thresholds used to report changes in a route metric
- A time delay before changes in a tracked object's state are reported to a client

Tracking Layer 2 Interfaces

You can create an object to track the line-protocol state of a Layer 2 interface. In this type of object tracking, the link-level operational status (UP or DOWN) of the interface is monitored.

When the link-level status goes down, the tracked resource status is considered to be DOWN; if the link-level status goes up, the tracked resource status is considered to be UP. For logical interfaces, such as port-channels or VLANs, the link-protocol status is considered to be UP if any physical interface under the logical interface is UP.

Tracking Layer 3 Interfaces

You can create an object that tracks the Layer 3 state (IPv4 or IPv6 routing status) of an interface.

- The Layer 3 status of an interface is UP only if the Layer 2 status of the interface is UP and the interface has a valid IP address.
- The Layer 3 status of an interface goes DOWN when its Layer 2 status goes down or the IP address is removed from the routing table.

Tracking IPv4 and IPv6 Routes

You can create an object that tracks an IPv4 or IPv6 route entry in the routing table. You specify a tracked route by its IPv4/IPv6 address and prefix-length, and optionally, by a VRF instance name if the route to be tracked is part of a VRF. The next-hop address is not part of the definition of the tracked object.

A tracked route matches a route in the routing table only if the exact address and prefix length match an entry in the routing table. For example, when configured as a tracked route, 10.0.0.0/24 does not match the routing table entry 10.0.0.0/8. If no route-table entry has the exact address and prefix length, the tracked route is considered to be DOWN.

In addition to the entry of a route in the routing table, you can configure how the status of a route is tracked in either the following ways:

- By the reachability of the route's next-hop router
- By comparing the UP or DOWN threshold for a route’s metric with current entries in the route table

Tracking Route Reachability

If you configure the reachability of an IP route entry as a tracked object, the UP/DOWN state of the route is determined by the entry of the next-hop address in the ARP cache. A tracked route is considered to be reachable if there is an ARP cache entry for the route's next-hop address. If the next-hop address in the ARP cache ages out for a route tracked for its reachability, an attempt is made to regenerate the ARP cache entry to see if the next-hop address appears before considering the route DOWN.

Tracking a Metric Threshold

If you configure a metric threshold to track a route, the UP/DOWN state of the tracked route is determined by the current metric for the route entered in the routing table.

To provide a common tracking interface for different clients, route metrics are scaled in the range 0 to 255, where 0 is connected and 255 is inaccessible. The scaled metric value communicated to a client always considers a lower value to have priority over a higher value. The resulting scaled value is compared against the threshold values to determine the state of a tracked route as follows:

- If the scaled metric for a route entry is less than or equal to the UP threshold, the state of a route is UP.

- If the scaled metric for a route is greater than or equal to the DOWN threshold or the route is not entered in the routing table, the state of a route is DOWN.

The UP and DOWN thresholds are user-configurable for each tracked route. The default UP threshold is 254; the default DOWN threshold is 255. The notification of a change in the state of a tracked object is sent when a metric value crosses a configured threshold.

The tracking process uses a protocol-specific resolution value to convert the actual metric in the routing table to a scaled metric in the range 0 to 255. The resolution value is user-configurable and calculates the scaled metric by dividing a route's cost by the resolution value set for the route type:

- For ISIS, you can set the resolution in the range 1 to 1000, where the default is 10.
- For OSPF, you can set the resolution in the range 1 to 1592, where the default is 1.
- The resolution value used to map static routes is not configurable. By default, FTOS assigns a metric of 0 to static routes.
- The resolution value used to map RIP routes is not configurable. The RIP hop-count is automatically multiplied by 16 to scale it; a RIP metric of 16 (unreachable) scales to 256, which considers the route to be DOWN. For example, to configure object tracking for a RIP route to be considered UP only if the RIP hop count is less than or equal to 4, you would configure the UP threshold to be 64 (4 x 16) and the DOWN threshold to be 65.

Setting Tracking Delays

You can configure an optional UP and/or DOWN timer for each tracked object to set the time delay before a change in the state of a tracked object is communicated to clients. The configured time delay starts when the state changes from UP to DOWN or vice-versa.

If the state of an object changes back to its former UP/DOWN state before the timer expires, the timer is cancelled and the client is not notified. If the timer expires and an object’s state has changed, a notification is sent to the client. For example, if the DOWN timer is running when an interface goes down and comes back up, the DOWN timer is cancelled and the client is not notified of the event.

If no delay is configured, a notification is sent immediately as soon as a change in the state of a tracked object is detected. The time delay in communicating a state change is specified in seconds.

VRRP Object Tracking

As a client, VRRP can track up to twenty objects (including route entries, and Layer 2 and Layer 3 interfaces) in addition to the twelve tracked interfaces supported for each VRRP group.

You can assign a unique priority-cost value from 1 to 254 to each tracked VRRP object or group interface. The priority cost is subtracted from the VRRP group priority if a tracked VRRP object is in a DOWN state. If a VRRP group router acts as owner-master, the run-time VRRP group priority remains fixed at 255 and changes in the state of a tracked object have no effect. For more information on how to track a VRRP object, see Track an Interface or Object on page 1139.

Note: In VRRP object tracking, the sum of the priority costs for all tracked objects and interfaces cannot equal or exceed the priority of the VRRP group.

Object Tracking Configuration

You can configure the following types of object tracking for a client:

- Tracking a Layer 2 Interface on page 681
- Tracking a Layer 3 Interface on page 682
- Tracking an IPv4/IPv6 Route on page 684

For a complete listing of all commands related to object tracking, refer to the *FTOS Command Line Interface*.

Tracking a Layer 2 Interface

You can create an object that tracks the line-protocol state of a Layer 2 interface and monitors its operational status (UP or DOWN). You can track the status of any of the following Layer 2 interfaces:

- 1-Gigabit Ethernet: Enter **gigabitethernet** *slot/port* in the **track interface** *interface* command (see Step 1 below).
- 10-Gigabit Ethernet: Enter **tengigabitethernet** *slot/port*.
- Port channel: Enter **port-channel** *number*, where valid port-channel numbers are:
  - For the C-Series and S-Series, 1 to 128
  - For the E-Series, 1 to 32 (EtherScale) and 1 to 255 (TeraScale and ExaScale)
- SONET: Enter **sonet** *slot/port*.
- VLAN: Enter **vlan** *vlan-id*, where valid VLAN IDs are from 1 to 4094.

A line-protocol object only tracks the link-level (UP/DOWN) status of a specified interface. When the link-level status goes down, the tracked object status is considered to be DOWN; if the link-level status is up, the tracked object status is considered to be UP.

To configure object tracking on the status of a Layer 2 interface, use the following commands. To remove object tracking on a Layer 2 interface, enter the **no track** *object-id* command.

Step	Task	Command Syntax	Command Mode
1	Configure object tracking on the line-protocol state of a Layer 2 interface.	**track** *object-id* **interface** *interface* **line-protocol** Valid object IDs are from 1 to 65535.	CONFIGURATION
2	(Optional) Configure the time delay used before communicating a change in the status of a tracked interface.	**delay** {[**up** *seconds*] [**down** *seconds*]} Valid delay times are from 0 to 180 seconds. Default: 0.	OBJECT TRACKING
3	(Optional) Identify the tracked object with a text description.	**description** *text* The text string can be up to 80 characters.	OBJECT TRACKING
4	(Optional) Display the tracking configuration and the tracked object's status.	**show track** *object-id*	EXEC Privilege

Figure 31-2. Command Example: track interface line-protocol

```
FTOS(conf)#track 100 interface gigabitethernet 7/1 line-protocol
FTOS(conf-track-100)#delay up 20
FTOS(conf-track-100)#description San Jose data center
FTOS(conf-track-100)#end
FTOS#show track 100

Track 100
  Interface GigabitEthernet 7/1 line-protocol
  Description: San Jose data center
  Line protocol is Up
   2 changes, last change 00:03:05
  Tracked by:
```

Tracking a Layer 3 Interface

You can create an object that tracks the routing status of an IPv4 or IPv6 Layer 3 interface. You can track the routing status of any of the following Layer 3 interfaces:

- 1-Gigabit Ethernet: Enter **gigabitethernet** *slot/port* in the **track interface** *interface* command (see Step 1 below).
- 10-Gigabit Ethernet: Enter **tengigabitethernet** *slot/port*.
- Port channel: Enter **port-channel** *number*, where valid port-channel numbers are:
  - For the C-Series and S-Series, 1 to 128
  - For the E-Series, 1 to 32 (EtherScale) and 1 to 255 (TeraScale and ExaScale)
- SONET: Enter **sonet** *slot/port*.
- VLAN: Enter **vlan** *vlan-id*, where valid VLAN IDs are from 1 to 4094.

For an IPv4 interface, a routing object only tracks the UP/DOWN status of the specified IPv4 interface (**track interface ip-routing** command).

- The status of an IPv4 interface is UP only if the Layer 2 status of the interface is UP and the interface has a valid IP address.
- The Layer 3 status of an IPv4 interface goes DOWN when its Layer 2 status goes down (for a Layer 3 VLAN, all VLAN ports must be down) or the IP address is removed from the routing table.

For an IPv6 interface, a routing object only tracks the UP/DOWN status of the specified IPv6 interface (**track interface ipv6-routing** command).

- The status of an IPv6 interface is UP only if the Layer 2 status of the interface is UP and the interface has a valid IPv6 address.
- The Layer 3 status of an IPv6 interface goes DOWN when its Layer 2 status goes down (for a Layer 3 VLAN, all VLAN ports must be down) or the IPv6 address is removed from the routing table.

To configure object tracking on the routing status of a Layer 3 interface, use the following commands. To remove object tracking on a Layer 3 IPv4/IPv6 interface, enter the **no track object-id** command.

Step	Task	Command Syntax	Command Mode
1	Configure object tracking on the routing status of an IPv4 or IPv6 interface.	**track** *object-id* **interface** *interface* {**ip routing** | **ipv6 routing**} Valid object IDs are from 1 to 65535.	CONFIGURATION
2	(Optional) Configure the time delay used before communicating a change in the status of a tracked interface.	**delay** {[**up** *seconds*] [**down** *seconds*]} Valid delay times are from 0 to 180 seconds. Default: 0.	OBJECT TRACKING
3	(Optional) Identify the tracked object with a text description.	**description** *text* The text string can be up to 80 characters.	OBJECT TRACKING
4	(Optional) Display the tracking configuration and the tracked object's status.	**show track** *object-id*	EXEC Privilege

Figure 31-3. Command Example: track interface ip routing

```
FTOS(conf)#track 101 interface gigabitethernet 7/2 ip routing
FTOS(conf-track-101)#delay up 20
FTOS(conf-track-101)#description NYC metro
FTOS(conf-track-101)#end
FTOS#show track 101

Track 101
  Interface GigabitEthernet 7/2 ip routing
  Description: NYC metro
  IP routing is Down (shutdown)
   2 changes, last change 00:03:23
  Tracked by:
```

Figure 31-4. Command Example: track interface ipv6 routing

```
FTOS(conf)#track 103 interface gigabitethernet 7/11 ipv6 routing
FTOS(conf-track-103)#description Austin access point
FTOS(conf-track-103)#end
FTOS#show track 103

Track 103
  Interface GigabitEthernet 7/11 ipv6 routing
  Description: Austin access point
  IPv6 routing is Down (shutdown)
   2 changes, last change 00:03:25
  Tracked by:
```

Tracking an IPv4/IPv6 Route

You can create an object that tracks the reachability or metric of an IPv4 or IPv6 route. You specify the route to be tracked by its address and prefix-length values. Optionally, for an IPv4 route you can enter a VRF instance name if the route is part of a VPN routing and forwarding (VRF) table. The next-hop address is not part of the definition of a tracked IPv4/IPv6 route.

In order for an route's reachability or metric to be tracked, the route must appear as an entry in the routing table. A tracked route is considered to match an entry in the routing table only if the exact IPv4 or IPv6 address and prefix length match an entry in the table. For example, when configured as a tracked route, 10.0.0.0/24 does not match the routing table entry 10.0.0.0/8. Similarly, for an IPv6 address, 3333:100:200:300:400::/80 does not match routing table entry 3333:100:200:300::/64. If no route-table entry has the exact IPv4/IPv6 address and prefix length, the tracked route is considered to be DOWN.

In addition to the entry of a route in the routing table, you can configure the UP/DOWN state of a tracked route to be determined in the following ways:

- By the reachability of the route's next-hop router

  The UP/DOWN state of the route is determined by the entry of the next-hop address in the ARP cache. A tracked route is considered to be reachable if there is an ARP cache entry for the route's next-hop address. If the next-hop address in the ARP cache ages out for a route tracked for its reachability, an attempt is made to regenerate the ARP cache entry to see if the next-hop address appears before considering the route DOWN.

- By comparing the threshold for a route's metric with current entries in the route table

  The UP/DOWN state of the tracked route is determined by the threshold for the current value of the route metric in the routing table.

  To provide a common tracking interface for different clients, route metrics are scaled in the range 0 to 255, where 0 is connected and 255 is inaccessible. The scaled metric value communicated to a client always considers a lower value to have priority over a higher value. The resulting scaled value is compared against the configured threshold values to determine the state of a tracked route as follows:

  - If the scaled metric for a route entry is less than or equal to the UP threshold, the state of a route is UP.
  - If the scaled metric for a route is greater than or equal to the DOWN threshold or the route is not entered in the routing table, the state of a route is DOWN.

  The UP and DOWN thresholds are user-configurable for each tracked route. The default UP threshold is 254; the default DOWN threshold is 255. The notification of a change in the state of a tracked object is sent when a metric value crosses a configured threshold.

The tracking process uses a protocol-specific resolution value to convert the actual metric in the routing table to a scaled metric in the range 0 to 255. The resolution value is user-configurable and calculates the scaled metric by dividing a route's cost by the resolution value set for the route type:

- For ISIS, you can set the resolution in the range 1 to 1000, where the default is 10.
- For OSPF, you can set the resolution in the range 1 to 1592, where the default is 1.
- The resolution value used to map static routes is not configurable. By default, FTOS assigns a metric of 0 to static routes.
- The resolution value used to map RIP routes is not configurable. The RIP hop-count is automatically multiplied by 16 to scale it. For example, a RIP metric of 16 (unreachable) scales to 256, which considers a route to be DOWN.

Tracking Route Reachability

To configure object tracking on the reachability of an IPv4 or IPv6 route, use the following commands. To remove object tracking, enter the **no track object-id** command.

Step	Task	Command Syntax	Command Mode
1	Configure object tracking on the reachability of an IPv4 or IPv6 route.	**track** *object-id* {**ip route** *ip-address/prefix-len* | **ipv6 route** *ipv6-address/prefix-len*} **reachability** [**vrf** *vrf-name*] Valid object IDs are from 1 to 65535. Enter an IPv4 address in dotted decimal format; valid IPv4 prefix lengths are from / 0 to /32. Enter an IPv6 address in X:X:X:X::X format; valid IPv6 prefix lengths are from / 0 to /128. (Optional) **E-Series only**: For an IPv4 route, you can enter a VRF name to specify the virtual routing table to which the tracked route belongs.	CONFIGURATION
2	(Optional) Configure the time delay used before communicating a change in the status of a tracked route.	**delay** {[**up** *seconds*] [**down** *seconds*]} Valid delay times are from 0 to 180 seconds. Default: 0.	OBJECT TRACKING
3	(Optional) Identify the tracked object with a text description.	**description** *text* The text string can be up to 80 characters.	OBJECT TRACKING
4	(Optional) Display the tracking configuration and the tracked object's status.	**show track** *object-id*	EXEC Privilege

Figure 31-5. Command Example: track ip route reachability

```
FTOS(conf)#track 104 ip route 10.0.0.0/8 reachability
FTOS(conf-track-104)#delay up 20 down 10
FTOS(conf-track-104)#end
FTOS#show track 104

Track 104
  IP route 10.0.0.0/8 reachability
  Reachability is Down (route not in route table)
   2 changes, last change 00:02:49
  Tracked by:

FTOS#configure
FTOS(conf)#track 4 ip route 3.1.1.0/24 reachability vrf vrf1
```

Figure 31-6. Command Example: track ipv6 route reachability

```
FTOS(conf)#track 105 ipv6 route 1234::/64 reachability
FTOS(conf-track-105)#delay down 5
FTOS(conf-track-105)#description Headquarters
FTOS(conf-track-105)#end
FTOS#show track 105

Track 105
  IPv6 route 1234::/64 reachability
  Description: Headquarters
  Reachability is Down (route not in route table)
   2 changes, last change 00:03:03
  Tracked by:
```

Tracking a Metric Threshold

To configure object tracking on the metric threshold of an IPv4 or IPv6 route, use the following commands. To remove object tracking, enter the **no track object-id** command.

Step	Task	Command Syntax	Command Mode
1	(Optional) Reconfigure the default resolution value used by the specified protocol to scale the metric for IPv4 or IPv6 routes.	**track resolution {ip route** | **ipv6 route} {isis** *resolution-value* **| ospf** *resolution-value***}** Range of resolution values: ISIS routes - 1 to 1000. Default: 1. OSPF routes - 1 to 1592. Default: 1.	CONFIGURATION
2	Configure object tracking on the metric of an IPv4 or IPv6 route.	**track** *object-id* {**ip route** *ip-address/prefix-len* | **ipv6 route** *ipv6-address/prefix-len*} **metric threshold** [**vrf** *vrf-name*] Valid object IDs are from 1 to 65535. Enter an IPv4 address in dotted decimal format.Valid IPv4 prefix lengths are from /0 to /32. Enter an IPv6 address in X:X:X:X::X format. Valid IPv6 prefix lengths are from /0 to /128. (Optional) **E-Series only**: For an IPv4 route, you can enter a VRF name.	CONFIGURATION

Step	Task	Command Syntax	Command Mode
3	(Optional) Configure the time delay used before communicating a change in the UP and/or DOWN status of a tracked route.	**delay** {[**up** *seconds*] [**down** *seconds*]} Valid delay times are from 0 to 180 seconds. Default: 0.	OBJECT TRACKING
4	(Optional) Identify the tracked object with a text description.	**description** *text* The text string can be up to 80 characters.	OBJECT TRACKING
5	(Optional) Configure the metric threshold for the UP and/or DOWN routing status to be tracked for the specified route.	**threshold metric** {[**up** *number*] [**down** *number*]} Default UP threshold: 254. The routing state is UP if the scaled route metric is less than or equal to the UP threshold. Default DOWN threshold: 255. The routing state is DOWN if the scaled route metric is greater than or equal to the DOWN threshold.	OBJECT TRACKING
6	(Optional) Display the tracking configuration.	**show track** *object-id*	EXEC Privilege

Figure 31-7. Command Example: track ip route metric threshold

```
FTOS(conf)#track 6 ip route 2.1.1.0/24 metric threshold
FTOS(conf-track-6)#delay down 20
FTOS(conf-track-6)#delay up 20
FTOS(conf-track-6)#description track ip route  metric
FTOS(conf-track-6)#threshold metric down 40
FTOS(conf-track-6)#threshold metric up 40
FTOS(conf-track-6)#exit
FTOS(conf)#track 10 ip route 3.1.1.0/24 metric threshold vrf vrf1
```

Figure 31-8. Command Example: track ipv6 route metric threshold

```
FTOS(conf)#track 8 ipv6 route 2::/64 metric threshold
FTOS(conf-track-8)#threshold metric up 30
FTOS(conf-track-8)#threshold metric down 40
```

Displaying Tracked Objects

You can display the currently configured objects used to track Layer 2 and Layer 3 interfaces, and IPv4 and IPv6 routes, by entering the following **show** commands:

- **show track** [*object-id* [**brief**] | **interface** [**brief**] [**vrf** *vrf-name*] | **ip route** [**brief**] [**vrf** *vrf-name*] | **resolution** | **vrf** *vrf-name* [**brief**] | **brief**]

  Use the **show track** command to display the configuration and status of currently tracked Layer 2 or Layer 3 interfaces, IPv4 or IPv6 routes, or a VRF instance. You can also display the currently configured per-protocol resolution values used to scale route metrics when tracking metric thresholds.

Figure 31-9. Command Example: show track

```
FTOS#show track

Track 1
  IP route 23.0.0.0/8 reachability
  Reachability is Down (route not in route table)
   2 changes, last change 00:16:08
  Tracked by:

Track 2
  IPv6 route 2040::/64 metric threshold
  Metric threshold is Up (STATIC/0/0)
   5 changes, last change 00:02:16
  Metric threshold down 255 up 254
  First-hop interface is GigabitEthernet 13/2
  Tracked by:
    VRRP GigabitEthernet 7/30 IPv6 VRID 1

Track 3
  IPv6 route 2050::/64 reachability
  Reachability is Up (STATIC)
   5 changes, last change 00:02:16
  First-hop interface is GigabitEthernet 13/2
  Tracked by:
    VRRP GigabitEthernet 7/30 IPv6 VRID 1

Track 4
  Interface GigabitEthernet 13/4 ip routing
  IP routing is Up
   3 changes, last change 00:03:30
  Tracked by:

Track 5
  IP route 192.168.0.0/24 reachability, Vrf: red
  Reachability is Up (CONNECTED)
   3 changes, last change 00:02:55
  First-hop interface is GigabitEthernet 13/4
  Tracked by:
```

Figure 31-10. Command Example: show track brief

```
Router# show track brief

ResId  Resource                    Parameter            State   LastChange
1      IP route reachability       10.16.0.0/16         Up      00:01:08
2      Interface line-protocol     Ethernet0/2          Down    00:05:00
3      Interface ip routing        VLAN100              Up      01:10:05
```

Figure 31-11. Command Example: show track resolution

```
FTOS#show track resolution

IP Route Resolution
  ISIS          1
  OSPF          1

IPv6 Route Resolution
  ISIS          1
  OSPF          1
```

Figure 31-12. Command Example: show track vrf

```
FTOS#show track vrf red

Track 5
  IP route 192.168.0.0/24 reachability, Vrf: red
  Reachability is Up (CONNECTED)
   3 changes, last change 00:02:39
  First-hop interface is GigabitEthernet 13/4
  Tracked by:
```

- **show running-config track** [*object-id*]

  Use the **show running-config track** command to display the tracking configuration of a specified object or all objects that are currently configured on the router.

Figure 31-13. Command Example: show running-config track

```
FTOS#show running-config track

track 1 ip route 23.0.0.0/8 reachability

track 2 ipv6 route 2040::/64 metric threshold
delay down 3
delay up 5
threshold metric up 200

track 3 ipv6 route 2050::/64 reachability

track 4 interface GigabitEthernet 13/4 ip routing

track 5 ip route 192.168.0.0/24 reachability vrf red

track resolution ip route isis 20
track resolution ip route ospf 10
```

32

Open Shortest Path First (OSPFv2 and OSPFv3)

Open Shortest Path First version 2 (OSPF for IPv4) is supported on platforms C E S

Open Shortest Path First version 3 (OSPF for IPv6) is supported on platforms C E

OSPF for IPv4 is supported on the E-Series ExaScale platform with FTOS 8.1.1.0; OSPF for IPv6 is supported on E-Series ExaScale with FTOS version 8.2.1.0 and later.

This chapter is intended to provide a general description of OSPFv2 (OSPF for IPv4) and OSPFv3 (OSPF for IPv6) as supported in the Force 10 Operating System (FTOS). It is not intended to provide a complete

Note: The fundamental mechanisms of OSPF (flooding, DR election, area support, SPF calculations, etc.) are the same between OSPFv2 and OSPFv3. Where there are differences between the two versions, they are identified and clarified. Except where identified, the information in this chapter applies to both protocol versions.

This chapter includes the following topics:

OSPF protocol standards are listed in Appendix , , on page 1239.

Protocol Overview

Open Shortest Path First (OSPF) routing is a link-state routing protocol that calls for the sending of Link-State Advertisements (LSAs) to all other routers within the same Autonomous System (AS) Areas. Information on attached interfaces, metrics used, and other variables is included in OSPF LSAs. As OSPF routers accumulate link-state information, they use the SPF algorithm (Shortest Path First algorithm) to calculate the shortest path to each node.

OSPF routers initially exchange HELLO messages to set up adjacencies with neighbor routers. The HELLO process is used to establish adjacencies between routers of the AS. It is not required that every router within the Autonomous System areas establish adjacencies. If two routers on the same subnet agree to become neighbors through the HELLO process, they begin to exchange network topology information in the form of Link State Advertisements (LSAs).

OSPFv3 runs on a per-link basis instead of on a per-IP-subnet basis. All neighbors on all link types are identified by Router ID (RID). In OSPFv2 neighbors on broadcast and NBMA links are identified by their interface addresses, while neighbors on other types of links are identified by RID. OSPFv3 removes this inconsistency, and all neighbors on all link types are identified by RID.

Note: OSPFv3 is not backward-compatible with OSPFv2; they can co-exist. To use OSPF with both IPv4 and IPv6, you must run both OSPFv2 and OSPFv3.

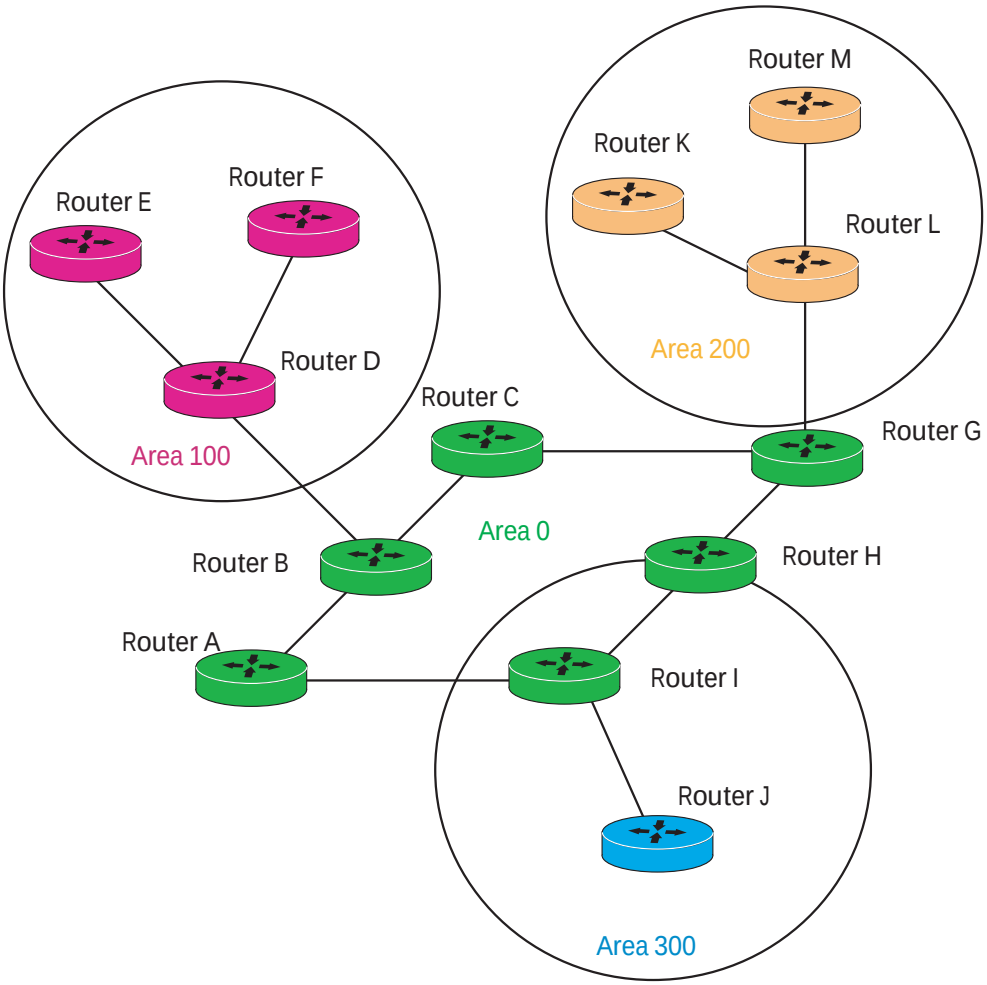
Autonomous System (AS) Areas

OSPF operate in a type of hierarchy. The largest entity within the hierarchy is the autonomous system (AS), which is a collection of networks under a common administration that share a common routing strategy. OSPF is an intra-AS (interior gateway) routing protocol, although it is capable of receiving routes from and sending routes to other ASs.

An AS can be divided into a number of areas, which are groups of contiguous networks and attached hosts. Routers with multiple interfaces can participate in multiple areas. These routers, Area Border Routers (ABRs), maintain separate databases for each area. Areas are a logical grouping of OSPF routers identified by an integer or dotted-decimal number.

Areas allow you to further organize your routers within in the AS. One or more areas are required within the AS. Areas are valuable in that they allow sub-networks to "hide" within the AS, thus minimizing the size of the routing tables on all routers. An area within the AS may not see the details of another Area's topology. AS areas are known by their area number or the router’s IP address.

Figure 32-1. Autonomous System Areas



Area Types

The **Backbone** of the network is Area 0. It is also called Area 0.0.0.0 and is the core of any Autonomous System (AS). All other areas must connect to Area 0. Areas can be defined in such a way that the backbone is not contiguous. In this case, backbone connectivity must be restored through virtual links. Virtual links are configured between any backbone routers that share a link to a non-backbone area and function as if they were direct links.

An OSPF backbone is responsible for distributing routing information between areas. It consists of all Area Border Routers, networks not wholly contained in any area, and their attached routers.

The Backbone is the only area with an default area number. All other areas can have their Area ID assigned in the configuration.

Figure 32-1 shows Routers A, B, C, G, H, and I are the Backbone.

A **Stub Area** (SA) does not receive external route information, except for the default route. These areas do receive information from inter-area (IA) routes. Note that all routers within an assigned Stub area must be configured as stubby, and no generate LSAs that do not apply. For example, a Type 5 LSA is intended for external areas and the Stubby area routers may not generate external LSAs. Stubby areas cannot be traversed by a virtual link.

A **Not-So-Stubby** Area (NSSA) can import AS external route information and send it to the Backbone. It cannot received external AS information from the Backbone or other areas. It can be traversed by a virtual link.

Totally Stubby Areas are referred to as No Summary areas in FTOS.

Networks and Neighbors

As a link-state protocol, OSPF sends routing information to other OSPF routers concerning the state of the links between them. The state (up or down) of those links is important.

Routers that share a link become neighbors on that segment. OSPF uses the hello protocol as a neighbor discovery and keep alive mechanism. After two routers are neighbors, they may proceed to exchange and synchronize their databases, which creates an adjacency.

Note: You can log adjacency state changes for OSPFv2 and v3 with the command **log-adjacency-changes** from ROUTER OSPF mode.

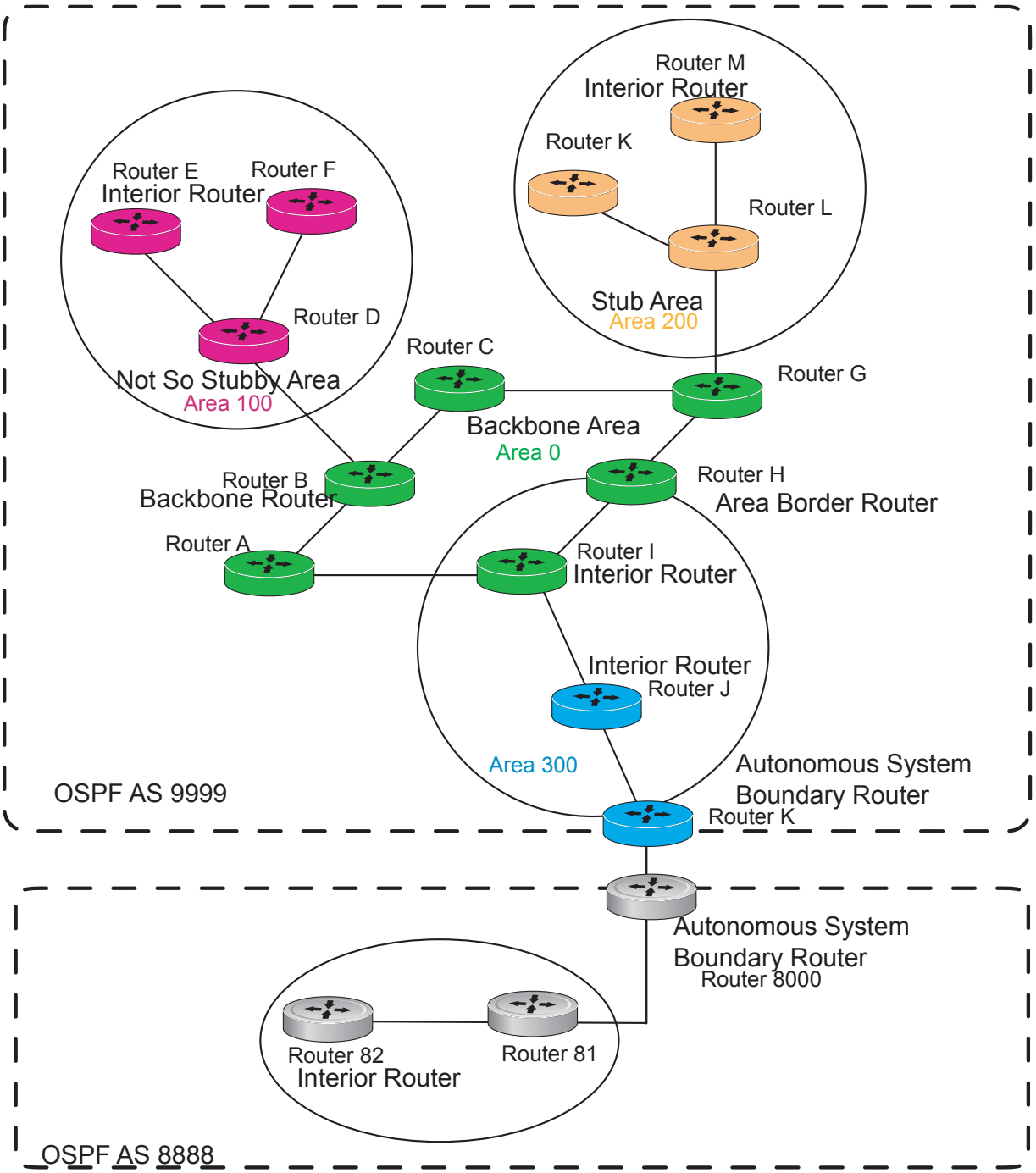
Router Types

Router types are attributes of the OSPF process. A given physical router may be a part of one or more OSPF processes. For example, a router connected to more than one area, receiving routing from a BGP process connected to another AS acts as both an Area Border Router and an Autonomous System Router.

Each router has a unique ID, written in decimal format (A.B.C.D). The router ID does not have to be associated with a valid IP address. However, Force 10 recommends that the router ID and the router's IP address reflect each other, to make troubleshooting easier.

Figure 32-2gives some examples of the different router designations.

Figure 32-2. OSPF Routing Examples



Backbone Router (BR)

A Backbone Router (BR) is part of the OSPF Backbone, Area 0. This includes all Area Border Routers (ABRs). It can also include any routers that connect only to the Backbone and another ABR, but are only part of Area 0, such as Router I in Figure 32-2 above.

Area Border Router (ABR)

Within an AS, an Area Border (ABR) connects one or more areas to the Backbone. The ABR keeps a copy of the link-state database for every area it connects to, so it may keep multiple copies of the link state database. An Area Border Router (ABR) takes information it has learned on one of its attached areas and can summarize it before sending it out on other areas it is connected to.

An ABR can connect to many areas in an AS, and is considered a member of each area it connects to.

Autonomous System Border Router (ASBR)

The Autonomous System Border Area Router (ASBR) connects to more than one AS, and exchanges information with the routers in other ASs. Generally the ASBR connects to a non-Interior Gate Protocol (IGP) such as BGP or uses static routes.

Internal Router (IR)

The Internal Router (IR) has adjacencies with ONLY routers in the same area, as Router E, M and I are shown in Figure 32-2.

Designated and Backup Designated Routers

OSPF elects a Designated Router and a Backup Designated router. Among other things, the designated router is responsible for generating LSAs for the entire multiaccess network. Designated routers allow a reduction in network traffic and in the size of the topological database.

- The Designated Router (DR) maintains a complete topology table of the network and sends the updates to the other routers via multicast. All routers in an area form a slave/master relationship with the DR. Every time a router sends an update, it sends it to the Designated Router (DR) and Backup Designated Router (BDR). The DR sends the update out to all other routers in the area.
- The Backup Designated Router (BDR) is the router that takes over if the DR fails.

Each router exchanges information with the DR and BDR. The DR and BDR relay the information to the other routers. On broadcast network segments the number of OSPF packets is further reduced by the DR and BDR sending such OSPF updates to a multicast IP address that all OSPF routers on the network segment are listening on.

These router designations are not the same ad the router IDs discussed earlier. The Designated and Backup Designated Routers are configurable in FTOS. If no DR or BDR is defined in FTOS, the system assigns them. OSPF looks at the priority of the routers on the segment to determine which routers are the DR and BDR. The router with the highest priority is elected the DR. If there is a tie, then the router with the higher Router ID takes precedence. After the DR is elected, the BDR is elected the same way. A router with a router priority set to zero is cannot become the DR or BDR.

Link-State Advertisements (LSAs)

A Link-State Advertisement (LSA) communicates the router's local routing topology to all other local routers in the same area.

- OSPFv3 can treat LSAs as having link-local flooding scope, or store and flood them as if they are understood, while ignoring them in their own SPF algorithms.
- OSPFv2 always discards unknown LSA types.

The LSA types supported by Force 10 are defined as follows:

- Type 1 - Router LSA
  - The router lists links to other routers or networks in the same area. Type 1 LSAs are flooded across their own area only. The Link-State ID of the Type 1 LSA is the originating router ID.
- Type 2 - Network LSA
  - The Designated Router (DR) in an area lists which routers are joined together within the area. Type 2 LSAs are flooded across their own area only. The Link-State ID of the Type 2 LSA is the IP interface address of the DR.
- Type 3 - Summary LSA (OSPFv2), Inter-Area-Prefix LSA (OSPFv3)
  - An Area Border Router (ABR) takes information it has learned on one of its attached areas and can summarize it before sending it out on other areas it is connected to. The Link-State ID of the Type 3 LSA is the destination network number.
- Type 4 - AS Border Router Summary LSA (OSPFv2), Inter-Area-Router LSA (OSPFv3)
  - In some cases, Type 5 External LSAs are flooded to areas where the detailed next-hop information may not be available. An Area Border Router will (ABR) flood the information for the router (i.e. the Autonomous System Border Router (ASBR) where the Type 5 advertisement originated. The Link-State ID for Type 4 LSAs is the router ID of the described ASBR.
- Type 5 - External LSA
  - These LSAs contain information imported into OSPF from other routing processes. They are flooded to all areas, except stub areas. The Link-State ID of the Type 5 LSA is the external network number.
- Type 7
  - Routers in a Not-So-Stubby-Area (NSSA) do not receive external LSAs from Area Border Routers (ABRs), but are allowed to send external routing information for redistribution. They use Type 7 LSAs to tell the ABRs about these external routes, which the Area Border Router then translates to Type 5 external LSAs and floods as normal to the rest of the OSPF network.
- Type 8 - Link LSA (OSPFv3)
  - This LSA carries the IPv6 address information of the local links.
- Type 9 - Link Local LSA (OSPFv2), Intra-Area-Prefix LSA (OSPFv3)
  - For OSPFv2, this is a link-local “opaque” LSA as defined by RFC2370.
  - For OSPFv3, this LSA carries the IPv6 prefixes of the router and network links.
- Type 11 - Grace LSA (OSPFv3)
  - For OSPFv3 only, this LSA is a link-local “opaque” LSA sent by a restarting OSPFv3 router during a graceful restart.

For all LSA types, there are 20-byte LSA headers. One of the fields of the LSA header is the Link-State ID.

Each router link is defined as one of four types: type 1, 2, 3, or 4. The LSA includes a link ID field that identifies, by the network number and mask, the object this link connects to.

Depending on the type, the link ID has different meanings.

- 1: point-to-point connection to another router neighboring router
- 2: connection to a transit network IP address of Designated Router
- 3: connection to a stub network IP network/subnet number
- 4: virtual link neighboring router ID

Virtual Links

In the case in which an area cannot be directly connected to Area 0, you must configure a virtual link between that area and Area 0. The two endpoints of a virtual link are ABRs, and the virtual link must be configured in both routers. The common non-backbone area to which the two routers belong is called a transit area. A virtual link specifies the transit area and the router ID of the other virtual endpoint (the other ABR).

A Virtual Link cannot be configured through a Stub Area or NSSA.

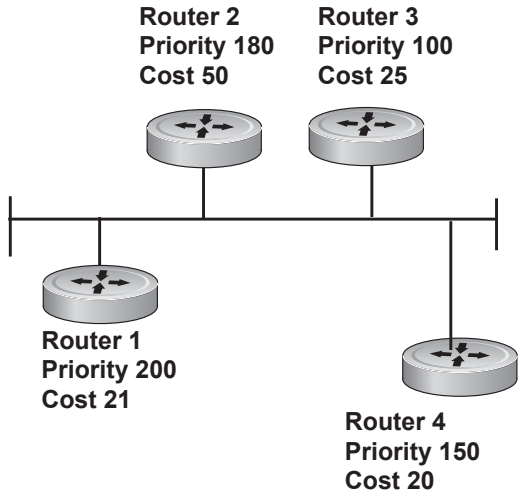
Router Priority and Cost

Router priority and cost is the method the system uses to “rate” the routers. For example, if not assigned, the system will select the router with the highest priority as the DR. The second highest priority is the BDR.

Priority is a numbered rating 0-255. The higher the number, the higher the priority.

Cost is a numbered rating 1-65535. The higher the number, the greater the cost. The cost assigned reflects the cost should the router fail. When a router fails and the cost is assessed, a new priority number results.

Figure 32-3. Priority and Costs Example



Router 1 selected by the system as DR.
Router 2 selected by the system as BDR.

If R1 fails, the system subtracts 21 from R1's priority number. R1's new priority is 179.

R2 as both the selected BDR and the now-highest priority, becomes the DR.

If R3 fails, the system subtracts 50 from its priority. R2's new priority is 130.

R4 is now the highest priority and becomes the DR.

Implementing OSPF with FTOS

FTOS supports up to 10,000 OSPF routes. Within that 10,000 up to 8,000 routes can be designated as external and up to 2,000 designated as inter/intra area routes.

FTOS version 7.8.1.0 and later support multiple OSPF processes (OSPF MP).

Prior to 7.8.1.0, FTOS supports 1 OSPFv2 and 1 OSPFv3 process ID per system. Recall that OSPFv2 and OSPFv3 can coexist but must be configured individually.

FTOS supports Stub areas, Totally Stub (No Summary) and Not So Stubby Areas (NSSAs) and supports the following LSAs, as discussed earlier in this document.

- Router (type 1)
- Network (type 2)
- Network Summary (type 3)
- AS Boundary (type 4)
- AS External (type 5)

- NSSA External (type 7)
- Opaque Link-local (type 9)

Graceful Restart

Graceful Restart for OSPFv2 is supported on C E and S platforms in Helper and Restart modes.

Graceful Restart for OSPFv3 is supported only on E T platforms in Helper and Restart modes.

When a router goes down without a Graceful Restart, there is a possibility for loss of access to parts of the network due to ongoing network topology changes. Additionally, LSA flooding and reconvergence can cause substantial delays. It is, therefore, desirable that the network maintain a stable topology if it is possible for data flow to continue uninterrupted.

OSPF Graceful Restart recognizes the fact that in a modern router, the control plane and data plane functionality are separate, restarting the control plane functionality (such as the failover of the active RPM to the backup in a redundant configuration), does not necessarily have to interrupt the forwarding of data packets. This behavior is supported because the forwarding tables previously computed by an active RPM have been downloaded into the Forwarding Information Base on the line cards (the data plane), and are still resident. For packets that have existing FIB/CAM entries, forwarding between ingress and egress ports/VLANs etc., can continue uninterrupted while the control plane OSPF process comes back to full functionality and rebuilds its routing tables.

When a router is attempting to restart gracefully, it originates the following link-local Grace LSAs to notify its helper neighbors that the restart process is beginning:

- An OSPFv2 router sends Type 9 LSAs.
- An OSPFv3 router sends Type 11 LSAs.

Type 9 and 11 LSAs include a grace period, which is the time period an OSPF router advertises to adjacent neighbor routers as the time to wait for it to return to full control plane functionality. During the grace period, neighbor OSPFv2 /v3 interfaces save the LSAs from the restarting OSPF interface. Helper neighbor routers continue to announce the restarting router as fully adjacent, as long as the network topology remains unchanged. When the restarting router completes its restart, it flushes the Type 9 and 11 LSAs, thereby notifying its neighbors that the restart is complete. This should happen before the grace period expires.

Dell Force10 routers support the following OSPF graceful restart functionality:

- Restarting role in which a router is enabled to perform its own graceful restart.
- Helper role in which the router's graceful restart function is to help a restarting neighbor router in its graceful restarts.
- Helper-reject role in which OSPF does not participate in the graceful restart of a neighbor.

  OSPFv2 supports “helper-only” and “restarting-only” roles. By default, both helper and restarting roles are enabled. OSPFv2 supports the helper-reject role globally on a router.

  OSPFv3 supports “helper-only” and “restarting-only” roles. The “helper-only” role is enabled by default. To enable the restarting role in addition to the “helper-only” role, you must configure a grace

period. You reconfigure OSPFv3 graceful restart to a “restarting-only” role when you enable the helper-reject role on an interface. OSPFv3 supports the helper-reject role on a per-interface basis.

Configuring helper-reject role on an OSPFv2 router or OSPFv3 interface enables the restarting-only role globally on the router or locally on the interface. In a helper-reject role, OSPF does not participate in the graceful restart of an adjacent OSPFv2/v3 router.

If multiple OSPF interfaces provide communication between two routers, after you configure helper-reject on one interface, all other interfaces between the two routers behave as if they are in the help-reject role.

- OSPFv2 and OSPFv3 support planned-only and/or unplanned-only restarts. The default is support for both planned and unplanned restarts.

  A planned restart occurs when you enter the **redundancy force-failover rpm** command to force the primary RPM to switch to the backup RPM. During a planned restart, OSPF sends out a Grace LSA before the system switches over to the backup RPM.

  An unplanned restart occurs when an unplanned event causes the active RPM to switch to the backup RPM, such as when an active process crashes, the active RPM is removed, or a power failure happens. During an unplanned restart, OSPF sends out a Grace LSA when the backup RPM comes online.

To display the configuration values for OSPF graceful restart, enter the following commands:

- For OSPFv2: **show run ospf**
- For OSPFv3: **show run ospf** and **show ipv6 ospf database database-summary**

Fast Convergence (OSPFv2, IPv4 only)

Fast Convergence allows you to define the speeds at which LSAs are originated and accepted, and reduce OSPFv2 end-to-end convergence time. FTOS enables you to accept and originate LSAa as soon as they are available to speed up route information propagation.

Note that the faster the convergence, the more frequent the route calculations and updates. This will impact CPU utilization and may impact adjacency stability in larger topologies.

Multi-Process OSPF (OSPFv2, IPv4 only)

Multi-Process OSPF is supported on platforms C E and S with FTOS version 7.8.1.0 and later, and is supported on OSPFv2 with IPv4 only.

Multi-Process OSPF allows multiple OSPFv2 processes on a single router. Multiple OSPFv2 processes allow for isolating routing domains, supporting multiple route policies and priorities in different domains, and creating smaller domains for easier management.

- The E-Series supports up to 28 OSPFv2 processes.
- The C-Series supports up to 6 OSPFv2 processes.
- The S-Series supports up to 3 OSPFv2 processes.

Each OSPFv2 process has a unique process ID and must have an associated Router ID. There must be an equal number of interfaces must be in Layer-3 mode for the number of processes created. For example, if 5 OSPFv2 processes are created on a system, there must be at least 5 interfaces assigned in Layer-3 mode.

Each OSPFv2 process is independent. If one process loses adjancency, the other processes continue to function/

Processing SNMP and Sending SNMP Traps

Though there are may be several OSPFv2 processes, only one process can process SNMP requests and send SNMP traps. The **mib-binding** command identifies one of the OSPVFv2 processes as the process responsible for SNMP management. If the **mib-binding** command is not specified, the first OSPFv2 process created manages the SNMP processes and traps.

RFC-2328 Compliant OSPF Flooding

In OSPF, flooding is the most resource-consuming task. The flooding algorithm described in RFC 2328 requires that OSPF flood LSAs on all interfaces, as governed by LSA's flooding scope. (Refer to Section 13 of the RFC.) When multiple direct links connect two routers, the RFC 2328 flooding algorithm generates significant redundant information across all links.

By default, FTOS implements an enhanced flooding procedure which dynamically and intelligently detects when to optimize flooding. Wherever possible, the OSPF task attempts to reduce flooding overhead by selectively flooding on a subset of the interfaces between two routers.

If RFC 2328 flooding behavior is required, enable it by using the command **flood-2328** in ROUTER OSPF mode. When enabled, this command configures FTOS to flood LSAs on all interfaces.

Confirm RFC 2328 flooding behavior by using the command **debug ip ospf packet** and look for output similar to the following:

Figure 32-4. Enabling RFC-2328 Compliant OSPF Flooding

```
00:10:41 : OSPF(1000:00):
Rcv. v:2 t:5(LSAck) l:64 Acks 2 rid:2.2.2.2          <-- Printed only for ACK packets
        aid:1500 chk:0xdbee aut:0 auk: keyid:0 from:Vl 1000
            LSType:Type-5 AS External id:160.1.1.0 adv:6.1.0.0 seq:0x8000000c
            LSType:Type-5 AS External id:160.1.2.0 adv:6.1.0.0 seq:0x8000000c
00:10:41 : OSPF(1000:00):
Rcv. v:2 t:5(LSAck) l:64 Acks 2 rid:2.2.2.2
        aid:1500 chk:0xdbee aut:0 auk: keyid:0 from:Vl 100
            LSType:Type-5 AS External id:160.1.1.0 adv:6.1.0.0 seq:0x8000000c
            LSType:Type-5 AS External id:160.1.2.0 adv:6.1.0.0 seq:0x8000000c
00:10:41 : OSPF(1000:00):
Rcv. v:2 t:4(LSUpd) l:100 rid:6.1.0.0                <-- No change in update packets
        aid:0 chk:0xccbd aut:0 auk: keyid:0 from:Gi 10/21
            Number of LSA:2
            LSType:Type-5 AS External(5) Age:1 Seq:0x8000000c id:170.1.1.0 Adv:6.1.0.0
                Netmask:255.255.255.0 fwd:0.0.0.0 E2, tos:0 metric:0
            LSType:Type-5 AS External(5) Age:1 Seq:0x8000000c id:170.1.2.0 Adv:6.1.0.0
                Netmask:255.255.255.0 fwd:0.0.0.0 E2, tos:0 metric:0
```

In FTOS Version, 7.5.1.0 use **show ip ospf** to confirm that RFC-2328 compliant OSPF flooding is enabled, as shown below.

Figure 32-5. Enabling RFC-2328 Compliant OSPF Flooding

```
FTOS#show ip ospf
Routing Process ospf 1 with ID 2.2.2.2
Supports only single TOS (TOS0) routes
It is an Autonomous System Boundary Router
It is Flooding according to RFC 2328
SPF schedule delay 5 secs, Hold time between two SPFs 10 secs
Number of area in this router is 1, normal 0 stub 0 nssa 1
--More--
```

OSPF ACK Packing

The OSPF ACK Packing feature bundles multiple LS acknowledgements in a single packet, significantly reducing the number of ACK packets transmitted when the number of LSAs increases. This feature also enhances network utilization and reduces the number of small ACK packets sent to a neighboring router. OSPF ACK packing is enabled by default, and non-configurable.

OSPF Adjacency with Cisco Routers

To establish an OSPF adjacency between Force 10 and Cisco routers, the hello interval and dead interval must be the same on both routers. In FTOS the OSPF dead interval value is, by default, set to 40 seconds, and is independent of the OSPF hello interval. Configuring a hello interval does not change the dead interval in FTOS. In contrast, the OSPF dead interval on a Cisco router is, by default, four times as long as the hello interval. Changing the hello interval on the Cisco router automatically changes the dead interval as well.

To ensure equal intervals between the routers, manually set the dead interval of the Dell Force10 router to match the Cisco configuration. Use the command "**ip ospf dead-interval <x>**" in interface mode:

Figure 32-6. Command Example: ip ospf intervals

```
FTOS(conf)#int gi 2/2
FTOS(conf-if-gi-2/2)#ip ospf hello-interval 20
FTOS(conf-if-gi-2/2)#ip ospf dead-interval 80

FTOS(conf-if-gi-2/2)#
```

Dead Interval Set at 4x Hello Interval

Figure 32-7. OSPF Configuration with intervals set

```
FTOS (conf-if-gi-2/2)#ip ospf dead-interval 20
FTOS (conf-if-gi-2/2)#do show ip os int gi1/3
GigabitEthernet 2/2 is up, line protocol is up
 Internet Address 20.0.0.1/24, Area 0
 Process ID 10, Router ID 1.1.1.2, Network Type BROADCAST, Cost: 1
 Transmit Delay is 1 sec, State DR, Priority 1
 Designated Router (ID) 1.1.1.2, Interface address 30.0.0.1
 Backup Designated Router (ID) 1.1.1.1, Interface address 30.0.0.2
 Timer intervals configured, Hello 20, Dead 80, Wait 20, Retransmit 5
 Hello due in 00:00:04
 Neighbor Count is 1, Adjacent neighbor count is 1
 Adjacent with neighbor 1.1.1.1 (Backup Designated Router)
FTOS (conf-if-gi-2/2)#
```

Dead Interval Set at 4x Hello Interval

For more information regarding this functionality or for assistance, go to www.force10networks.com/support.

Configuration Requirements

The interfaces must be in Layer-3 mode (assigned an IP address) and enabled so that they can send and receive traffic. The OSPF process must know about these interfaces. To make the OSPF process aware of these interfaces, they must be assigned to OSPF areas.

OSPF must be configured GLOBALLY on the system in CONFIGURATION mode.

OSPF features and functions are assigned to each router using the CONFIG-INTERFACE commands for each interface.

Note: By default, OSPF is disabled

Configuration Task List for OSPFv2 (OSPF for IPv4)

Open Shortest Path First version 2 (OSPF for IPv4) is supported on platforms C E S

1. Configure a physical interface. Assign an IP address, physical or loopback, to the interface to enable Layer 3 routing.

2. Enable OSPF globally. Assign network area and neighbors.
3. Add interfaces or configure other attributes.

The following configuration steps include two mandatory steps and several optional ones:

- Enable OSPFv2 (mandatory)
- Enable Multi-Process OSPF
- Assign an OSPFv2 area (mandatory)
- Enable OSPFv2 on interfaces
- Configure stub areas
- Configure OSPF Stub-Router Advertisement
- Enable passive interfaces
- Enable fast-convergence
- Change OSPFv2 parameters on interfaces
- Enable OSPFv2 authentication
- Enable OSPFv2 graceful restart
- Configure virtual links
- Redistribute routes
- Troubleshooting OSPFv2

For a complete listing of all commands related to OSPFv2, refer to the OSPF section in the *FTOS Command Line Interface* document.

Enable OSPFv2

Assign an IP address to an interface (physical or Loopback) to enable Layer 3 routing. By default OSPF, like all routing protocols, is disabled.

You *must* configure at least one interface for Layer 3 before enabling OSPFv2 globally.

If implementing, Multi-Process OSPF, you must create an equal number of Layer 3 enabled interfaces and OSPF Process IDs. For example, if you create 4 OSPFv2 process IDs, you must have 4 interfaces with Layer 3 enabled.

Use these commands on one of the interfaces to enable OSPFv2 routing.

Step	Command Syntax	Command Mode	Usage
1	**ip address** *ip-address mask*	CONFIG-INTERFACE	Assign an IP address to an interface. Format: A.B.C.D/M If using a Loopback interface, refer to Loopback Interfaces on page 427.
2	**no shutdown**	CONFIG-INTERFACE	Enable the interface.

Return to CONFIGURATION mode to enable the OSPF process. The OSPF Process ID is the identifying number assigned to the OSPF process, and the Router ID is the IP address associated with the OSPF process. .

Command Syntax	Command Mode	Usage
router ospf *process-id [vrf {vrf name}]*	CONFIGURATION	Enable the OSPFv2 process globally. Range: 0-65535 *vrf name*: Enter the VRF key word and instance name to tie the OSPF instance to the VRF. All network commands under this OSPF instance are subsequently tied to the VRF instance.

Once the OSPF process and the VRF are tied together, the OSPF Process ID cannot be used again in the system.

If you try to enter an OSPF Process ID, or if you try to enable more OSPF processes than available Layer 3 interfaces, prior to assigning an IP address to an interface and setting the no shutdown command, you will see the following message.

Message 1

```
C300(conf)#router ospf 1
% Error: No router ID available.
```

In CONFIGURATION ROUTER OSPF mode, assign the Router ID. The Router ID is not required to be the router's IP address. Dell Force10 recommends using the IP address as the Router ID for easier management and troubleshooting.

Command Syntax	Command Mode	Usage
router-id *ip address*	CONFIG-ROUTER-OSPF-id	Assign the Router ID for the OSPFv2 process. IP Address: A.B.C.D

Use the **no router ospf** *process-id* command syntax in the CONFIGURATION mode to disable OSPF.

Use the **clear ip ospf** *process-id* command syntax in EXEC Privilege mode to reset the OSPFv2 process.

Use the **show ip ospf** *process-id* command in EXEC mode (Figure 408) to view the current OSPFv2 status.

Figure 32-8. Command Example: show ip ospf *process-id*

```
FTOS#show ip ospf 55555
Routing Process ospf 55555 with ID 10.10.10.10
Supports only single TOS (TOS0) routes
SPF schedule delay 5 secs, Hold time between two SPFs 10 secs
Number of area in this router is 0, normal 0 stub 0 nssa 0
FTOS#
```

Enable Multi-Process OSPF

Multi-Process OSPF allows multiple OSPFv2 processes on a single router. The following list shows the number of processes supported on each platform type.

- The E-Series supports up to 30 OSPFv2 processes.
- The C-Series supports up to 6 OSPFv2 processes.
- The S-Series supports up to 4 OSPFv2 processes.

Follow the same steps as above, when configuring a single OSPF process. Repeat them as often as necessary for the desired number of processes. Once the process is created, all other configurations apply as usual,

Step	Command Syntax	Command Mode	Usage
1	**ip address** *ip-address mask*	CONFIG-INTERFACE	Assign an IP address to an interface. Format: A.B.C.D/M If using a Loopback interface, refer to Loopback Interfaces on page 427.
2	**no shutdown**	CONFIG-INTERFACE	Enable the interface.

Return to CONFIGURATION mode to enable the OSPF process. The OSPF Process ID is the identifying number assigned to the OSPF process, and the Router ID is the IP address associated with the OSPF process. .

Command Syntax	Command Mode	Usage
router ospf *process-id* [**vrf** *{vrf name}]*	CONFIGURATION	Enable the OSPFv2 process globally. Range: 0-65535 *vrf name*: Enter the VRF key word and instance name to tie the OSPF instance to the VRF. All network commands under this OSPF instance are subsequently tied to the VRF instance.

Once the OSPF process and the VRF are tied together, the OSPF Process ID cannot be used again in the system.

If you try to enable more OSPF processes than available Layer 3 interfaces you will see the following message.

Message 2

```
C300(conf)#router ospf 1
% Error: No router ID available.
```

In CONFIGURATION ROUTER OSPF mode, assign the Router ID. The Router ID is not required to be the router's IP address. Dell Force10 recommends using the IP address as the Router ID for easier management and troubleshooting.

Command Syntax	Command Mode	Usage
router-id *ip address*	CONFIG-ROUTER-OSPF-id	Assign the Router ID for the OSPFv2 process. IP Address: A.B.C.D

Use the **no router ospf** *process-id* command syntax in the CONFIGURATION mode to disable OSPF.

Use the **clear ip ospf** *process-id* command syntax in EXEC Privilege mode to reset the OSPFv2 process.

Assign an OSPFv2 area

After OSPFv2 is enabled, assign the interface to an OSPF area. Set up OSPF Areas and enable OSPFv2 on an interface with the **network** command.

You must have at least one AS area: Area 0. This is the Backbone Area. If your OSPF network contains more than one area, you must also configure a backbone area (Area ID 0.0.0.0). Any area besides Area 0 can have any number ID assigned to it.

The OSPFv2 process evaluates the **network** commands in the order they are configured. Assign the network address that is most explicit first to include all subnets of that address. For example, if you assign the network address 10.0.0.0 /8, you cannot assign the network address 10.1.0.0 /16 since it is already included in the first network address.

When configuring the **network** command, you must configure a network address and mask that is a superset of the IP subnet configured on the Layer-3 interface to be used for OSPFv2.

Use this command in CONFIGURATION ROUTER OSPF mode to set up each neighbor and OSPF area. The Area can be assigned by a number or with an IP interface address.

Command Syntax	Command Mode	Usage
network *ip-address mask* **area** *area-id*	CONFIG-ROUTER-OSPF-id	Enable OSPFv2 on an interface and assign an network address range to a specific OSPF area. IP Address Format: A.B.C.D/M Area ID Range: 0-65535 or A.B.C.D/M

Enable OSPFv2 on interfaces

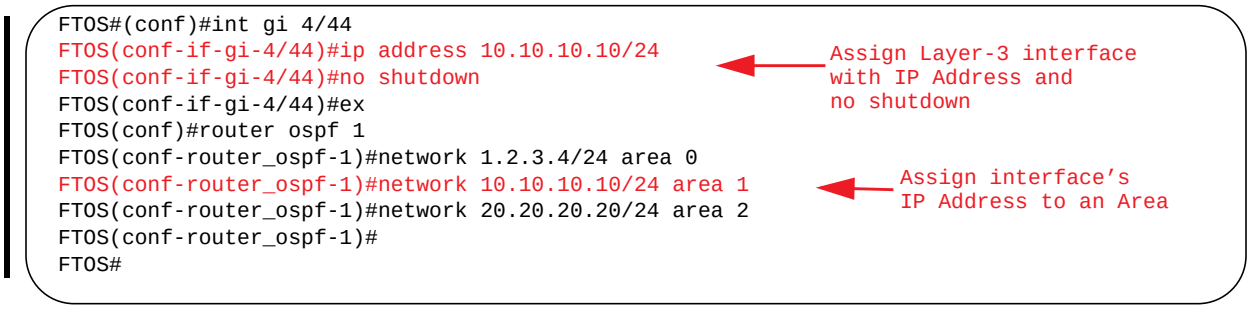
Each interface must have OSPFv2 enabled on it. It must be configured for Layer 3 protocol, and not be shutdown. OSPFv2 can also be assigned to a loopback interface as a virtual interface.

OSPF functions and features, such as MD5 Authentication, Grace Period, Authentication Wait Time, etc, are assigned on a per interface basis.

Note: If using features like MD5 Authentication, ensure all the neighboring routers are also configured for MD5.

Figure 32-9 presents an example of assigning an IP address to an interface and then assigning an OSPFv2 area that includes that Layer-3 interface's IP address.

Figure 32-9. Configuring an OSPF Area Example



Force 10 recommends that the OSPFv2 Router ID be the interface IP addresses for easier management and troubleshooting.

Use the **show config** command in CONFIGURATION ROUTER OSPF mode to view the configuration.

OSPF, by default, sends hello packets out to all physical interfaces assigned an IP address that are a subset of a network on which OSPF is enabled. Use the **show ip ospf interface** command (Figure 410) to view the interfaces currently active and the areas assigned to the interfaces.

Figure 32-10. Command Example: show ip ospf *process-id* interface

```
FTOS>show ip ospf 1 interface

GigabitEthernet 12/17 is up, line protocol is up
  Internet Address 10.2.2.1/24, Area 0.0.0.0
  Process ID 1, Router ID 11.1.2.1, Network Type BROADCAST, Cost: 1
  Transmit Delay is 1 sec, State DR, Priority 1
  Designated Router (ID) 11.1.2.1, Interface address 10.2.2.1
  Backup Designated Router (ID) 0.0.0.0, Interface address 0.0.0.0
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:04
  Neighbor Count is 0, Adjacent neighbor count is 0

GigabitEthernet 12/21 is up, line protocol is up
  Internet Address 10.2.3.1/24, Area 0.0.0.0
  Process ID 1, Router ID 11.1.2.1, Network Type BROADCAST, Cost: 1
  Transmit Delay is 1 sec, State BDR, Priority 1
  Designated Router (ID) 13.1.1.1, Interface address 10.2.3.2
  Backup Designated Router (ID) 11.1.2.1, Interface address 10.2.3.1
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:05
  Neighbor Count is 1, Adjacent neighbor count is 1
    Adjacent with neighbor 13.1.1.1 (Designated Router)
FTOS>
```

Loopback interfaces also assist in the OSPF process. OSPF will pick the highest interface address as the router-id and a loopback interface address has a higher precedence than other interface addresses.

Figure 32-11 gives an example of the **show ip ospf** *process-id interface* command with a Loopback interface.

Figure 32-11. Command Example: show ip ospf *process-id* interface

```
FTOS#show ip ospf 1 int

GigabitEthernet 13/23 is up, line protocol is up
  Internet Address 10.168.0.1/24, Area 0.0.0.1
  Process ID 1, Router ID 10.168.253.2, Network Type BROADCAST, Cost: 1
  Transmit Delay is 1 sec, State DROTHER, Priority 1
  Designated Router (ID) 10.168.253.5, Interface address 10.168.0.4
  Backup Designated Router (ID) 192.168.253.3, Interface address 10.168.0.2
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 00:00:08
  Neighbor Count is 3, Adjacent neighbor count is 2
    Adjacent with neighbor 10.168.253.5 (Designated Router)
    Adjacent with neighbor 10.168.253.3 (Backup Designated Router)

Loopback 0 is up, line protocol is up
  Internet Address 10.168.253.2/32, Area 0.0.0.1
  Process ID 1, Router ID 10.168.253.2, Network Type LOOPBACK, Cost: 1
Loopback interface is treated as a stub Host.
```

Configure stub areas

OSPF supports different types of LSAs to help reduce the amount of router processing within the areas. Type 5 LSAs are not flooded into stub areas; the Area Border Router (ABR) advertises a default route into the stub area to which it is attached. Stub area routers use the default route to reach external destinations

To ensure connectivity in your OSPFv2 network, never configure the backbone area as a stub area.

Use these commands in the following sequence, starting in EXEC Privilege mode to configure a stub area.

Step	Command Syntax	Command Mode	Usage
1	**show ip ospf** *process-id* **[vrf** *vrf name]* **database database-summary**	EXEC Privilege	Review all areas after they were configured to determine which areas are NOT receiving type 5 LSAs. *vrf name*: Show only the OSPF information tied to the VRF process.
2	**configure**	EXEC Privilege	Enter the CONFIGURATION mode.
3	**router ospf** *process-id* **[vrf** *{vrf name}]*	CONFIGURATION	Enter the ROUTER OSPF mode. Process ID is the ID assigned when configuring OSPFv2 globally (page 58). *vrf name*: Enter the VRF key word and instance name to tie the OSPF instance to the VRF. All network commands under this OSPF instance are subsequently tied to the VRF instance.
4	**area** *area-id* **stub** [**no-summary**]	CONFIG-ROUTER-OSPF-id	Configure the area as a stub area. Use the **no-summary** keywords to prevent transmission in to the area of summary ASBR LSAs. Area ID is the number or IP address assigned when creating the Area (page 60).

Use the **show ip ospf database** *process-id* **database-summary** command syntax (Figure 413) in the EXEC Privilege mode To view which LSAs are transmitted.

Figure 32-12. Command Example: show ip ospf *process-id* database database-summary

```
FTOS#show ip ospf 34 database database-summary

            OSPF Router with ID (10.1.2.100) (Process ID 34)

Area ID         Router   Network S-Net   S-ASBR  Type-7   Subtotal
2.2.2.2         1        0       0       0       0        1
3.3.3.3         1        0       0       0       0        1
FTOS#
```

To view information on areas, use the **show ip ospf** *process-id* command in the EXEC Privilege mode.

Configure OSPF Stub-Router Advertisement

Configure OSPF Stub-Router Advertisement is supported on platforms: C E

When you bring a new router onto an OSPF network, you can configure the router to function as a stub area by globally reconfiguring the OSPF link cost so that other routers do not use a path that forwards traffic destined to other networks through the new router for a specified time until the router's switching and routing functions are up and running, and the routing tables in network routers have converged.

By using the **max-metric router-lsa** command, you force the link cost of all OSPF non-stub links to the maximum link cost (65535) for a specified time. The advertisement of this maximum metric causes other routers to assign a cost to the new router that is higher than the cost of using an alternate path. Because of the high cost assigned to paths that pass through the new router, other routers will not use a path through the new router as a transit path to forward traffic to other networks.

Use the **max-metric router-lsa** command to gracefully shut down or reload a router without dropping packets destined for other networks.

Command Syntax	Command Mode	Usage
max-metric router-lsa [**on-startup** {*announce-time* | **wait-for-bgp** [*wait-time*]}]	ROUTER OSPF	**E-Series ExaScale only**: Configure the maximum cost of 65535 on a new router so that it always functions as a stub router in the network and OSPF traffic destined to other networks is not routed on paths which pass through the router. **on-startup** *announce-time* specifies the time (in seconds) following boot-up during which the maximum cost (65535) for transmitting OSPF traffic on router interfaces is announced in LSAs and the router functions as a stub router. Range: 5 to 86400 seconds. **on-startup wait-for-bgp [***wait-time***]** enables the router to announce the maximum metric in OSPF LSAs until the BGP routing table converges with updated routes. Default: 600 seconds. You can also specify the time (in seconds) that the router waits for the BGP routing table to converge before it stops advertising the maximum cost in LSAs and advertises the router's currently configured OSPF cost. Range: 5 to 86400 seconds.

Note: If you enter the **max-metric router-lsa** command without an option (**on-startup announce-time** or **on-startup wait-for-bgp** [**wait-time**]), the maximum metric of 65535 is always announced in LSAs sent by the router.

Enable passive interfaces

A passive interface is one that does not send or receive routing information. Enabling passive interface suppresses routing updates on an interface. Although the passive interface will neither send nor receive routing updates, the network on that interface will still be included in OSPF updates sent via other interfaces.

Use the following command in the ROUTER OSPF mode to suppress the interface's participation on an OSPF interface. This command stops the router from sending updates on that interface.

Command Syntax	Command Mode	Usage
passive-interface {default | interface}	CONFIG-ROUTER-OSPF-id	Specify whether all or some of the interfaces will be passive. **Default** enabled passive interfaces on ALL interfaces in the OSPF process. Entering the physical interface type, slot, and number enable passive interface on only the identified interface. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information (**e.g. passive-interface gi 2/1**). • For a port channel, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale (**e.g. passive-interface po 100**) • For a SONET interface, enter the keyword **sonet** followed by the slot/port information (**e.g. passive-interface so 2/2**). • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information (e**.g. passive-interface ten 2/3**). • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094 (**e.g. passive-interface vlan 2222)**. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS. The default keyword sets all interfaces on this OSPF process as passive. The passive interface can be removed from select interfaces using the **no passive-interface interface** command while **passive interface default** is configured.

To enable both receiving and sending routing updates, enter the **no passive-interface interface** command.

When you configure a passive interface, the **show ip ospf** *process-id* **interface** command (Figure 413) adds the words "`passive interface`" to indicate that hello packets are not transmitted on that interface.

Figure 32-13. Command Example: show ip ospf *process-id* interface

```
FTOS#show ip ospf 34 int

GigabitEthernet 0/0 is up, line protocol is down
  Internet Address 10.1.2.100/24, Area 1.1.1.1
  Process ID 34, Router ID 10.1.2.100, Network Type BROADCAST, Cost: 10
  Transmit Delay is 1 sec, State DOWN, Priority 1
  Designated Router (ID) 10.1.2.100, Interface address 0.0.0.0
  Backup Designated Router (ID) 0.0.0.0, Interface address 0.0.0.0
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    Hello due in 13:39:46
  Neighbor Count is 0, Adjacent neighbor count is 0

GigabitEthernet 0/1 is up, line protocol is down
  Internet Address 10.1.3.100/24, Area 2.2.2.2
  Process ID 34, Router ID 10.1.2.100, Network Type BROADCAST, Cost: 10
  Transmit Delay is 1 sec, State DR, Priority 1
  Designated Router (ID) 10.1.2.100, Interface address 10.1.3.100
  Backup Designated Router (ID) 0.0.0.0, Interface address 0.0.0.0
  Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
    No Hellos (Passive interface)   <-- Interface is not running the OSPF protocol.
  Neighbor Count is 0, Adjacent neighbor count is 0

Loopback 45 is up, line protocol is up
  Internet Address 10.1.1.23/24, Area 2.2.2.2
  Process ID 34, Router ID 10.1.2.100, Network Type LOOPBACK, Cost: 1
Loopback interface is treated as a stub Host.
FTOS#
```

Enable fast-convergence

The fast-convergence CLI sets the minimum origination and arrival LSA parameters to zero (0), allowing rapid route calculation. When fast-convergence is disabled, origination and arrival LSA parameters are set to 5 seconds and 1 second, respectively.

Setting the convergence parameter (1-4) indicates the actual convergence level. Each convergence setting adjusts the LSA parameters to zero, but the fast-convergence parameter setting allows for even finer tuning of the convergence speed. The higher the number, the faster the convergence. Use the following command in the ROUTER OSPF mode to enable or disable fast-convergence.

Command Syntax	Command Mode	Usage
fast-convergence {*number*}	CONFIG-ROUTER-OSPF-id	Enable OSPF fast-convergence and specify the convergence level. **Parameter: 1-4** The higher the number, the faster the convergence. When disabled, the parameter is set at 0 (Figure 32-15).
		Note: A higher convergence level can result in occasional loss of OSPF adjacency. Generally, convergence level 1 meets most convergence requirements. Higher convergence levels should only be selected following consultation with Dell Force10 technical support.

Figure 32-14 shows the convergence settings when fast-convergence is enabled and Figure 32-15 shows settings when fast-convergence is disabled. These displays appear with the **show ip ospf** command.

Figure 32-14. Command Example: show ip ospf *process-id* (fast-convergence enabled)

```
FTOS(conf-router_ospf-1)#fast-converge 2
FTOS(conf-router_ospf-1)#ex
FTOS(conf)#ex
FTOS#show ip ospf 1
Routing Process ospf 1 with ID 192.168.67.2
Supports only single TOS (TOS0) routes
SPF schedule delay 5 secs, Hold time between two SPFs 10 secs
Convergence Level 2
Min LSA origination 0 secs, Min LSA arrival 0 secs
Number of area in this router is 0, normal 0 stub 0 nssa 0
FTOS#
```

Fast-converge parameter setting

LSA Parameters

Figure 32-15. Command example: show ip ospf *process-id* (fast-convergence disabled)

```
FTOS#(conf-router_ospf-1)#no fast-converge
FTOS#(conf-router_ospf-1)#ex
FTOS#(conf)#ex
FTOS##show ip ospf 1
Routing Process ospf 1 with ID 192.168.67.2
Supports only single TOS (TOS0) routes
SPF schedule delay 5 secs, Hold time between two SPFs 10 secs
Convergence Level 0
Min LSA origination 5 secs, Min LSA arrival 1 secs
Number of area in this router is 0, normal 0 stub 0 nssa 0
FTOS#
```

Fast-converge parameter setting

LSA Parameters

Change OSPFv2 parameters on interfaces

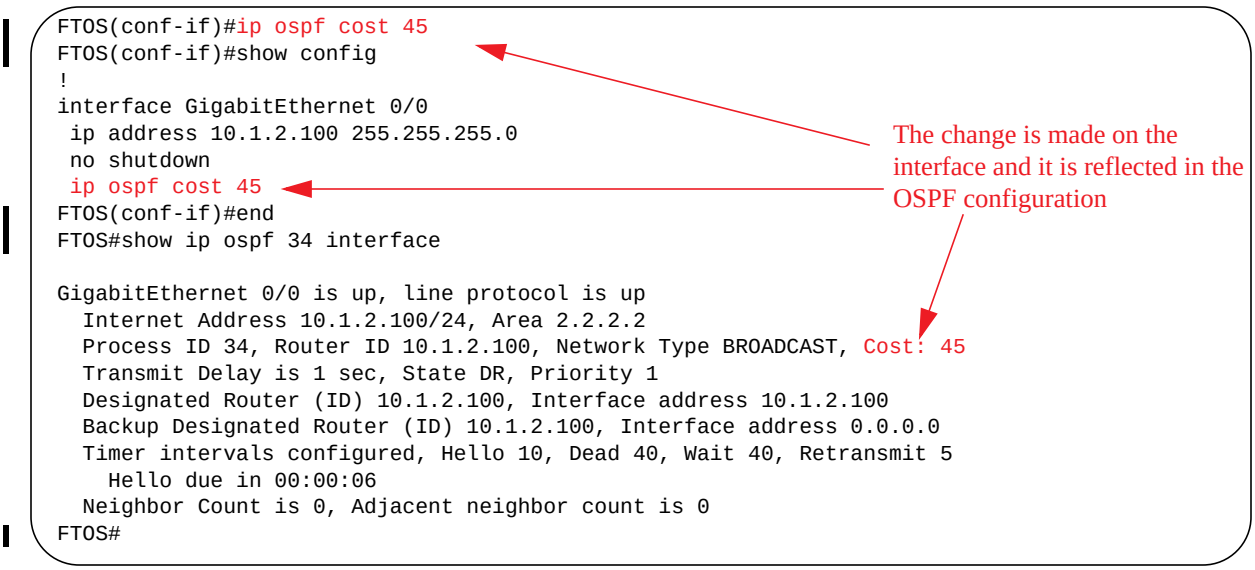
In FTOS, you can modify the OSPF settings on the interfaces. Some interface parameter values must be consistent across all interfaces to avoid routing errors. For example, you must set the same time interval for the hello packets on all routers in the OSPF network to prevent misconfiguration of OSPF neighbors.

Use any or all of the following commands in CONFIGURATION INTERFACE mode to change OSPFv2 parameters on the interfaces:

Command Syntax	Command Mode	Usage
ip ospf *cost*	CONFIG-INTERFACE	Change the cost associated with OSPF traffic on the interface. Cost: 1 to 65535 (default depends on the interface speed).
ip ospf dead-interval *seconds*	CONFIG-INTERFACE	Change the time interval the router waits before declaring a neighbor dead. Configure Seconds range: 1 to 65535 (default is 40 seconds).
	The dead interval must be four times the hello interval. The dead interval must be the same on all routers in the OSPF network.	
ip ospf hello-interval *seconds*	CONFIG-INTERFACE	Change the time interval between hello-packet transmission. Seconds range: from 1 to 65535 (default is 10 seconds).
	The hello interval must be the same on all routers in the OSPF network.	
ip ospf message-digest-key *keyid* **md5** *key*	CONFIG-INTERFACE	Use the MD5 algorithm to produce a message digest or key, which is sent instead of the key. Keyid range: 1 to 255 Key: a character string
	Be sure to write down or otherwise record the Key. You cannot learn the key once it is configured. You must be careful when changing this key.	
ip ospf priority *number*	CONFIG-INTERFACE	Change the priority of the interface, which is used to determine the Designated Router for the OSPF broadcast network. Number range: 0 to 255 (the default is 1).
ip ospf retransmit-interval *seconds*	CONFIG-INTERFACE	Change the retransmission interval between LSAs. Seconds range: from 1 to 65535 (default is 5 seconds).
	The retransmit interval must be the same on all routers in the OSPF network.	
ip ospf transmit-delay *seconds*	CONFIG-INTERFACE	Change the wait period between link state update packets sent out the interface. Seconds range: from 1 to 65535 (default is 1 second).
	The transmit delay must be the same on all routers in the OSPF network.	

Use the **show config** command in CONFIGURATION INTERFACE mode (Figure 32-16) to view interface configurations. Use the **show ip ospf interface** command in EXEC mode to view interface status in the OSPF process.

Figure 32-16. Changing the OSPF Cost Value on an Interface



Enable OSPFv2 authentication

Use the following commands in CONFIGURATION INTERFACE mode to enable or change various OSPF authentication parameters:

Command Syntax	Command Mode	Usage
ip ospf authentication-key *key*	CONFIG-INTERFACE	Set clear text authentication scheme on the interface. Configure a *key* that is a text string no longer than eight characters. All neighboring routers must share the same password to exchange OSPF information.
ip ospf auth-change-wait-time *seconds*	CONFIG-INTERFACE	Set the authentication change wait time in *seconds* between 0 and 300 for the interface. This is the amount of time OSPF has available to change its interface authentication type. During the auth-change-wait-time, OSPF sends out packets with both the new and old authentication schemes. This transmission stops when the period ends. The default is 0 seconds.

Enable OSPFv2 graceful restart

Graceful Restart is enabled for the global OSPF process. Use these commands to configure OSPFv2 graceful restart. Refer to Graceful Restart on page 700 for feature details.

The Dell Force10 implementation of OSPFv2 graceful restart enables you to specify:

- **grace period**—the length of time the graceful restart process can last before OSPF terminates it.

- **helper-reject neighbors**—the router ID of each restart router that does not receive assistance from the configured router.
- **mode**—the situation or situations that trigger a graceful restart.
- **role**—the role or roles the configured router can perform.

Note: By default, OSPFv2 graceful restart is disabled.

You enable OSPFv2 graceful restart in CONFIGURATION ROUTER OSPF mode.

Command Syntax	Command Mode	Usage
graceful-restart grace-period *seconds*	CONFIG-ROUTER-OSPF-id	Enable OSPFv2 graceful-restart globally and set the grace period. Seconds range: between 40 and 3000 This is the period of time that an OSPFv2 router's neighbors will advertise it as fully adjacent, regardless of the synchronization state, during a graceful restart. OSPFv2 terminates this process when the grace period ends.
graceful-restart helper-reject *router-id*	CONFIG-ROUTER-OSPF-id	Enter the Router ID of the OSPFv2 helper router from which the router does not accept graceful restart assistance. This applies to the specified router only. IP Address: A.B.C.D
graceful-restart mode [planned-only | unplanned-only]	CONFIG-ROUTER-OSPF-id	Specify the operating mode in which graceful-restart functions. FTOS supports the following options: • Planned-only. The OSPFv2 router supports graceful-restart for planned restarts only. A planned restart is when the user manually enters a fail-over command to force the primary RPM over to the secondary RPM. During a planned restart, OSPF sends out a Grace LSA before the system switches over to the secondary RPM. OSPF also is notified that a planned restart is happening. • Unplanned-only. The OSPFv2 router supports graceful-restart for only unplanned restarts. During an unplanned restart, OSPF sends out a Grace LSA once the secondary RPM comes online. By default, OSPFv2 supports both planned and unplanned restarts. Selecting one or the other mode restricts OSPFv2 to the single selected mode.
graceful-restart role [helper-only | restart-only]	CONFIG-ROUTER-OSPF-id	Configure the graceful restart role or roles that this OSPFv2 router performs. FTOS supports the following options: • Helper-only. The OSPFv2 router supports graceful-restart only as a helper router. • Restart-only. The OSPFv2 router supports graceful-restart only during unplanned restarts. By default, OSPFv2 supports both restarting and helper roles. Selecting one or the other role restricts OSPFv2 to the single selected role.

When you configure a graceful restart on an OSPFv2 router, the **show run ospf** command (Figure 32-17) displays information similar to the following.

Figure 32-17. Command Example: show run ospf

```
FTOS#show run ospf
!
router ospf 1
 graceful-restart grace-period 300
 graceful-restart role helper-only
 graceful-restart mode unplanned-only
 graceful-restart helper-reject 10.1.1.1
 graceful-restart helper-reject 20.1.1.1
 network 10.0.2.0/24 area 0
FTOS#
```

Use the following command to disable OSPFv2 graceful-restart after you have enabled it.

Command Syntax	Command Mode	Usage
no graceful-restart grace-period	CONFIG-ROUTER-OSPF-id	Disable OSPFv2 graceful-restart. Returns OSPF graceful-restart to its default state.

For more information on OSPF graceful restart, refer to the *FTOS Command Line Interface Reference*.

Configure virtual links

Areas within OSPF must be connected to the backbone area (Area ID 0.0.0.0). If an OSPF area does not have a direct connection to the backbone, at least one virtual link is required. Virtual links must be configured on an ABR connected to the backbone.

- hello-interval: help packet
- retransmit-interval: LSA retransmit interval
- transmit-delay: LSA transmission delay
- dead-interval: dead router detection time
- authentication-key: authentication key
- message-digest-key: MD5 authentication key

Use the following command in CONFIGURATION ROUTER OSPF mode to configure virtual links.

Command Syntax	Command Mode	Usage
area *area-id* **virtual-link** *router-id* [hello-interval *seconds* | retransmit-interval *seconds* | transmit-delay *seconds* | dead-interval *seconds* | authentication-key *key* | message-digest-key *keyid* md5 *key*]	CONFIG-ROUTER-OSPF-id	Configure the optional parameters of a virtual link: • Area ID: assigned earlier (0-65535 or A.B.C.D) • Router ID: IP address associated with the virtual link neighbor • Hello Interval Seconds: 1-8192 (default 10) • Retransmit Interval Seconds: 1-3600 (default 5) • Transmit Delay Seconds: 1-3600 (default 1) • Dead Interval Seconds: 1-8192 (default 40) • Authentication Key: 8 characters • Message Digest Key: 1-255 • MD5 Key: 16 characters Only the Area ID and Router ID require configuration to create a virtual link. If no other parameter is entered, the defaults are used. Use EITHER the Authentication Key or the Message Digest (MD5) key.

Use the **show ip ospf** *process-id* **virtual-links** command (Figure 32-18) in the EXEC mode to view the virtual link.

Figure 32-18. Command Example: show ip ospf *process-id* virtual-links

```
FTOS#show ip ospf 1 virtual-links

Virtual Link to router 192.168.253.5 is up
    Run as demand circuit
    Transit area 0.0.0.1, via interface GigabitEthernet 13/16, Cost of using 2
    Transmit Delay is 1 sec, State POINT_TO_POINT,
        Timer intervals configured, Hello 10, Dead 40, Wait 40, Retransmit 5
        Hello due in 00:00:02
FTOS#
```

Filter routes

To filter routes, use prefix lists. OSPF applies prefix lists to incoming or outgoing routes. Incoming routes must meet the conditions of the prefix lists, and if they do not, OSPF does not add the route to the routing table. Configure the prefix list in CONFIGURATION PREFIX LIST mode prior to assigning it to the OSPF process.

Command Syntax	Command Mode	Usage
ip prefix-list *prefix-name*	CONFIGURATION	Create a prefix list and assign it a unique name. You are in PREFIX LIST mode.

Command Syntax	Command Mode	Usage
seq *sequence-number* {deny |permit} *ip-prefix* [ge min-prefix-length] [le max-prefix-length]	CONFIG- PREFIX LIST	Create a prefix list with a sequence. number and a deny or permit action. The optional parameters are: **ge** min-prefix-length: is the minimum prefix length to be matched (0 to 32). **le** *max-prefix-length:* is the maximum prefix length to be matched (0 to 32).

For configuration information on prefix lists, refer to *IP Access Control Lists, Prefix Lists, and Route-maps* chapter in the *FTOS Configuration Guide*.

Use the following commands in CONFIGURATION-ROUTER OSPF mode to apply prefix lists to incoming or outgoing OSPF routes

Command Syntax	Command Mode	Usage
distribute-list *prefix-list-name* **in** [*interface*]	CONFIG-ROUTER-OSPF-id	Apply a configured prefix list to incoming OSPF routes.
distribute-list *prefix-list-name* **out** [**connected** | **isis** | **rip** | **static**]	CONFIG-ROUTER-OSPF-id	Assign a configured prefix list to outgoing OSPF routes.

Redistribute routes

You can add routes from other routing instances or protocols to the OSPF process. With the **redistribute** command syntax, you can include RIP, static, or directly connected routes in the OSPF process.

Note: Do not route iBGP routes to OSPF unless there are route-maps associated with the OSPF redistribution.

Use the following command in CONFIGURATION- ROUTER-OSPF mode to redistribute routes:

Command Syntax	Command Mode	Usage
redistribute {**bgp** | **connected** | **isis** | **rip** | **static**} [**metric** *metric-value* | **metric-type** *type-value*] [**route-map** *map-name*] [**tag** *tag-value*]	CONFIG-ROUTER-OSPF-id	Specify which routes will be redistributed into OSPF process. Configure the following required and optional parameters: • **bgp, connected**, **isis**, **rip**, or **static**: enter one of the keyword to redistribute those routes. **rip** is supported only on E-Series. • **metric** *metric-value* range: 0 to 4294967295. • **metric-type** *metric-type*: 1 for OSPF external route type 1 or 2 for OSPF external route type 2. • **route-map** *map-name*: enter a name of a configured route map. • **tag** *tag-value* range: 0 to 4294967295.

To view the current OSPF configuration, use the **show running-config ospf** command in the EXEC mode or the **show config** command in the ROUTER OSPF mode

Figure 32-19. Command Example: show config

```
FTOS(conf-router_ospf)#show config
!
router ospf 34
 network 10.1.2.32 0.0.0.255 area 2.2.2.2
 network 10.1.3.24 0.0.0.255 area 3.3.3.3
 distribute-list dilling in
FTOS(conf-router_ospf)#
```

Troubleshooting OSPFv2

FTOS has several tools to make troubleshooting easier. Be sure to check the following, as these are typical issues that interrupt an OSPFv2 process. Note that this is not a comprehensive list, just some examples of typical troubleshooting checks.

- Has OSPF been enabled globally?
- Is the OSPF process active on the interface?
- Are adjacencies established correctly?
- Are the interfaces configured for Layer 3 correctly?
- Is the router in the correct area type?
- Have the routes been included in the OSPF database?
- Have the OSPF routes been included in the routing table (not just the OSPF database)?

Some useful troubleshooting commands are:

- show interfaces
- show protocols
- debug IP OSPF events and/or packets
- show neighbors
- show virtual links
- show routes

Note: If you are using Multi-Process OSPF, you must enter the Process ID to view information regarding a specific OSPF process. If you do not enter the Process ID, only the first configured process is listed.

Use the **show running-config ospf** command to see the state of all the enabled OSPFv2 processes.

Command Syntax	Command Mode	Usage
show running-config ospf	EXEC Privilege	View the summary of all OSPF process IDs enables on the router.

Figure 32-20. Command Example: show running-config ospf

```
FTOS#show run ospf
!
router ospf 3
!
router ospf 4
 router-id 4.4.4.4
 network 4.4.4.0/28 area 1
!
router ospf 5
!
router ospf 6
!
router ospf 7
 mib-binding
!
router ospf 8
!
router ospf 90
 area 2 virtual-link 4.4.4.4
 area 2 virtual-link 90.90.90.90 retransmit-interval 300
!
ipv6 router ospf 999
 default-information originate always
 router-id 10.10.10.10
```

Use the following commands in EXEC Privilege mode to get general route and links status information.

Command Syntax	Command Mode	Usage
show ip route summary	EXEC Privilege	View the summary information of the IP routes
show ip ospf database	EXEC Privilege	View the summary information for the OSPF database

Use the following command in EXEC Privilege mode to view the OSPFv2 configuration for a neighboring router:

Command Syntax	Command Mode	Usage
show ip ospf neighbor	EXEC Privilege	View the configuration of OSPF neighbors.

Use the following command in EXEC Privilege mode to configure the debugging options of an OSPFv2 process:

Command Syntax	Command Mode	Usage
debug ip ospf *process-id* [**event** | **packet** | **spf**]	EXEC Privilege	View debug messages. To view debug messages for a specific OSPF process ID, enter **debug ip ospf** *process-id.* If you do not enter a process ID, the command applies to the first OSPF process. To view debug messages for a specific operation, enter one of the optional keywords: • **event**: View OSPF event messages • **packet**: View OSPF packets. • **spf**: View shortest path first (spf) information.

To display a summary of the information stored in the OSPFv2 database of the router, enter the **show ip ospf database database-summary** command. Note that the number of Type-9 Grace LSAs received from restarting neighbor OSPFv2 routers are also displayed.

Command Syntax	Command Mode	Usage
show ip ospf database database-summary	EXEC Privilege	View a summary of OSPFv2 database information.

Figure 32-21. Command Example: show ip ospf database database-summary

```
FTOS#show ip ospf database database-summary
!
OSPF Router with ID (200.1.1.1) (Process ID 1)

      Area ID        Router Net    S-Net  S-ASBR Type7  Type9  Type10 Total  ChSum
      0               4      3      3000   0      0      1      0      3008   0x5e69164
```

Sample Configurations for OSPFv2

The following configurations are examples for enabling OSPFv2. These are not comprehensive directions. They are intended to give you a some guidance with typical configurations.

You can copy and paste from these examples to your CLI. Be sure you make the necessary changes to support your own IP Addresses, Interfaces, Names, etc.

Basic OSPFv2 Router Topology

The following illustration is a sample basic OSPFv2 topology.

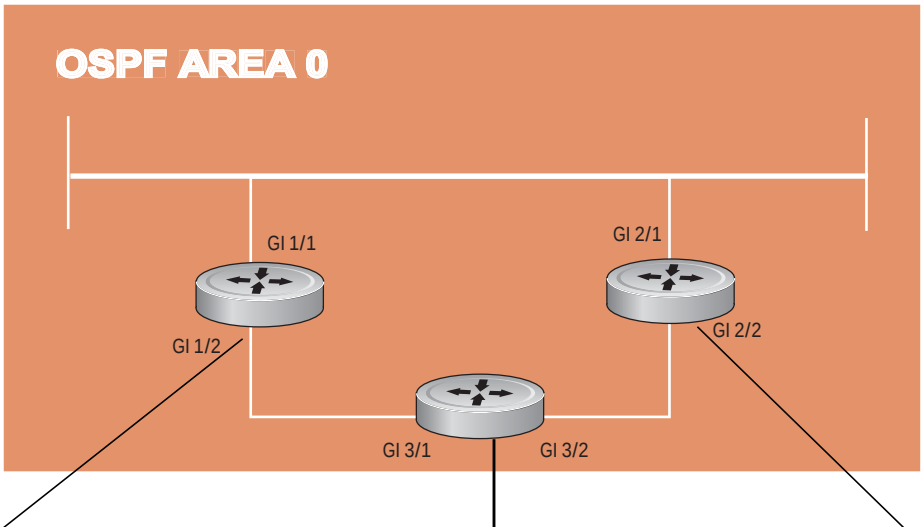


Figure 32-22. Basic topology and CLI commands for OSPFv2

```
router ospf 11111
 network 10.0.11.0/24 area 0
 network 10.0.12.0/24 area 0
 network 192.168.100.0/24 area 0
!
interface GigabitEthernet 1/1
 ip address 10.1.11.1/24
 no shutdown
!
interface GigabitEthernet 1/2
 ip address 10.2.12.2/24
 no shutdown
!
interface Loopback 10
 ip address 192.168.100.100/24
 no shutdown
```

```
router ospf 33333
 network 192.168.100.0/24 area 0
 network 10.0.13.0/24 area 0
 network 10.0.23.0/24 area 0
!
interface Loopback 30
 ip address 192.168.100.100/24
 no shutdown
!
interface GigabitEthernet 3/1
 ip address 10.1.13.3/24
 no shutdown
!
interface GigabitEthernet 3/2
 ip address 10.2.13.3/24
no shutdown
```

```
router ospf 22222
 network 192.168.100.0/24 area 0
 network 10.2.21.0/24 area 0
 network 10.2.22.0/24 area 0
!
interface Loopback 20
 ip address 192.168.100.20/24
 no shutdown
!
interface GigabitEthernet 2/1
 ip address 10.2.21.2/24
 no shutdown
!
interface GigabitEthernet 2/2
 ip address 10.2.22.2/24
 no shutdown
```

Configuration Task List for OSPFv3 (OSPF for IPv6)

Open Shortest Path First version 3 (OSPF for IPv6) is supported on platforms C E

The configuration options of OSPFv3 are the same as those for OSPFv2, but may be configured with differently labeled commands. Process IDs and areas need to be specified. Interfaces and addresses need to be included in the process. Areas can be defined as stub or totally stubby.

The interfaces must be in IPv6 Layer-3 mode (assigned an IPv6 IP address) and enabled so that they can send and receive traffic. The OSPF process must know about these interfaces. To make the OSPF process aware of these interfaces, they must be assigned to OSPF areas.

TheOSPFv3 **ipv6 ospf area** command enables OSPFv3 on the interface and places the interface in an area. With OSPFv2, two commands are required to accomplish the same tasks: the **router ospf** command to create the OSPF process, then the **network area** command to enable OSPF on an interface. Note that the OSPFv2 **network area** command can enable OSPF on multiple interfaces with the single command, while the OSPFv3 **ipv6 ospf area** command must be configured on each interface that will be running OSPFv3.

All IPv6 addresses on an interface are included in the OSPFv3 process that is created on the interface.

OSPFv3 for IPv6 is enabled by specifying an OSPF Process ID and an Area in the INTERFACE mode. If an OSPFv3 process has not yet been created, it is created automatically. All IPv6 addresses configured on the interface are included in the specified OSPF process.

Note: IPv6 and OSPFv3 do *not* support Multi-Process OSPF. Only a single OSPFv3 process is can be enabled.

- Enable IPv6 Unicast Routing
- Assign IPv6 addresses on an interface
- Assign Area ID on interface
- Assign OSPFv3 Process ID and Router ID Globally
- Configure stub areas
- Configure Passive-Interface
- Redistribute routes
- Configure a default route
- (Optional) Enable OSPFv3 graceful restart
- (Optional) OSPFv3 Authentication Using IPsec

Enable IPv6 Unicast Routing

Command Syntax	Command Mode	Usage
ipv6 unicast routing	CONFIGURATION	Enables IPv6 unicast routing globally.

Assign IPv6 addresses on an interface

Command Syntax	Command Mode	Usage
ipv6 address *ipv6 address*	CONF-INT-type slot/port	Assign IPv6 address to the interface. IPv6 addresses are normally written as eight groups of four hexadecimal digits, where each group is separated by a colon (:). FORMAT: A:B:C::F/128
no shutdown	CONF-INT-type slot/port	Bring the interface up.

Assign Area ID on interface

Command Syntax	Command Mode	Usage
ipv6 ospf *process-id* **area** *area-id*	CONF-INT-type slot/port	Assign the OSPFv3 process and an OSPFv3 area to this interface. *process-id*: The Process ID number assigned above. *area-id*: the area ID for this interface.

The **ipv6 ospf area** command enables OSPFv3 on an interface and places the interface in the specified area. Additionally, it creates the OSPFv3 process with ID on the router. OSPFv2 required two commands are required to accomplish the same tasks: the **router ospf** command to create the OSPF process, then the **network area** command to enable OSPFv2 on an interface. Note that the OSPFv2 **network area** command can enable OSPFv2 on multiple interfaces with the single command, whereas the OSPFv3 **ipv6 ospf area** command must be configured on each interface that will be running OSPFv3.

Assign OSPFv3 Process ID and Router ID Globally

Command Syntax	Command Mode	Usage
ipv6 router ospf {*process ID}*	CONFIGURATION	Enable the OSPFv3 process globally and enter OSPFv3 mode. Range: 0-65535
router-id *{number}*	CONF-IPV6-ROUTER-OSPF	Assign the Router ID for this OSPFv3 process *number*: IPv4 address Format: A.B.C.D **Note:** The router-id for an OSPFv3 router is entered as an IPv4 IP address.

Configure stub areas

Command Syntax	Command Mode	Usage
area *area-id* **stub** [**no-summary**]	CONF-IPV6-ROUTER-OSPF	Configure the area as a stub area. Use the **no-summary** keywords to prevent transmission in to the area of summary ASBR LSAs. *Area ID* is a number or IP address assigned when creating the Area. The Area ID can be represented as a number between 0 – 65536 if a dotted decimal format is assigned, rather than an IP address.

Configure Passive-Interface

Use the following command to suppress the interface's participation on an OSPFv3 interface. This command stops the router from sending updates on that interface.

Command Syntax	Command Mode	Usage
passive-interface {type slot/port}	CONF-IPV6-ROUTER-OSPF	Specify whether some or all some of the interfaces will be passive. **Interface** identifies the specific interface that will be passive. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information (**e.g. passive-interface gi 2/1**). • For a port channel, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale (**e.g. passive-interface po 100**) • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information (**e.g. passive-interface ten 2/3**). • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094 (**e.g. passive-interface vlan 2222)**. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

To enable both receiving and sending routing updates, enter the **no passive-interface interface** command.

When you configure a passive interface, the **show ipv6 ospf interface** command adds the words "`passive interface`" to indicate that hello packets are not transmitted on that interface.

Redistribute routes

You can add routes from other routing instances or protocols to the OSPFv3 process. With the **redistribute** command syntax, you can include RIP, static, or directly connected routes in the OSPF process.

Command Syntax	Command Mode	Usage
redistribute {**bgp** | **connected** | **static**} [**metric** *metric-value* | **metric-type** *type-value*] [**route-map** *map-name*] [**tag** *tag-value*]	CONF-IPV6-ROUTER-OSPF	Specify which routes will be redistributed into OSPF process. Configure the following required and optional parameters: • **bgp, connected**, or **static**: enter one of the keyword to redistribute those routes. • **metric** *metric-value* range: 0 to 4294967295. • **metric-type** *metric-type*: 1 for OSPFv3 external route type 1 OR 2 for OSPFv3 external route type 2. • **route-map** *map-name*: enter a name of a configured route map. • **tag** *tag-value* range: 0 to 4294967295.

Configure a default route

Configure FTOS to generate a default external route into the OSPFv3 routing domain.

Command Syntax	Command Mode	Usage
default-information originate [**always** [**metric** *metric-value*] [**metric-type** *type-value*]] [**route-map** *map-name*]	CONF-IPV6-ROUTER-OSPF	Specify the information for the default route. Configure the following required and optional parameters: • **always**: indicate that default route information must always be advertised • **metric** *metric-value* range: 0 to 4294967295. • **metric-type** *metric-type*: 1 for OSPFv3 external route type 1 OR 2 for OSPFv3 external route type 2. • **route-map** *map-name*: enter a name of a configured route map.

Enable OSPFv3 graceful restart

Graceful Restart for OSPFv3 is supported only on platform E(T). Refer to Graceful Restart on page 700 for more information on the feature.

By default, OSPFv3 graceful restart is disabled and functions only in a helper role to help restarting neighbor routers in their graceful restarts when it receives a Grace LSA.

To enable OSPFv3 graceful restart, you must enter the **ipv6 router ospf process-id** command to enter OSPFv3 configuration mode and then configure a grace period using the **graceful-restart grace-period** command. The grace period is the length of time that OSPFv3 neighbors continue to advertise the restarting router as though it is fully adjacent. When graceful restart is enabled (restarting role), an OSPFv3 restarting expects its OSPFv3 neighbors to help when it restarts by not advertising the broken link.

When you enable the helper-reject role on an interface with the **ipv6 ospf graceful-restart helper-reject** command, you reconfigure OSPFv3 graceful restart to function in a "restarting-only" role. OSPFv3 does not participate in the graceful restart of a neighbor. (Note that you enter the **ipv6 ospf graceful-restart helper-reject** command in Interface configuration mode.)

Command Syntax	Command Mode	Usage
graceful-restart grace-period *seconds*	CONF-IPV6-ROUTER-OSPF	Enable OSPFv3 graceful restart globally by setting the grace period (in seconds). Valid values are from 40 to 1800 seconds.
ipv6 ospf graceful-restart helper-reject	INTERFACE	Configure an OSPFv3 interface to not act upon the Grace LSAs that it receives from a restarting OSPFv3 neighbor.
graceful-restart mode [planned-only | unplanned-only]	CONF-IPV6-ROUTER-OSPF	Specify the operating mode and type of events that trigger a graceful restart: • Planned-only. The OSPFv3 router supports graceful restart only for planned restarts. A planned restart is when you manually enter a **redundancy force-failover rpm** command to force the primary RPM over to the secondary RPM. During a planned restart, OSPFv3 sends out a Grace LSA before the system switches over to the secondary RPM. OSPFv3 is notified that a planned restart is happening. • Unplanned-only. The OSPFv3 router supports graceful-restart only for unplanned restarts. During an unplanned restart, OSPFv3 sends out a Grace LSA once the secondary RPM comes online. Default: Both planned and unplanned restarts trigger an OSPFv3 graceful restart. Selecting one or the other mode restricts OSPFv3 to the single selected mode.

To disable OSPFv3 graceful restart when it is enabled, enter the following command:

Command Syntax	Command Mode	Usage
no graceful-restart grace-period	CONF-IPV6-ROUTER-OSPF	Disable OSPFv3 graceful-restart.

To display information on the use and configuration of OSPFv3 graceful restart, enter any of the following commands:

Command Syntax	Command Mode	Usage
show run ospf	EXEC Privilege	Display the graceful-restart configuration for OSPFv2 and OSPFv3 (Figure 32-23).
show ipv6 ospf database grace-lsa	EXEC Privilege	Display the Type-11 Grace LSAs sent and received on an OSPFv3 router (Figure 32-24).
show ipv6 ospf database database-summary	EXEC Privilege	Display the currently configured OSPFv3 parameters for graceful restart (Figure 32-25).

Figure 32-23. Command Example: show run ospf

```
FTOS#show run ospf
!
router ospf 1
 router-id 200.1.1.1
 log-adjacency-changes
 graceful-restart grace-period 180
 network 20.1.1.0/24 area 0
 network 30.1.1.0/24 area 0
!
ipv6 router ospf 1
 log-adjacency-changes
 graceful-restart grace-period 180
```

Figure 32-24. Command Example: show ipv6 ospf database database-summary

```
FTOS#show ipv6 ospf database database-summary
!
OSPFv3 Router with ID (200.1.1.1) (Process ID 1)

Process 1 database summary
Type                    Count/Status
Oper Status              1
Admin Status             1
Area Bdr Rtr Status      0
AS Bdr Rtr Status        1
AS Scope LSA Count       0
AS Scope LSA Cksum sum   0
Originate New LSAS       73
Rx New LSAS              114085
Ext LSA Count            0
Rte Max Eq Cost Paths    5
GR grace-period          180
GR mode                  planned and unplanned

Area 0 database summary
Type                    Count/Status
Brd Rtr Count            2
AS Bdr Rtr Count         2
LSA count                12010
Summary LSAs             1
Rtr LSA Count            4
Net LSA Count            3
Inter Area Pfx LSA Count 12000
Inter Area Rtr LSA Count 0
Group Mem LSA Count      0
Type-7 LSA count         0
Intra Area Pfx LSA Count 3
Intra Area TE LSA Count  0
```

Figure 32-25. Command Example: show ipv6 ospf database grace-lsa

```
FTOS#show ipv6 ospf database grace-lsa
!
Type-11 Grace LSA (Area 0)

LS Age              : 10
Link State ID       : 6.16.192.66
Advertising Router  : 100.1.1.1
LS Seq Number       : 0x80000001
Checksum            : 0x1DF1
Length              : 36
Associated Interface : Gi 5/3
Restart Interval    : 180
Restart Reason      : Switch to Redundant Processor
```

OSPFv3 Authentication Using IPsec

OSPFv3 Authentication Using IPsec is supported only on platform: E T

Starting in release 8.4.2.0, OSPFv3 uses the IP Security (IPsec) to provide authentication for OSPFv3 packets. IPsec authentication ensures security in the transmission of OSPFv3 packets between IPsec-enabled routers.

IPsec is a set of protocols developed by the IETF to support secure exchange of packets at the IP layer. IPsec supports two encryption modes: Transport and Tunnel. Transport mode encrypts only the data portion (payload) of each packet, but leaves the header untouched. The more secure Tunnel mode encrypts both the header and the payload. On the receiving side, an IPsec-compliant device decrypts each packet.

Note: FTOS supports only transport encryption mode in OSPFv3 authentication with IPsec.

With IPsec-based authentication, crypto images are used to include the IPsec secure socket application programming interface (API) required for use with OSPFv3.

To ensure integrity, data origin authentication, detection and rejection of replays, and confidentiality of the packet, RFC 4302 and RFC 4303 propose using two security protocols - AH (authentication header) and ESP (encapsulating security payload). For OSPFv3, these two IPsec protocols provide interoperable, high-quality cryptographically-based security.

- The IPsec authentication header is used in packet authentication to verify that data is not altered during transmission and ensures that users are communicating with the intended individual or organization. The authentication header is inserted after the IP header with a value of 51. AH provides integrity and validation of data origin by authenticating every OSPFv3 packet. For detailed information on the IP AH protocol, refer to RFC 4302.
- The encapsulating security payload encapsulates data, enabling the protection of data that follows in the datagram. ESP provides authentication and confidentiality of every packet. The ESP extension header is designed to provide a combination of security services for both IPv4 and IPv6. The ESP header is inserted after the IP header and before the next layer protocol header in transport mode. It is possible that the ESP header is inserted between the next layer protocol header and encapsulated IP header in tunnel mode. The tunnel mode is not supported in FTOS. For detailed information on the IP ESP protocol, refer to RFC 4303.

In OSPFv3 communication, IPsec provides security services between a pair of communicating hosts or security gateways using either AH or ESP. In an authentication policy on an interface or in an OSPF area, AH and ESP are used alone; in an encryption policy, AH and ESP may be used together. The difference between the two mechanisms is the extent of the coverage. ESP only protects IP header fields if they are encapsulated by ESP.

The set of IPsec protocols that are employed for authentication and encryption and the ways in which they are employed is user-dependent. When IPsec is correctly implemented and deployed, it does not adversely affect users or hosts. AH and ESP are designed to be cryptographic algorithm-independent.

OSPFv3 Authentication using IPsec: Configuration Notes

OSPFv3 authentication using IPsec is implemented according to the specifications in RFC 4552, including:

- To use IPsec, you configure an authentication (using AH) or encryption (using ESP) security policy on an interface or in an OSPFv3 area. Each security policy consists of a security policy index (SPI) and the key used to validate OSPFv3 packets. After IPsec is configured for OSPFv3, IPsec operation is invisible to the user.
  - Only one security protocol (AH or ESP) can be enabled at a time on an interface or for an area. IPsec AH is enabled with the **ipv6 ospf authentication** command; IPsec ESP is enabled with the **ipv6 ospf encryption** command.
  - The security policy configured for an area is inherited by default on all interfaces in the area.
  - The security policy configured on an interface overrides any area-level configured security for the area to which the interface is assigned.
  - The configured authentication or encryption policy is applied to all OSPFv3 packets transmitted on the interface or in the area. The IPsec security associations (SAs) are the same on inbound and outbound traffic on an OSPFv3 interface.
  - There is no maximum AH or ESP header length because the headers have fields with variable lengths.
- Manual key configuration is supported in an authentication or encryption policy (dynamic key configuration using the Internet Key Exchange (IKE) protocol is not supported).
- In an OSPFv3 authentication policy:
  - AH is used to authenticate OSPFv3 headers and certain fields in IPv6 headers and extension headers.
  - MD5 and SHA1 authentication types are supported; encrypted and unencrypted keys are supported.
- In an OSPFv3 encryption policy:
  - Both encryption and authentication are used.
  - IPsec security associations (SAs) are supported only in transport mode (tunnel mode is not supported).
  - ESP with null encryption is supported for authenticating only OSPFv3 protocol headers.
  - ESP with non-null encryption is supported for full confidentiality.
  - 3DES, DES, AES-CBC, and NULL encryption algorithms are supported; encrypted and unencrypted keys are supported.

Note: You may encrypt all keys on a router by using the **service password-encryption** command in global configuration mode. However, this command does not provide a high level of network security. To enable key encryption in an IPsec security policy at an interface or area level, specify **7** for [*key-encryption-type*] when you enter the **ipv6 ospf authentication ipsec** or **ipv6 ospf encryption ipsec** command.

- To configure an IPsec security policy for authenticating or encrypting OSPFv3 packets on a physical, port-channel, or VLAN interface or OSPFv3 area, perform any of the following tasks:
  - Configuring IPsec Authentication on an Interface
  - Configuring IPsec Encryption on an Interface

- Configuring IPsec Authentication for an OSPFv3 Area
- Configuring IPsec Encryption for an OSPFv3 Area
- Displaying OSPFv3 IPsec Security Policies

Configuring IPsec Authentication on an Interface

Prerequisite: Before you enable IPsec authentication on an OSPFv3 interface, you must first enable IPv6 unicast routing globally, configure an IPv6 address and enable OSPFv3 on the interface, and assign it to an area (see Configuration Task List for OSPFv3 (OSPF for IPv6) on page 726).

To configure IPsec authentication on an interface, enter the following command:

Command Syntax	Command Mode	Usage
ipv6 ospf authentication {**null** | **ipsec spi** *number* {**MD5** | **SHA1**} [*key-encryption-type*] *key*}	INTERFACE	Enable IPsec authentication for OSPFv3 packets on an IPv6-based interface, where: **null causes an authentication policy configured for the area to not be inherited on the interface**. **ipsec spi** *number* is the Security Policy index (SPI) value. Range: 256 to 4294967295. **MD5** | **SHA1 specifies the authentication type: Message Digest 5 (MD5) or Secure Hash Algorithm 1 (SHA-1).** *key-encryption-type* **(optional) specifies if the key is encrypted. Valid values: 0 (key is not encrypted) or 7 (key is encrypted).** *key* **specifies the text string** used in authentication. All neighboring OSPFv3 routers must share the same key to exchange information. For MD5 authentication, the key must be 32 hex digits (non-encrypted) or 64 hex digits (encrypted). For SHA-1 authentication, the key must be 40 hex digits (non-encrypted) or 80 hex digits (encrypted).

An SPI value must be unique to one IPsec security policy (authentication or encryption) on the router. You must configure the same authentication policy (same SPI and key) on each OSPFv3 interface in a link.

To remove an IPsec authentication policy from an interface, enter the **no ipv6 ospf authentication ipsec spi** *number* command. To remove null authentication on an interface to allow the interface to inherit the authentication policy configured for the OSPFv3 area, enter the **no ipv6 ospf authentication null** command.

To display the configuration of IPsec authentication policies on the router, enter the **show crypto ipsec policy** command. To display the security associations set up for OSPFv3 interfaces in authentication policies, enter the **show crypto ipsec sa ipv6** command.

Configuring IPsec Encryption on an Interface

Prerequisite: Before you enable IPsec encryption on an OSPFv3 interface, you must first enable IPv6 unicast routing globally, configure an IPv6 address and enable OSPFv3 on the interface, and assign it to an area (see Configuration Task List for OSPFv3 (OSPF for IPv6) on page 726).

To configure IPsec encryption on an interface, enter the following command

Command Syntax	Command Mode	Usage
ipv6 ospf encryption {**null** | **ipsec spi** *number* **esp** *encryption-algorithm* [*key-encryption-type*] *key* *authentication-algorithm* [*key-authentication-type*] *key* }	INTERFACE	Enable IPsec encryption for OSPFv3 packets on an IPv6-based interface, where: **null causes an encryption policy configured for the area to not be inherited on the interface**. **ipsec spi** *number* is the Security Policy index (SPI) value. Range: 256 to 4294967295. **esp** *encryption-algorithm* **specifies the encryption algorithm used with ESP. Valid values are: 3DES, DES, AES-CBC, and NULL. For AES-CBC, only the AES-128 and AES-192 ciphers are supported.** *key* **specifies the text string** used in the encryption. All neighboring OSPFv3 routers must share the same key to decrypt information. Required lengths of a non-encrypted or encrypted key are: 3DES - 48 or 96 hex digits; DES - 16 or 32 hex digits; AES-CBC - 32 or 64 hex digits for AES-128 and 48 or 96 hex digits for AES-192. *key-encryption-type* **(optional) specifies if the key is encrypted. Valid values: 0 (key is not encrypted) or 7 (key is encrypted).** *authentication-algorithm* **specifies the encryption authentication algorithm to use. Valid values are MD5** or **SHA1.** *key* **specifies the text string** used in used in authentication. All neighboring OSPFv3 routers must share the same key to exchange information. For MD5 authentication, the key must be 32 hex digits (non-encrypted) or 64 hex digits (encrypted). For SHA-1 authentication, the key must be 40 hex digits (non-encrypted) or 80 hex digits (encrypted). *key-authentication-type* **(optional) specifies if the authentication key is encrypted. Valid values: 0 or 7**.

Note that when you configure encryption with the **ipv6 ospf encryption ipsec** command, you enable both IPsec encryption and authentication. However, when you enable authentication on an interface with the **ipv6 ospf authentication ipsec** command, you do not enable encryption at the same time.

An SPI value must be unique to one IPsec security policy (authentication or encryption) on the router. You must configure the same authentication policy (same SPI and key) on each OSPFv3 interface in a link.

To remove an IPsec encryption policy from an interface, enter the **no ipv6 ospf encryption ipsec spi** *number* command. To remove null encryption on an interface to allow the interface to inherit the encryption policy configured for the OSPFv3 area, enter the **no ipv6 ospf encryption null** command.

To display the configuration of IPsec encryption policies on the router, enter the **show crypto ipsec policy** command. To display the security associations set up for OSPFv3 interfaces in encryption policies, enter the **show crypto ipsec sa ipv6** command.

Configuring IPsec Authentication for an OSPFv3 Area

Prerequisite: Before you enable IPsec authentication on an OSPFv3 area, you must first enable OSPFv3 globally on the router (see Configuration Task List for OSPFv3 (OSPF for IPv6) on page 726).

To configure IPsec authentication for an OSPFv3 area, enter the following command in global configuration mode:

Command Syntax	Command Mode	Usage
area-*id* **authentication ipsec spi** *number* {**MD5** | **SHA1**} [*key-encryption-type*] *key*	CONF-IPV6-ROUTER-OSPF	Enable IPsec authentication for OSPFv3 packets in an area, where: **area** *area-id* specifies the area for which OSPFv3 traffic is to be authenticated. For *area-id*, you can enter a number or an IPv6 prefix. **spi** *number* is the Security Policy index (SPI) value. Range: 256 to 4294967295. **MD5** | **SHA1 specifies the authentication type: message digest 5 (MD5) or Secure Hash Algorithm 1 (SHA-1).** *key-encryption-type* **(optional) specifies if the key is encrypted. Valid values: 0 (key is not encrypted) or 7 (key is encrypted).** *key* **specifies the text string** used in authentication. All neighboring OSPFv3 routers must share the same key to exchange information. For MD5 authentication, the key must be 32 hex digits (non-encrypted) or 64 hex digits (encrypted). For SHA-1 authentication, the key must be 40 hex digits (non-encrypted) or 80 hex digits (encrypted).

An SPI value must be unique to one IPsec security policy (authentication or encryption) on the router. You must configure the same authentication policy (same SPI and key) on each interface in an OPSFv3 link.

If you have enabled IPsec encryption in an OSPFv3 area with the **area encryption** command, you cannot use the **area authentication** command in the area at the same time.

The configuration of IPsec authentication on an interface-level takes precedence over an area-level configuration. If you remove an interface configuration, an area authentication policy that has been configured is applied to the interface.

To remove an IPsec authentication policy from an OSPFv3 area, enter the **no area** *area-id* **authentication ipsec spi** *number* command.

To display the configuration of IPsec authentication policies on the router, enter the **show crypto ipsec policy** command.

Configuring IPsec Encryption for an OSPFv3 Area

Prerequisite: Before you enable IPsec encryption in an OSPFv3 area, you must first enable OSPFv3 globally on the router (see Configuration Task List for OSPFv3 (OSPF for IPv6) on page 726).

To configure IPsec encryption in an OSPFv3 area, enter the following command in global configuration mode:

Command Syntax	Command Mode	Usage
area *area-id* **encryption ipsec spi** *number* **esp** *encryption-algorithm* [*key-encryption-type*] *key* *authentication-algorithm* [*key-authentication-type*] *key*	CONF-IPV6-ROUTER-OSPF	Enable IPsec encryption for OSPFv3 packets in an area, where: **area** *area-id* specifies the area for which OSPFv3 traffic is to be encrypted. For *area-id*, you can enter a number or an IPv6 prefix. **spi** *number* is the Security Policy index (SPI) value. Range: 256 to 4294967295. **esp** *encryption-algorithm* **specifies the encryption algorithm used with ESP. Valid values are: 3DES, DES, AES-CBC, and NULL. For AES-CBC, only the AES-128 and AES-192 ciphers are supported.** *key* **specifies the text string** used in the encryption. All neighboring OSPFv3 routers must share the same key to decrypt information. Required lengths of a non-encrypted or encrypted key are: 3DES - 48 or 96 hex digits; DES - 16 or 32 hex digits; AES-CBC - 32 or 64 hex digits for AES-128 and 48 or 96 hex digits for AES-192. *key-encryption-type* **(optional) specifies if the key is encrypted. Valid values: 0 (key is not encrypted) or 7 (key is encrypted).** *authentication-algorithm* **specifies the authentication algorithm to use for encryption. Valid values are MD5** or **SHA1.** *key* **specifies the text string** used in authentication. All neighboring OSPFv3 routers must share the same key to exchange information. For MD5 authentication, the key must be 32 hex digits (non-encrypted) or 64 hex digits (encrypted). For SHA-1 authentication, the key must be 40 hex digits (non-encrypted) or 80 hex digits (encrypted). *key-authentication-type* **(optional) specifies if the authentication key is encrypted. Valid values: 0 or 7**.

An SPI value must be unique to one IPsec security policy (authentication or encryption) on the router. You must configure the same encryption policy (same SPI and keys) on each interface in an OPSFv3 link.

Note that when you configure encryption with the **area encryption** command, you enable both IPsec encryption and authentication. However, when you enable authentication on an area with the **area authentication** command, you do not enable encryption at the same time.

If you have enabled IPsec authentication in an OSPFv3 area with the **area authentication** command, you cannot use the **area encryption** command in the area at the same time.

The configuration of IPsec encryption on an interface-level takes precedence over an area-level configuration. If you remove an interface configuration, an area encryption policy that has been configured is applied to the interface.

To remove an IPsec encryption policy from an OSPFv3 area, enter the **no area** *area-id* **encryption ipsec spi** *number* command.

To display the configuration of IPsec encryption policies on the router, enter the **show crypto ipsec policy** command.

Displaying OSPFv3 IPsec Security Policies

To display the configuration of IPsec authentication and encryption policies, enter the following command:

Command Syntax	Command Mode	Usage
show crypto ipsec policy [**name** *name*]	EXEC Privilege	Display the AH and ESP parameters configured in IPsec security policies, including the SPI number, key, and algorithms used. *name* **displays configuration details about a specified policy**.

Figure 32-26. Command Example: show crypto ipsec policy

```
FTOS#show crypto ipsec policy

Crypto IPSec client security policy data

Policy name             : OSPFv3-1-502
Policy refcount         : 1
Inbound  ESP SPI        : 502 (0x1F6)
Outbound ESP SPI        : 502 (0x1F6)
Inbound  ESP Auth Key   : 123456789a123456789b123456789c12
Outbound ESP Auth Key   : 123456789a123456789b123456789c12
Inbound  ESP Cipher Key : 123456789a123456789b123456789c123456789d12345678
Outbound ESP Cipher Key : 123456789a123456789b123456789c123456789d12345678
Transform set           : esp-3des esp-md5-hmac


Crypto IPSec client security policy data

Policy name             : OSPFv3-1-500
Policy refcount         : 2
Inbound  AH SPI         : 500 (0x1F4)
Outbound AH SPI         : 500 (0x1F4)
Inbound  AH Key         : bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba8ed8bb5efe91e97e
Outbound AH Key         : bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba8ed8bb5efe91e97e
Transform set           : ah-md5-hmac


Crypto IPSec client security policy data

Policy name             : OSPFv3-0-501
Policy refcount         : 1
Inbound  ESP SPI        : 501 (0x1F5)
Outbound ESP SPI        : 501 (0x1F5)
Inbound  ESP Auth Key   :
bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba8ed8bb5efe91e97eb7c0c30808825fb5
Outbound ESP Auth Key   :
bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba8ed8bb5efe91e97eb7c0c30808825fb5
Inbound  ESP Cipher Key : bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba10345a1039ba8f8a
Outbound ESP Cipher Key : bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba10345a1039ba8f8a
Transform set           : esp-128-aes esp-sha1-hmac
```

In this encryption policy, the keys are not encrypted.

In this authentication policy, the keys are encrypted.

In this encryption policy, the keys are encrypted.

To display the IPsec security associations (SAs) used on OSPFv3 interfaces, enter the following command:

Command Syntax	Command Mode	Usage
show crypto ipsec sa ipv6 [**interface interface**]	EXEC Privilege	Displays security associations set up for OSPFv3 links in IPsec authentication and encryption policies on the router. To display information on the SAs used on a specific interface, enter **interface** *interface*, where *interface* is one of the following values: For a 1-Gigabit Ethernet interface, enter **GigabitEthernet** *slot*/*port*. For a Port Channel interface, enter **port-channel** *number*. Valid port-channel numbers (on an E-Series TeraScale): 1 to 255. For a 10-Gigabit Ethernet interface, enter **TenGigabitEthernet** *slot*/*port*. For a VLAN interface, enter **vlan** *vlan-id.* Valid VLAN IDs: 1 to 4094

Figure 32-27. Command Example: show crypto ipsec sa ipv6

```
FTOS#show crypto ipsec sa ipv6

Interface: TenGigabitEthernet 0/0
  Link Local address: fe80::201:e8ff:fe40:4d10
  IPSecv6 policy name: OSPFv3-1-500

  inbound ah sas
   spi : 500 (0x1f4)
    transform : ah-md5-hmac
    in use settings : {Transport, }
    replay detection support : N
    STATUS : ACTIVE

  outbound ah sas
   spi : 500 (0x1f4)
    transform : ah-md5-hmac
    in use settings : {Transport, }
    replay detection support : N
    STATUS : ACTIVE

  inbound esp sas

  outbound esp sas


Interface: TenGigabitEthernet 0/1
  Link Local address: fe80::201:e8ff:fe40:4d11
  IPSecv6 policy name: OSPFv3-1-600

  inbound ah sas

  outbound ah sas

  inbound esp sas
   spi : 600 (0x258)
    transform : esp-des esp-sha1-hmac
    in use settings : {Transport, }
    replay detection support : N
    STATUS : ACTIVE

  outbound esp sas
   spi : 600 (0x258)
    transform : esp-des esp-sha1-hmac
    in use settings : {Transport, }
    replay detection support : N
    STATUS : ACTIVE
```

Troubleshooting OSPFv3

FTOS has several tools to make troubleshooting easier. Be sure to check the following, as these are typical issues that interrupt the OSPFv3 process. Note that this is not a comprehensive list, just some examples of typical troubleshooting checks.

- Has OSPF been enabled globally?
- Is the OSPF process active on the interface?
- Are adjacencies established correctly?
- Are the interfaces configured for Layer 3 correctly?
- Is the router in the correct area type?
- Have the routes been included in the OSPF database?
- Have the OSPF routes been included in the routing table (not just the OSPF database)?

Some useful troubleshooting commands are:

- show ipv6 interfaces
- show ipv6 protocols
- debug IPv6 OSPF events and/or packets
- show ipv6 neighbors
- show virtual links
- show ipv6 routes

Use the following commands in EXEC Privilege mode to get general route and links status information.

Command Syntax	Command Mode	Usage
show ipv6 route summary	EXEC Privilege	View the summary information of the IPv6 routes
show ipv6 ospf database	EXEC Privilege	View the summary information for the OSPFv3 database

Use the following command in EXEC Privilege mode to view the OSPF configuration for a neighboring router:

Command Syntax	Command Mode	Usage
show ipv6 ospf neighbor	EXEC Privilege	View the configuration of OSPFv3 neighbors.

Use the following command in EXEC Privilege mode to configure the debugging options of an OSPFv3 process:

Command Syntax	Command Mode	Usage
debug ipv6 ospf [**event** | **packet**] {type slot/port}	EXEC Privilege	View debug messages for all OSPFv3 interfaces. • **event**: View OSPF event messages. • **packet**: View OSPF packets. • For a Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information (**e.g. passive-interface gi 2/1**). • For a port channel, enter the keyword **port-channel** followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale (**e.g. passive-interface po 100**) • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information (**e.g. passive-interface ten 2/3**). • For a VLAN, enter the keyword **vlan** followed by a number from 1 to 4094 (**e.g. passive-interface vlan 2222)**. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

33

PIM Dense-Mode

PIM Dense-Mode is supported on platforms: C E S

PIM-Dense Mode (PIM-DM) is a multicast protocol that directs routers to forward multicast traffic to all subnets until the router receives a request to stop; this behavior is the opposite of PIM-Sparse Mode, which does not forward multicast traffic to a subnet until the traffic is specifically requested using a PIM Join message.

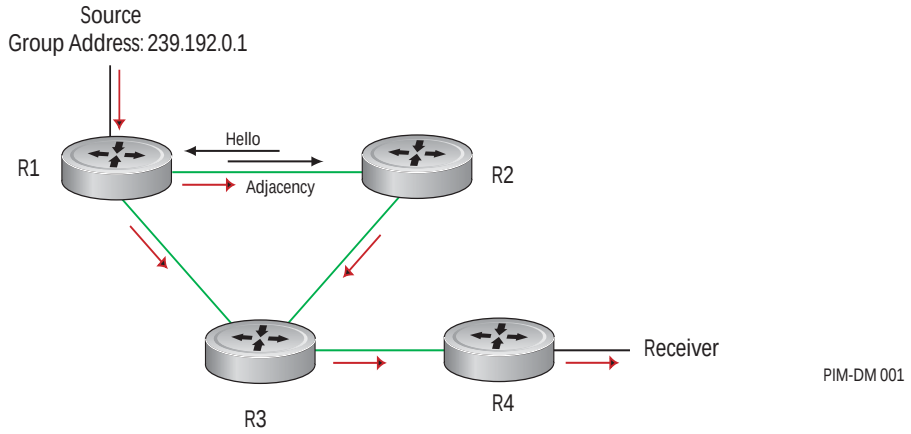
Implementation Information

- E-Series supports a maximum of 511 PIM interfaces and 50K multicast entries including (*,G), (S,G), and (S,G,rpt) entries. There is no limit on the number of PIM neighbors E-Series can have.
- FTOS reduces the number of control messages sent between multicast routers by bundling Join and Prune requests in the same message.
- FTOS supports PIM-DM on physical, VLAN, and port-channel interfaces.

Protocol Overview

PIM-DM routers form adjacencies with their neighbors by sending periodic hello messages to the all-PIM-routers address 224.0.0.13 out of all PIM-DM-enabled interfaces. By default, PIM-DM routers assume that every subnet has at least one receiver. When a router receives traffic from a particular source and for a particular group, it creates a (S,G) entry and lists all interfaces directly connected to a PIM-DM neighbor as an outgoing interface, thus recreating a unique distribution tree called a shortest path tree (SPT) to the source that includes all subnets.

Figure 33-1. Multicast Flooding in a PIM-DM Network



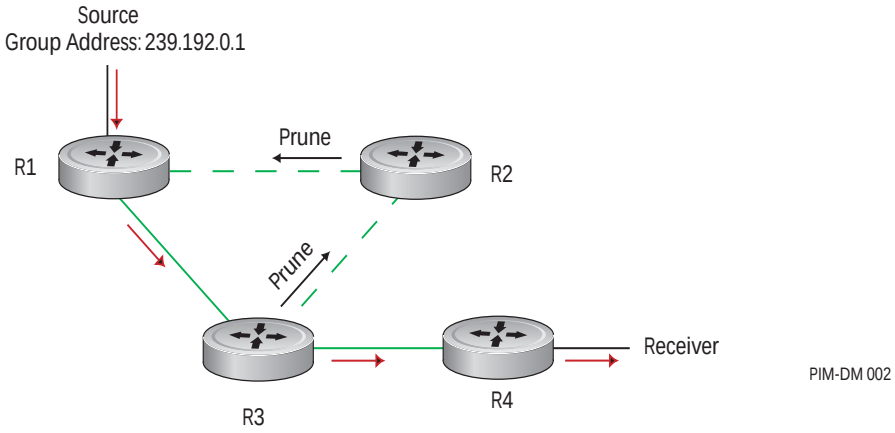
Refusing Multicast Traffic

If a PIM-DM router has no receivers for a group, it refuses multicast traffic by sending a PIM Prune message to address 224.0.0.13 out of the source interface. The upstream neighbor receives the prune message and determines if it has any remaining neighbors downstream. If it does not, it propagates the prune message upstream out of the source interface. Likewise, all remaining routers between the receiver and the source independently determine whether to propagate the prune message until no router receives unwanted traffic for the (S,G).

Any router that receives multicast traffic on a port that does not lead back to the source (via the PIM-DM selected path) also generates a prune message.

In Figure 33-1, R3 receives multicast traffic by two paths. In Figure 33-2, PIM-DM selects only one path for the reverse path forwarding (RFP) check and generates a prune message so that routers upstream stop sending traffic for the group. R2 then has no PIM-DM neighbors downstream and so sends a prune message to R1.

Figure 33-2. Refusing Multicast Traffic in a PIM-DM Network

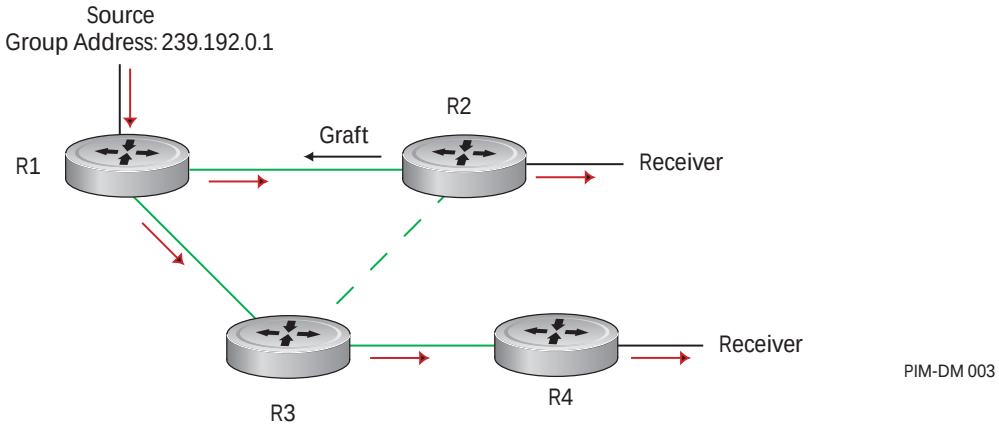


When a router receives a prune message, it flags the relevant (S,G) entry and sets a timer. If the timer expires, it begins flooding traffic out of the interface, and downstream routers must again evaluate whether to prune itself from the tree. To prevent the timer from expiring, while the source is sending traffic for the (S,G), the first-hop router periodically sends an (S,G) state-refresh messages down the entire SPT. Router that set the prune flag for the (S,G) entry reset the timer when they receive the message.

Requesting Multicast Traffic

When a new receiver joins a multicast group it sends an IGMP Membership Report to its gateway router. The gateway router sends a PIM Graft message to its upstream neighbor, which sets a forwarding flag, and propagates the graft message upstream, as shown in Figure 33-3. All remaining routers between the receiver and the source also set a forwarding flag and propagate the graft message so that the receiver begins receiving traffic for the (S,G).

Figure 33-3. Requesting Multicast Traffic in a PIM-DM Network



Configure PIM-DM

Configuring PIM-DM is a two-step process:

1. Enable multicast routing using the command **ip multicast-routing** from CONFIGURATION mode.
2. Enable PIM-DM on an interface. See page 750.

Related Configuration Tasks

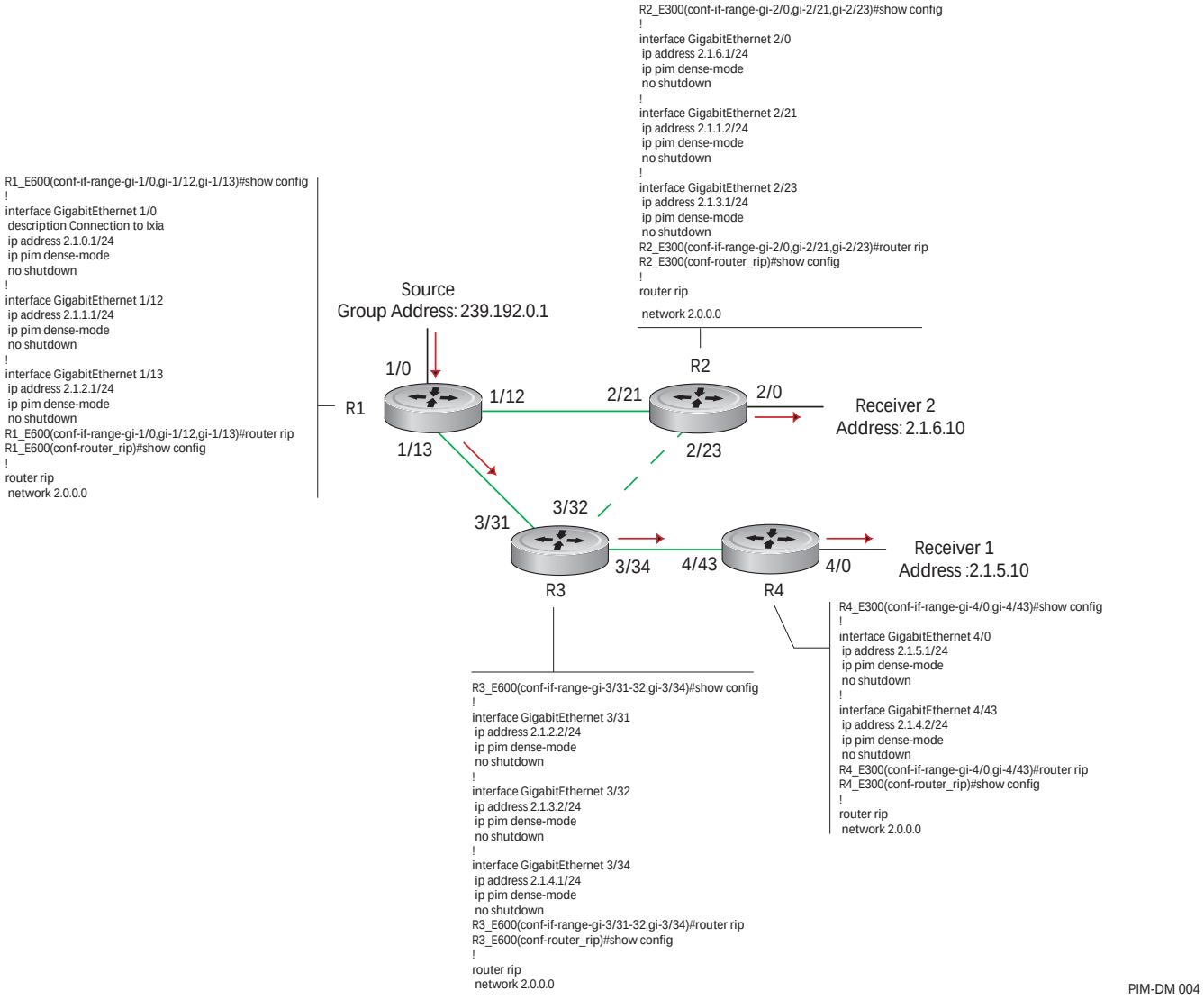
- Clear the PIM TIB using the command **clear ip pim tib** from EXEC Privilege mode.
- Debug PIM-DM by displaying control activity (packets, events, timers, etc.) using the command **debug ip pim** from EXEC Privilege mode.

Enable PIM-DM

To enable PIM-DM:

Step	Task	Command	Command Mode
1	Enable multicast routing on the system.	**ip multicast-routing**	CONFIGURATION
2	Enable PIM-Dense Mode on each interface that will participate in PIM-DM, as shown in Figure 33-4.	**ip pim dense-mode**	INTERFACE

Figure 33-4. Enabling PIM-DM



Display which interfaces are enabled with PIM-DM using the command **show ip pim interface** from EXEC Privilege mode, as shown in Figure 33-5.

Figure 33-5. Viewing PIM-SM Enabled Interfaces

```
R1_E600(conf)#do show ip pim interface
Address          Interface Ver/   Nbr    Query  DR         DR
                           Mode   Count  Intvl  Prio
2.1.0.1          Gi 1/0    v2/D   0      30     1          2.1.0.1
2.1.1.1          Gi 1/12   v2/D   1      30     1          2.1.1.1
2.1.2.1          Gi 1/13   v2/D   1      30     1          2.1.2.1
R1_E600(conf)#
```

Display PIM neighbors for each interface using the command **show ip pim neighbor** from EXEC Privilege mode, as shown in Figure 33-6.

Figure 33-6. Viewing PIM Neighbors Command Example

```
R1_E600(conf)#do show ip pim neighbor
Neighbor                 Interface      Uptime/Expires     Ver  DR Prio/Mode GR
Address
2.1.1.2                  Gi 1/12        01:43:51/00:01:35  v2   0  / D
2.1.2.2                  Gi 1/13        02:00:46/00:01:41  v2   0  / D
R1_E600(conf)#
```

Display the PIM routing table using the command **show ip pim tib** from EXEC privilege mode, as shown in Figure 33-7.

Figure 33-7. Viewing the PIM Multicast Routing Table

```
------------------------------------ Router 1 ------------------------------------------
R1_E600(conf)#do show ip pim tib

PIM Multicast Routing Table
Flags: D - Dense, S - Sparse, C - Connected, L - Local, P - Pruned,
       R - RP-bit set, F - Register flag, T - SPT-bit set, J - Join SPT,
       M - MSDP created entry, A - Candidate for MSDP Advertisement
       K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(*, 239.192.0.1), uptime 00:57:57, expires 00:00:00, flags: D
  Incoming interface: Null, RPF neighbor 0.0.0.0
  Outgoing interface list:

(2.1.0.10, 239.192.0.1), uptime 00:02:22, expires 00:01:17, flags: F
  Incoming interface: GigabitEthernet 1/0, RPF neighbor 0.0.0.0
  Outgoing interface list:
    GigabitEthernet 1/13  Forward/Sparse   2w4d/Never
    GigabitEthernet 1/12  Prune/Sparse   2w4d/00:01:13
------------------------------------ Router 2 ------------------------------------------
[output omitted]
(*, 239.192.0.1), uptime 00:05:23, expires 00:00:00, flags: D
  Incoming interface: Null, RPF neighbor 0.0.0.0
  Outgoing interface list:
    GigabitEthernet 2/0  Forward/Dense   00:00:03/Never
    GigabitEthernet 2/21  Forward/Dense   1d1h/Never
    GigabitEthernet 2/23  Forward/Dense   1d1h/Never

(2.1.0.10, 239.192.0.1), uptime 00:05:23, expires 00:00:00, flags:
  Incoming interface: GigabitEthernet 2/21, RPF neighbor 2.1.1.1
  Outgoing interface list:
    GigabitEthernet 2/0  Forward/Sparse   00:00:03/Never
------------------------------------ Router 3 ------------------------------------------
[output omitted]
(*, 239.192.0.1), uptime 00:05:06, expires 00:00:00, flags: D
  Incoming interface: Null, RPF neighbor 0.0.0.0
  Outgoing interface list:
    GigabitEthernet 3/31  Forward/Dense   6d0h/Never
    GigabitEthernet 3/32  Forward/Dense   6d0h/Never
    GigabitEthernet 3/34  Forward/Dense   6d0h/Never

(2.1.0.10, 239.192.0.1), uptime 00:05:06, expires 00:02:19, flags:
  Incoming interface: GigabitEthernet 3/31, RPF neighbor 2.1.2.1
  Outgoing interface list:
    GigabitEthernet 3/34  Forward/Sparse   6d0h/Never
    GigabitEthernet 3/32  Prune/Sparse   6d0h/7101w3d
------------------------------------ Router 4 ------------------------------------------
[output omitted]

(*, 239.192.0.1), uptime 00:02:36, expires 00:00:00, flags: D
  Incoming interface: Null, RPF neighbor 0.0.0.0
  Outgoing interface list:
    GigabitEthernet 4/0  Forward/Dense   00:00:34/Never
    GigabitEthernet 4/43  Forward/Dense   6d0h/Never

(2.1.0.10, 239.192.0.1), uptime 00:02:36, expires 00:03:24, flags:
  Incoming interface: GigabitEthernet 4/43, RPF neighbor 2.1.4.1
  Outgoing interface list:
    GigabitEthernet 4/0  Forward/Sparse   00:00:34/Never
```

34

PIM Sparse-Mode

PIM Sparse-Mode is supported on platforms: C E S

PIM-SM is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

PIM-Sparse Mode (PIM-SM) is a multicast protocol that forwards multicast traffic to a subnet only upon request using a PIM Join message; this behavior is the opposite of PIM-Dense Mode, which forwards multicast traffic to all subnets until it receives a request to stop.

Implementation Information

- The Dell Force10 implementation of PIM-SM is based on the IETF *Internet Draft draft-ietf-pim-sm-v2-new-05*.
- C-Series supports a maximum of 31 PIM interfaces and 4K multicast entries including (*,G), and (S,G) entries. There is no limit on the number of PIM neighbors C-Series can have.
- S-Series supports a maximum of 31 PIM interfaces and 2K multicast entries including (*,G), and (S,G) entries. There is no limit on the number of PIM neighbors S-Series can have.
- E-Series supports a maximum of 511 PIM interfaces and 50K multicast entries including (*,G), (S,G), and (S,G,rpt) entries. There is no limit on the number of PIM neighbors E-Series can have.
- The SPT-Threshold is zero, which means that the last-hop designated router (DR) joins the shortest path tree (SPT) to the source upon receiving the first multicast packet.
- FTOS reduces the number of control messages sent between multicast routers by bundling Join and Prune requests in the same message.
- FTOS supports PIM-SM on physical, VLAN, and port-channel interfaces.
- FTOS supports 2000 IPv6 multicast forwarding entries, with up to 128 PIM-SSM neighbors/interfaces.
- On VLAN interfaces, PIM-SM is supported on C-Series, E-Series, and S-Series platforms.
- IPv6 Multicast is not supported on SONET interfaces.

Protocol Overview

To distribute the same traffic to multiple receivers, PIM-SM creates a tree extending from a root, called the Rendezvous Point (RP), down branches that extend to the nodes which have requested the traffic. Nodes requesting the same traffic belong to the same *multicast group*.

Initially, a single PIM-SM tree called a *shared tree* to distribute traffic. It is called *shared* because all traffic for the group, regardless of the source, or the location of the source, must pass through the RP. The shared tree is unidirectional; that is, all multicast traffic flows only from the RP to the receivers. Once a receiver receives traffic from the RP, PM-SM switches to shortest path trees (SPT) to forward multicast traffic, which connects the receiver directly to the source. Each multicast group has an RP and a unidirectional shared tree (group-specific shared tree).

Requesting Multicast Traffic

A host requesting multicast traffic for a particular group sends an IGMP Join message to its gateway router. The gateway router is then responsible for joining the shared tree to the RP (RPT) so that the host can receive the requested traffic.

1. Upon receiving an IGMP Join message, the receiver gateway router (last-hop DR) creates a (*,G) entry in its multicast routing table for the requested group. The interface on which the join message was received becomes the outgoing interface associated with the (*,G) entry.
2. The last-hop DR sends a PIM Join message to the RP. All routers along the way, including the RP, create an (*,G) entry in their multicast routing table, and the interface on which the message was received becomes the outgoing interface associated with the (*,G) entry. This process constructs an RPT branch to the RP.
3. If a host on the same subnet as another multicast receiver sends an IGMP report for the same multicast group, the gateway takes no action. If a router between the host and the RP receives a PIM Join message for which it already has a (*,G) entry, the interface on which the message was received is added to the outgoing interface list associated with the (*,G) entry, and the message is not (and does not need to be) forwarded towards the RP.

Refusing Multicast Traffic

A host requesting to leave a multicast group sends an IGMP Leave message to the last-hop DR. If the host is the only remaining receiver for that group on the subnet, the last-hop DR is responsible for sending a PIM Prune message up the RPT to prune its branch to the RP.

1. Upon receiving an IGMP Leave message, the gateway removes the interface on which it is received from the outgoing interface list of the (*,G) entry. If the (*,G) entry has no remaining outgoing interfaces, multicast traffic for that group is no longer forwarded to that subnet.
2. If the (*,G) entry has no remaining outgoing interfaces, the last-hop DR sends a PIM Prune message to towards the RP. All routers along the way remove the interface on which the message was received from the outgoing interface list of the (*,G) entry. If on any router there is at least one outgoing interface listed for that (*,G) entry, the Prune message is not forwarded.

Sending Multicast Traffic

With PIM-SM, all multicast traffic must initially originate from the RP. A source must unicast traffic to the RP so that the RP can learn about the source and create an SPT to it. Then the last-hop DR may create an SPT directly to the source.

1. The source gateway router (first-hop DR) receives the multicast packets and creates an (S,G) entry in its multicast routing table. The first-hop DR encapsulates the initial multicast packets in PIM Register packets and unicasts them to the RP.
2. The RP decapsulates the PIM Register packets and forwards them if there are any receivers for that group. The RP sends a PIM Join message towards the source. All routers between the RP and the source, including the RP, create an (S,G) entry and list the interface on which the message was received as an outgoing interface, thus recreating a SPT to the source.
3. Once the RP starts receiving multicast traffic via the (S,G) it unicasts a Register-Stop message to the first-hop DR so that multicast packets are no longer encapsulated in PIM Register packets and unicast. Upon receiving the first multicast packet from a particular source, the last-hop DR sends a PIM Join message to the source to create an SPT to it.
4. There are two paths, then, between the receiver and the source, a direct SPT and an RPT. One router will receive a multicast packet on two interfaces from the same source in this case; this router prunes the shared tree by sending a PIM Prune message to the RP that tells all routers between the source and the RP to remove the outgoing interface from the (*,G) entry, and tells the RP to prune its SPT to the source with a Prune message.

FTOS Behavior: When the router creates an SPT to the source, there are then two paths between the receiver and the source, the SPT and the RPT. Until the router can prune itself from the RPT, the receiver receives duplicate multicast packets which may cause disruption. Therefore, the router must prune itself from the RPT as soon as possible. FTOS optimizes the shared to shortest-path tree switchover latency by copying and forwarding the first (S,G) packet received on the SPT to the PIM task immediately upon arrival. The arrival of the (S,G) packet confirms for PIM that the SPT is created, and that it can prune itself from the shared tree.

Important Points to Remember

- If a loopback interface with a /32 mask is used as the RP, you must enable PIM Sparse-mode on the interface.

Configure PIM-SM

Configuring PIM-SM is a two-step process:

1. Enable IPv4 or IPv6 multicast routing using the command [**ip | ipv6**] **multicast-routing** from CONFIGURATION mode.
2. Select a Rendezvous Point, or let PIM elect an RP. See page 760.

3. Enable PIM-SM on an interface. See page 758.

Related Configuration Tasks

- Configurable S,G Expiry Timers on page 759
- Configure a Static Rendezvous Point on page 760
- Elect an RP using the BSR Mechanism on page 762
- Configure a Designated Router on page 763
- Create Multicast Boundaries and Domains on page 763
- Set a Threshold for Switching to the SPT on page 764
- PIM-SM Graceful Restart on page 764
- First Packet Forwarding for Lossless Multicast on page 765
- Monitoring PIM on page 766

Enable PIM-SM

You must enable PIM-SM on each participating interface:

IP Version	Task	Command	Command Mode
IPv4	Enable PIM-Sparse Mode on an interface.	**ip pim sparse-mode**	INTERFACE
IPv6	Enable PIM-Sparse Mode on an interface.	**ipv6 pim sparse-mode**	INTERFACE

Display which interfaces are enabled with PIM-SM using the command **show** [**ip | ipv6**] **pim interface** from EXEC Privilege mode, as shown in Figure 34-1.

Figure 34-1. Viewing PIM-SM Enabled Interfaces

```
FTOS#show ip pim interface
Address          Interface VIFindex Ver/   Nbr    Query  DR    DR
                                    Mode   Count  Intvl  Prio
189.87.5.6       Gi 4/11   0x2      v2/S   1      30     1     127.87.5.6
189.87.3.2       Gi 4/12   0x3      v2/S   1      30     1     127.87.3.5
189.87.31.6      Gi 7/11   0x0      v2/S   0      30     1     127.87.31.6
189.87.50.6      Gi 7/13   0x4      v2/S   1      30     1     127.87.50.6
FTOS#
```

Note: You can influence the selection of the Rendezvous Point by enabling PIM-Sparse Mode on a loopback interface and assigning a low IP address.

Display PIM neighbors for each interface using the command **show** [**ip | ipv6**] **pim neighbor** from EXEC Privilege mode, as shown in Figure 34-2.

Figure 34-2. Viewing PIM Neighbors Command Example

```
FTOS#show ip pim neighbor
Neighbor          Interface      Uptime/Expires     Ver  DR
Address                                                 Prio/Mode
127.87.5.5        Gi 4/11        01:44:59/00:01:16  v2   1  / S
127.87.3.5        Gi 4/12        01:45:00/00:01:16  v2   1  / DR
127.87.50.5       Gi 7/13        00:03:08/00:01:37  v2   1  / S
FTOS#
```

Display the PIM routing table using the command **show** [**ip | ipv6**] **pim tib** from EXEC privilege mode, as shown in Figure 34-3.

Figure 34-3. Viewing the PIM Multicast Routing Table

```
FTOS#show ip pim tib

PIM Multicast Routing Table
Flags: D - Dense, S - Sparse, C - Connected, L - Local, P - Pruned,
       R - RP-bit set, F - Register flag, T - SPT-bit set, J - Join SPT,
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(*, 192.1.2.1), uptime 00:29:36, expires 00:03:26, RP 10.87.2.6, flags: SCJ
  Incoming interface: GigabitEthernet 4/12, RPF neighbor 10.87.3.5
  Outgoing interface list:
    GigabitEthernet 4/11
    GigabitEthernet 7/13

(10.87.31.5, 192.1.2.1), uptime 00:01:24, expires 00:02:26, flags: FT
  Incoming interface: GigabitEthernet 7/11, RPF neighbor 0.0.0.0
  Outgoing interface list:
    GigabitEthernet 4/11
    GigabitEthernet 4/12
    GigabitEthernet 7/13
--More--
```

Configurable S,G Expiry Timers

By default S, G entries expire in 210 seconds. You can configure a global expiry time (for all (S,G) entries) or configure a expiry time for a particular entry. If both are configured, the ACL supercedes the global configuration for the specified entries.

When an expiry time created, deleted, or updated, the changes are applied when the keep alive timer refreshes.

To configure a global expiry time:

Task	Command	Command Mode
Enable global expiry timer for S, G entries Range 211-86400 seconds Default: 210	**ip pim sparse-mode sg-expiry-timer** *seconds*	CONFIGURATION

Configure the expiry time for a particular (S,G) entry:

Step	Task	Command Syntax	Command Mode
1	Create an Extended ACL	**ip access-list extended** *access-list-name*	CONFIGURATION
2	Specify the source and group to which the timer will be applied using extended ACLs with permit rules only.	[**seq** *sequence-number*] **permit ip** *source-address/mask* **| any** | **host** *source-address*} {*destination-address/mask* | **any** | **host** *destination-address*}	CONFIG-EXT-NACL
3	Set the expiry time for a specific (S,G) entry (Figure 34-4). Range 211-86400 seconds Default: 210	**ip pim sparse-mode sg-expiry-timer** *seconds* **sg-list** *access-list-name*	CONFIGURATION

Note: The expiry time configuration is nullified, and the default global expiry time is used if:

- an ACL is specified for an in the **ip pim sparse-mode sg-expiry-timer** command, but the ACL has not been created or is a standard ACL.
- if the expiry time is specified for an (S,G) entry in a deny rule.

Figure 34-4. Configuring an (S,G) Expiry Time

```
FTOS(conf)#ip access-list extended SGtimer
FTOS(config-ext-nacl)#permit ip 10.1.2.3/24 225.1.1.0/24
FTOS(config-ext-nacl)#permit ip any 232.1.1.0/24
FTOS(config-ext-nacl)#permit ip 100.1.1.0/16 any
FTOS(config-ext-nacl)#show conf
!
ip access-list extended SGtimer
 seq 5 permit ip 10.1.2.0/24 225.1.1.0/24
 seq 10 permit ip any 232.1.1.0/24
 seq 15 permit ip 100.1.0.0/16 any
FTOS(config-ext-nacl)#exit

FTOS(conf)#ip pim sparse-mode sg-expiry-timer 1800 sg-list SGtimer
```

Display the expiry time configuration using the **show running-configuration** [**acl | pim**] command from EXEC Privilege mode.

Configure a Static Rendezvous Point

The rendezvous point is a PIM-enabled interface on a router that acts as the root a group-specific tree; every group must have an RP.

Identify an RP by the IP address of a PIM-enabled or loopback interface using the command **ip pim rp-address**, as shown in Figure 34-5.

Figure 34-5. Electing a Rendezvous Point

```
FTOS#sh run int loop0
!
interface Loopback 0
 ip address 1.1.1.1/32
 ip pim sparse-mode
 no shutdown
FTOS#sh run pim
!
ip pim rp-address 1.1.1.1 group-address 224.0.0.0/4
```

Override Bootstrap Router Updates

PIM-SM routers need to know the address of the RP for each group for which they have (*,G) entry. This address is obtained automatically through the bootstrap router (BSR) mechanism or a static RP configuration.

If you have configured a static RP for a group, use the option **override** with the command [**ip | ipv6**] **pim rp-address** to override bootstrap router updates with your static RP configuration. If you do not use this option, the RPs advertised in the BSR updates take precedence over any statically configured RPs.

Display the assigned RP for a group using the command **show** [**ip | ipv6**] **pim rp** from EXEC privilege mode, as shown in Figure 34-6.

Figure 34-6. Displaying the Rendezvous Point for a Multicast Group

```
FTOS#show ip pim rp
Group            RP
225.0.1.40       165.87.50.5
226.1.1.1        165.87.50.5
```

Display the assigned RP for a group range (group-to-RP mapping) using the command **show ip pim rp mapping** command in EXEC privilege mode

Figure 34-7. Display the Rendezvous Point for a Multicast Group Range

```
FTOS#show ip pim rp mapping
PIM Group-to-RP Mappings
Group(s): 224.0.0.0/4, Static
  RP: 165.87.50.5, v2
```

IP Version	Task	Command Syntax	Command Mode
IPv4	Override bootstrap router RP election results with your static RP configuration.	**ip pim rp-address**	CONFIGURATION
IPv4	Display the assigned RP for a group.	**show ip pim rp**	EXEC Privilege
IPv4	Display the assigned RP for a group range (group-to-RP mapping).	**show ip pim rp mapping**	EXEC Privilege

IP Version	Task	Command Syntax	Command Mode
IPv6	Override bootstrap router RP election results with your static RP configuration.	**ipv6 pim rp-address**	CONFIGURATION
IPv6	Display the assigned RP for a group.	**show ipv6 pim rp**	EXEC Privilege
IPv6	Display the assigned RP for a group range (group-to-RP mapping).	**show ipv6 pim rp mapping**	EXEC Privilege

Elect an RP using the BSR Mechanism

Every PIM router within a domain must map a particular multicast group address to the same RP. The group-to-RP mapping may be statically or dynamically configured. RFC 5059 specifies a dynamic, self-configuring method called the Bootstrap Router (BSR) mechanism, by which an RP is elected from a pool of RP candidates (C-RPs).

Some routers within the domain are configured to be C-RPs. Other routers are configured to be Bootstrap Router candidates (C-BSRs); one router is elected the BSR for the domain and is responsible for forwarding Bootstrap messages (BSM) containing the results of the RP election to the other routers in the domain.

The RP election process is as follows:

1. C-BSRs flood their candidacy throughout the domain in a BSM. Each message contains a BSR priority value, and the C-BSR with the highest priority value becomes the BSR.
2. Each C-RP unicasts periodic Candidate-RP-Advertisements to the BSR. Each message contains an RP priority value and the group ranges for which it is a C-RP.
3. The BSR determines the most efficient and stable group-to-RP mappings, which is called the *RP-Set*.
4. The BSR floods the RP-Set throughout the domain periodically in case new C-RPs are announced, or an RP failure occurs.

IP Version	Task	Command Syntax	Command Mode
IPv4	Make a PIM router a BSR candidate.	**ip pim bsr-candidate**	CONFIGURATION
	Make a PIM router a RP candidate.	**ip pim rp-candidate**	CONFIGURATION
	Display Bootstrap Router information.	**show ip pim bsr-router**	EXEC Privilege
IPv6	Make a PIM router a BSR candidate.	**ipv6 pim bsr-candidate**	CONFIGURATION
	Make a PIM router a RP candidate.	**ipv6 pim rp-candidate**	CONFIGURATION
	Display Bootstrap Router information.	**show ipv6 pim bsr-router**	EXEC Privilege

Configure a Designated Router

Multiple PIM-SM routers might be connected to a single LAN segment. One of these routers is elected to act on behalf of directly connected hosts. This router is the Designated Router (DR).

The DR is elected using hello messages. Each PIM router learns about its neighbors by periodically sending a hello message out of each PIM-enabled interface. Hello messages contain the IP address of the interface out of which it is sent and a DR priority value. The router with the greatest priority value is the DR. If the priority value is the same for two routers, then the router with the greatest IP address is the DR. By default the DR priority value is 192, so the IP address determines the DR.

IP Version	Task	Command Syntax	Command Mode
IPv4	Assign a DR priority value.	**ip pim dr-priority** *value*	INTERFACE
IPv4	Change the interval at which a router sends hello messages.	**ip pim query-interval** *seconds*	INTERFACE
IPv4	Display the current value of DR parameters.	**show ip pim interface**	EXEC Privilege
IPv6	Assign a DR priority value.	**ipv6 pim dr-priority** *value*	INTERFACE
IPv6	Change the interval at which a router sends hello messages.	**ipv6 pim query-interval** *seconds*	INTERFACE
IPv6	Display the current value of DR parameters.	**show ipv6 pim interface**	EXEC Privilege

Create Multicast Boundaries and Domains

A PIM domain is a contiguous set of routers that all implement PIM and are configured to operate within a common boundary defined by PIM Multicast Border Routers (PMBRs). PMBRs connect each PIM domain to the rest of the internet.

Create multicast boundaries and domains by filtering inbound and outbound Bootstrap Router (BSR) messages per interface, using the [**ip | ipv6**] **pim bsr-border** command. This command is applied to the subsequent inbound and outbound updates. Already existing BSR advertisements are removed by timeout.

Remove candidate RP advertisements using the **clear** [**ip | ipv6**] **pim rp-mapping** command.

IP Version	Task	Command Syntax	Command Mode
IPv4	Filter inbound and outbound Bootstrap Router messages per interface.	**ip pim bsr-border**	INTERFACE
	Remove candidate RP advertisements.	**clear ip pim rp-mapping**	EXEC PRIVILEGE

IP Version	Task	Command Syntax	Command Mode
IPv6	Filter inbound and outbound Bootstrap Router messages per interface.	**ipv6 pim bsr-border**	INTERFACE
	Remove candidate RP advertisements.	**clear ip pim rp-mapping**	EXEC PRIVILEGE

Set a Threshold for Switching to the SPT

Set a Threshold for Switching to the SPT is available only on platform: E

Initially, a single PIM-SM tree called a shared tree to distribute traffic. It is called shared because all traffic for the group, regardless of the source, or the location of the source, must pass through the RP. The shared tree is unidirectional; that is, all multicast traffic flows only from the RP to the receivers. Once a receiver receives traffic from the RP, PM-SM switches to *shortest path trees* (SPT) to forward multicast traffic, which connects the receiver directly to the source.

You can configure PIM to switch over to the SPT when the router receives multicast packets at or beyond a specified rate.

IP Version	Task	Command Syntax	Command Mode
IPv4	Configure PIM to switch over to the SPT when the multicast packet rate is at or beyond a specified rate. The keyword infinity directs PIM to never switch to the SPT.	**ip pim spt-threshold** {**value	infinity**} Default: 10 kbps	CONFIGURATION
IPv6	Configure PIM to switch over to the SPT when the multicast packet rate is at or beyond a specified rate. The keyword infinity directs PIM to never switch to the SPT.	**ip pim spt-threshold** {**value	infinity**} Default: 10 kbps	CONFIGURATION

PIM-SM Graceful Restart

PIM-SM Graceful Restart is supported only on platform E

PIM-SM Graceful Restart is supported only on platform E(X) with FTOS 8.2.1.0 and later.

When a PIM neighbor restarts and the liveliness timer for that neighbor expires, the join/prune states received from the neighbor expire, and the corresponding interfaces are removed from the outgoing list of multicast entries. The effect of this is that active multicast sessions are brought down.

FTOS supports PIM-SM graceful restart based on the GenID. Per RFC 4601, hello messages should contain a Generation_Identifier option, which contains a randomly generated value (GenID) that is regenerated each time PIM forwarding is started or restarted on the interface, including when the router restarts. When a router receives from a neighbor a hello message with a new GenID, any old hello information about that neighbor should be discarded and superseded by the information from the new hello message.

FTOS supports graceful restart based on the GenID. A Dell Force10 PIM router announces its graceful restart capability to its neighbors up front as an option in its hello messages.

If a graceful-restart capable router recognizes that a graceful-restart capable neighbor has restarted, it preserves the state from the neighbor and continues forwarding multicast traffic while the neighbor restarts.

- The router holds on to the entries learned from the neighbor for the graceful restart interval. If it does not receive a hello from the neighbor within this time, it purges all state associated with the neighbor.
- If the neighbor restarts and sends a hello with a new GenID before this interval expires, the router sends a join message towards the neighbor for the relevant entries.

If a graceful-restart capable router restarts, the router preserves all multicast entries in hardware until it receives and consolidates joins from its graceful-restart capable neighbors. The router is not taken off the forwarding path during restart.

Enable PIM-SM graceful restart (non-stop forwarding capability) using the command **ip pim graceful-restart nsf** from CONFIGURATION mode. There are two options with this command:

- **restart-time** is the time required by the Dell Force10 system to restart. The default value is 180 seconds.
- **stale-entry-time** is the maximum amount of time that the Dell Force10 system preserves entries from a restarting neighbor. The default value is 60 seconds.

In helper-only mode, the system preserves the PIM states of a neighboring router while the neighbor gracefully restarts, but the Dell Force10 system allows itself to be taken off the forwarding path if it restarts. Enable this mode using the command **ip pim graceful-restart helper-only**. This mode takes precedence over any graceful restart configuration.

First Packet Forwarding for Lossless Multicast

When the Dell Force10 system is the RP, packets arriving before an (S,G) entry is created are soft forwarded using the (*,G) entry. This provides for zero multicast packet loss on FTOS with two exceptions:

1. These packets can be soft (slow path) forwarded to receivers at a maximum rate of 70 packets/second. Incoming packets beyond this rate are dropped.
2. When the system is both the source-DR and the RP, in some cases, packet loss, packet reordering, or duplicate packets might occur.

To prevent these delivery errors you must statically map the potential incoming interfaces for the (*,G) entries via the CLI. When you create this mapping, (*,G) entries are programmed in hardware. Packets are then fast forwarded starting with the first packet, and the potential for these delivery errors is avoided.

Step	Task	Command Syntax	Command Mode
1	Create a standard access-list that permits one or more IGMP groups. `FTOS(config-std-nacl)#show config` `!` `ip access-list standard map1` `seq 5 permit 224.0.0.0/24`	**ip access-list standard** *name*	CONFIGURATION
2	Apply the ACL to an interface on which there might be a source for a group specified in the ACL. This command maps the incoming interface to the (*,G) entry so that the entry can be programmed into hardware.	**ip pim ingress-interface-map** *std-access-list*	INTERFACE

Monitoring PIM

The PIM MIB is supported only on platform E

FTOS fully supports the PIM MIB as specified in RFC 5060 with some exceptions.

- The following tables are not supported:
  - pimBidirDFElectionTable
  - pimAnycastRPSetTable
- The OIDs related to InvalidRegisterMsgs reflect the last received invalid register message. Similarly, the OIDs related to InvalidJoinPruneMsgs reflect the last received invalid Join or Prune message.
- OIDs which refer to any timer show the time that the timer started; it is 0 otherwise.

PIM-SM and IGMP Snooping: Usage Notes

PIM-SM is supported with IGMP snooping. Figure 34-10 shows the egress ports used for outgoing multicast traffic when you enable different combinations of PIM-SM DR flooding and IGMP snooping flooding on a switch/router.

When you enable PIM-SM and IGMP snooping at the same time:

- The IGMP report is forwarded on the port that connects to the PIM DR.
- The port that connects to the PIM DR port and ports on which IGMP queries are received are chosen as IGMP router ports. It is recommended that you configure the both the IGMP querier and the PIM DR in the same router to avoid unnecessary flooding.

- It is recommended that you do not enable IGMP snooping on a PIM-SM snooping-enabled VLAN interface unless until it is necessary for VLAN operation.

Table 34-1. Egress Ports Used for Multicast Traffic with PIM-SM and IGMP Snooping

Multicast Traffic	PIM-SM and IGMP Snooping Configuration	Egress Ports
Known multicast data packets	PIM-SM snooping DR flood with IGMP snooping flood	PIM-SM snooped VLANs PIM DR port IGMP snooped ports IGMP mrouter ports
	PIM-SM snooping DR flood with no IGMP snooping flood	PIM-SM snooped VLANs PIM DR port IGMP snooped ports IGMP mrouter ports
	IGMP flood with no PIM-SM snooping DR flood	PIM-SM snooped VLANs IGMP snooped ports IGMP mrouter ports
	No PIM-SM snooping DR flood and no IGMP snooping flood	PIM-SM snooped VLANs IGMP snooped ports IGMP mrouter ports
Unknown multicast data packets	PIM-SM snooping DR flood with IGMP snooping flood	PIM DR port
	PIM-SM snooping DR flood with no IGMP snooping flood	PIM DR port IGMP mrouter ports
	IGMP flood with no PIM-SM snooping DR flood	None
	No PIM-SM snooping DR flood and no IGMP snooping flood	IGMP mrouter ports

For information on how to enable PIM-SM snooping and disable PIM DR flooding, refer to PIM-SM Snooping on page 767.

PIM-SM Snooping

PIM-SM Snooping is supported on VLAN interfaces on platform: 

In a Layer 2 VLAN, a switch normally floods multicast traffic to all member ports in the VLAN. PIM-SM snooping is designed to restrict multicast traffic to only the PIM-SM-enabled routers and IGMP hosts that should receive the traffic. When PIM-SM snooping is enabled, the switch learns the multicast ports on which receivers are listening through PIM hello and PIM-SM join/prune messages, and forwards multicast traffic only to the VLAN ports connected to valid receivers.

Feature Overview

PIM-SM snooping functions in a Layer 2 network in which multiple routers are interconnected by a switch, such as an IXP where Internet service providers (ISPs) exchange Internet traffic between their networks. By default, the switch floods multicast traffic to all VLAN member ports, regardless of whether there are multicast receivers downstream that are joined to a multicast group.

When you enable PIM-SM snooping, the switch receives PIM hello and PIM-SM join/prune messages, and determines which multicast ports are connected to receivers to which multicast packets should be forwarded. Multicast data is forwarded only to VLAN member ports on which there are valid downstream receivers.

- Using PIM hello messages, the switch learns about PIM neighbors and builds a database for the VLAN and port on which the packets are received. The PIM Snooping neighbor database is the same one used for PIM-SM.

  Each neighbor entry stores the physical or port-channel port on which a hello message from a neighbor is received. PIM hello messages are flooded to all VLAN member ports, except the port on which the message was received. The PIM designated router on the VLAN is learned from the snooped PIM hello packets.
- A PIM-SM snooping-enabled switch will proxy the join/prune messages it receives to minimize the messages it sends upstream. The switch consumes the join/prune messages received from downstream neighbors and initiates join/prune messages towards upstream neighbors.

  All other PIM protocol messages are flooded to VLAN member ports. PIM join/prune messages to non-existent upstream neighbors are silently dropped.

  PIM-SM join/prune messages towards an upstream neighbor are sent only to the port corresponding to the upstream router in the join message.

  PIM (S,G,Rpt) prune and (S,G,Rpt) join messages are snooped and managed according to the PIM-SM RFC.
- The switch creates and maintains a tree topology with the state of PIM neighbors in the tree information base (TIB). Each PIM snooping-enabled VLAN has its own neighbor tree in the TIB.

  The PIM (*,G) TIB state maintains the list of multiple upstream neighbors for joins initiated by downstream routers towards the rendezvous point (RP). The PIM (*,G,) TIB adds all other upstream neighbor ports to its Outgoing Interface list, except the port to which the join was forwarded, to trigger assert conditions.

  The PIM (S,G) TIB state maintains the list of multiple upstream neighbors for joins initiated by downstream routers towards the source.
- If the PIM designated-router (DR) flood is not disabled (default setting; see Disable PIM Designated-Router Flooding on page 772):
  - Multicast traffic is transmitted on the egress port towards the PIM DR if the port is not the incoming interface.
  - Multicast traffic for an unknown group is sent on the port towards the PIM DR. When DR flooding is disabled, multicast traffic for an unknown group is dropped.
- Multicast traffic for known multicast group addresses, such as Local Network Control Block and Internetwork Control Block (as defined in RFC 5771), is flooded to all VLAN member ports.

- In the downstream PIM TIB, states and timers are maintained for each VLAN and member port. The downstream outgoing-interface timers for each valid (*,G) and (S,G) entry are started for each VLAN/port and upstream neighbor combination: (*port*,*,G,*neighbor*) or (*port*,S,G,*neighbor*), where *port* is a downstream port and *neighbor* is the upstream neighbor.
  - A timer is removed when a timer times out or a prune message is received for a specific VLAN member port.
  - PIM-SM snooping does not use the unicast Real Time Monitor (RTM) to forward snooped packets.
  - All multicast route and router information is timed-out based on the hold-time indicated in the PIM hello and PIM-SM join/prune control packets.

Configuration Notes and Restrictions

The following conditions apply when you configure and use PIM snooping on a switch:

- PIM-SM snooping is deployed in a Layer 2 environment and is mutually exclusive with PIM multicast routing. If you enable PIM-SM snooping, you cannot enable PIM-SM or PIM-DM. If you enable PIM-SM snooping, you cannot enable PIM-SM or PIM-DM
- PIM-SM snooping is supported with IGMP snooping, and forwards the IGMP report on the port that connects to the PIM DR.

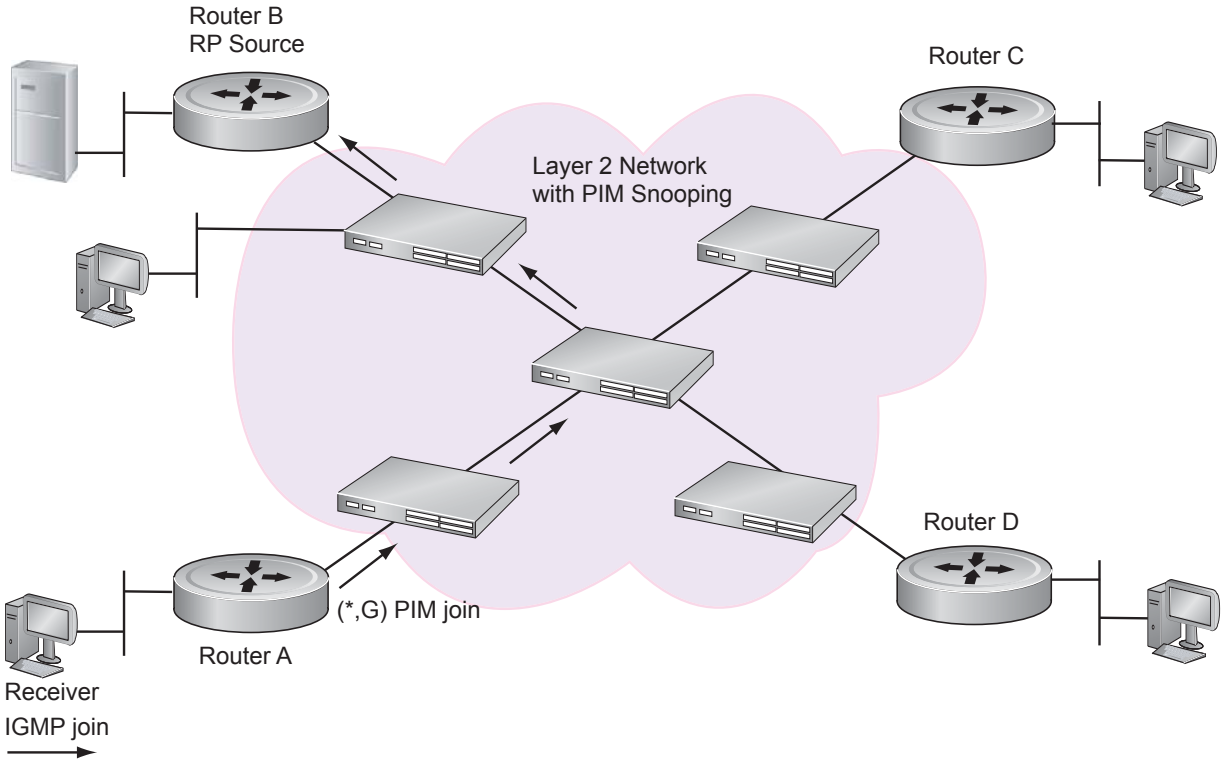
  It is recommended that you do not enable IGMP snooping on a PIM-SM snooping-enabled VLAN interface unless until it is necessary for VLAN operation.

  For information on the egress ports used for outgoing multicast traffic when you enable PIM-SM snooping and IGMP snooping at the same time on a VLAN interface, see PIM-SM and IGMP Snooping: Usage Notes on page 766.
- PIM-SM snooping listens to PIM hello and PIM-SM join and prune messages while maintaining the VLAN- and port-specific information in multicast packets that are snooped.

PIM-SM Snooping Example

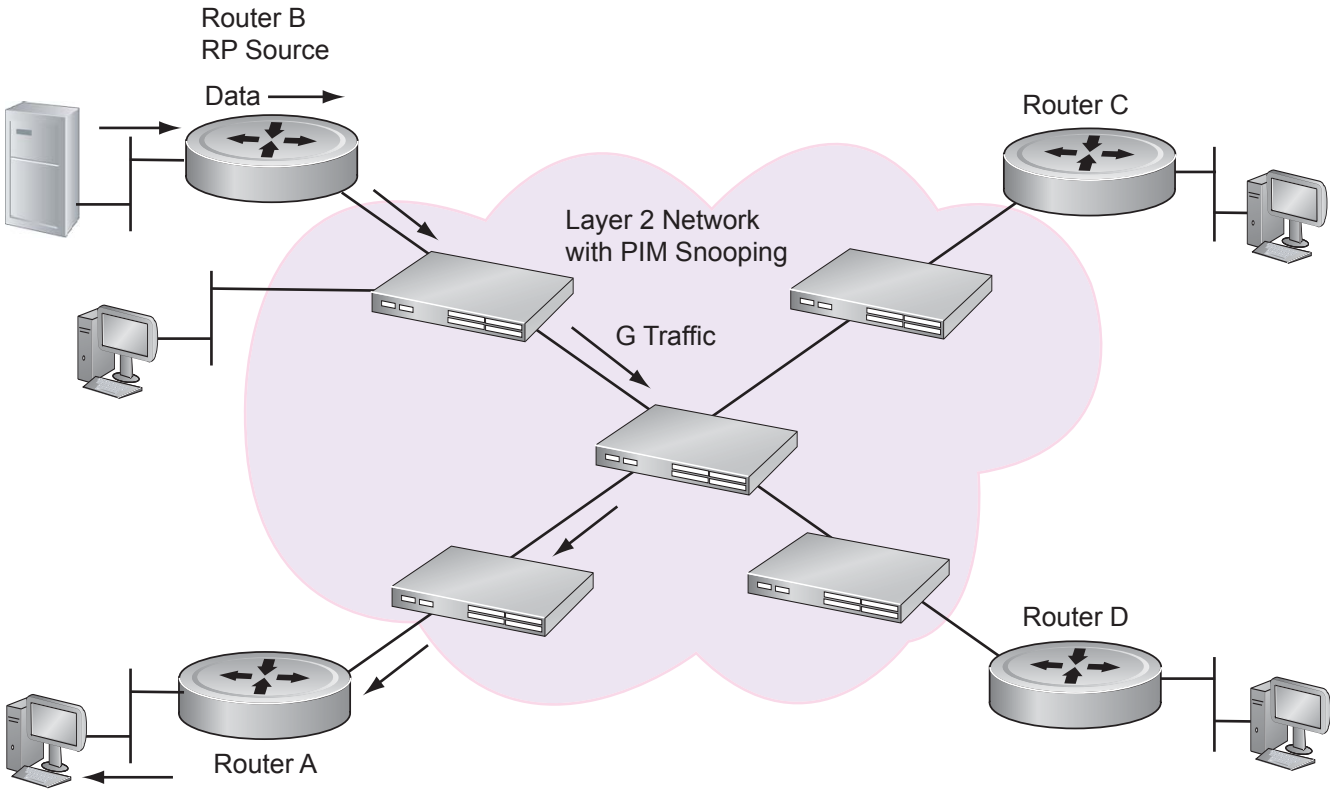
Figure 34-8 shows an example with PIM-SM snooping enabled. When Router A sends a join message to Router B, the switches forward the join message only to Router B without flooding the message to other connected routers, such as Routers C and D.

Figure 34-8. PIM-SM Snooping: Join Message Flow



Similarly, in Figure 34-8, when PIM-SM snooping is enabled and multicast data is sent to VLAN members of group G, the switches forward the data traffic from the server attached to Router B only to the router (Router A) in the multicast group that should receive it. Without PIM-SM snooping, the switches would flood the data to all connected routers, including Routers C and D.

Figure 34-9. PIM-SM Snooping: Data Forwarding



PIM-SM Snooping Configuration

You can enable PIM-SM snooping globally on a switch or on individual VLANs. PIM-SM snooping is not enabled by default and does not require an IP address, PIM-DM, or PIM-SM.

PIM-SM snooping and PIM multicast routing are mutually exclusive: PIM-SM snooping cannot be enabled on a switch/router if PIM-SM or PIM-DM is enabled.

If enabled at the global level, PIM-SM snooping is automatically enabled on all VLANs on the switch unless the **no ip pim snooping** command has been entered on a VLAN interface.

If enabled at the VLAN level, PIM-SM snooping requires that you also enter the **no shutdown** command to enable the interface.

Enable PIM Snooping

To enable PIM-SM snooping on all VLAN interfaces on a switch, enter the following command.

Task	Command	Command Mode
Enable PIM-SM snooping globally on a switch.	**ip pim snooping enable**	CONFIGURATION

To enable PIM-SM snooping on a VLAN interface, enter the following commands:

Task	Command	Command Mode
Enable PIM-SM snooping on a VLAN interface.	**ip pim snooping**	VLAN INTERFACE
Enable the interface.	**no shutdown**	VLAN INTERFACE

Disable PIM Designated-Router Flooding

By default, when you enable PIM-SM snooping, a switch floods all multicast traffic to the PIM designated router (DR), including unnecessary multicast packets. To minimize the traffic sent over the network to the designated router, you can disable designated-router flooding.

When designated-router flooding is disabled, PIM-SM snooping only forwards multicast traffic, which belongs to a multicast group for which the switch receives a join request, on the port connected towards the designated router.

To disable designated-router flooding for PIM-SM snooping, enter the **no ip pim snooping dr-flood** command:

Task	Command	Command Mode
Disable flooding of multicast packets to the designated-router.	**no ip pim snooping dr-flood**	CONFIGURATION

Verify PIM-SM Snooping

To display information about PIM-SM snooping operation, enter one of the following **show** commands:

Task	Command	Command Mode
Display information about PIM neighbors discovered by PIM-SM snooping.	**show ip pim snooping neighbor** [**vlan** *vlan-id*] Figure 34-10	EXEC Privilege
Display information about PIM group members and states stored in the tree information base (TIB) that was discovered by PIM-SM snooping.	**show ip pim snooping tib** [**vlan** *vlan-id*] [*group-address* [*source-address*]] Figure 34-11	EXEC Privilege
Display information about the VLAN interfaces on which PIM-SM snooping is configured.	**show ip pim snooping interface** [**vlan** *vlan-id*] Figure 34-12	EXEC Privilege
Display information about the current operation of PIM-SM snooping globally on the switch or on a specified VLAN.	**show ip pim summary** Figure 34-13	EXEC Privilege VLAN INTERFACE
Display information on the multicast routes discovered on VLANs configured for PIM-SM snooping.	**show ip mroute snooping** [**vlan** *vlan-id*] [*group-address* [*source-address*]] Figure 34-14	EXEC Privilege
Display information about the current configuration of PIM-SM snooping on the switch.	**show running-config pim** Figure 34-15	EXEC Privilege
Display the current configuration of PIM-SM snooping on a VLAN.	**show configuration** Figure 34-16	VLAN INTERFACE

To clear tree information learned through PIM-SM snooping from the PIM TIB, enter the **clear ip pim snooping tib** command.

To clear information on the multicast routes learned through PIM-SM snooping from the IPv4 multicast snooping table, enter the **clear ip mroute snooping** command.

The following examples show the PIM-SM snooping output displayed for these **show** commands.

Figure 34-10. PIM-SM snooping: show ip pim snooping neighbor

```
FTOS#show ip pim snooping neighbor

Neighbor              Interface            Uptime/Expires     Ver  DR Prio
Address
165.87.32.2           Vl 2 [Gi 4/13 ]      00:04:03/00:01:42  v2   1
165.87.32.10          Vl 2 [Gi 4/11 ]      00:00:46/00:01:29  v2   0
165.87.32.12          Vl 2 [Gi 4/20 ]      00:00:51/00:01:24  v2   0
```

Figure 34-11. PIM-SM snooping: show ip pim snooping tib

```
FTOS#show ip pim snooping tib

PIM Multicast Snooping Table
Flags: J/P - (*,G) Join/Prune, j/p - (S,G) Join/Prune
       SGR-P - (S,G,R) Prune
Timers: Uptime/Expires
* : Inherited port

(*, 225.1.2.1), uptime 00:00:01, expires 00:02:59, RP 165.87.70.1, flags: J
  Incoming interface: Vlan 2, RPF neighbor 0.0.0.0
  Outgoing interface list:
    GigabitEthernet 4/11  RPF 165.87.32.2           00:00:01/00:02:59
    GigabitEthernet 4/13  Upstream Port             -/-

FTOS#show ip pim snooping tib vlan 2 225.1.2.1 165.87.1.7

PIM Multicast Snooping Table
Flags: J/P - (*,G) Join/Prune, j/p - (S,G) Join/Prune
       SGR-P - (S,G,R) Prune
Timers: Uptime/Expires
* : Inherited port

(165.87.1.7, 225.1.2.1), uptime 00:00:08, expires 00:02:52, flags: j
  Incoming interface: Vlan 2, RPF neighbor 0.0.0.0
  Outgoing interface list:
    GigabitEthernet 4/11  Upstream Port             -/-
    GigabitEthernet 4/13  DR Port                   -/-
    GigabitEthernet 4/20  RPF 165.87.32.10          00:00:08/00:02:52
```

Figure 34-12. PIM-SM snooping: show ip pim snooping interface

```
FTOS#show ip pim snooping interface
Interface   Ver   Nbr      DR         DR
                  Count    Prio
Vlan 2      v2    3        1          165.87.32.2
```

Figure 34-13. PIM-SM snooping: show ip pim summary

```
FTOS#show ip pim summary

PIM TIB version 495
Uptime 22:44:52
Entries in PIM-TIB/MFC : 2/2

Active Modes :
        PIM-SNOOPING

Interface summary:
        1 active PIM interface
        0 passive PIM interfaces
        3 active PIM neighbors

TIB summary:
        1/1 (*,G) entries in PIM-TIB/MFC
        1/1 (S,G) entries in PIM-TIB/MFC
        0/0 (S,G,Rpt) entries in PIM-TIB/MFC

        0 PIM nexthops
        0 RPs
        0 sources
        0 Register states

Message summary:
        2582/2583 Joins sent/received
        5/0 Prunes sent/received
        0/0 Candidate-RP advertisements sent/received
        0/0 BSR messages sent/received
        0/0 State-Refresh messages sent/received
        0/0 MSDP updates sent/received
        0/0 Null Register messages sent/received
        0/0 Register-stop messages sent/received

Data path event summary:
        0 no-cache messages received
        0 last-hop switchover messages received
        0/0 pim-assert messages sent/received
        0/0 register messages sent/received

Memory usage:
        TIB             : 3768 bytes
        Nexthop cache   : 0 bytes
        Interface table : 992 bytes
        Neighbor table  : 528 bytes
        RP Mapping      : 0 bytes
```

Figure 34-14. PIM-SM snooping: show ip mroute snooping

```
FTOS#show ip mroute snooping

IPv4 Multicast Snooping Table

(*, 224.0.0.0), uptime 17:46:23
  Incoming vlan: Vlan 2
  Outgoing interface list:
    GigabitEthernet 4/13

(*, 225.1.2.1), uptime 00:04:16
  Incoming vlan: Vlan 2
  Outgoing interface list:
    GigabitEthernet 4/11
    GigabitEthernet 4/13

(165.87.1.7, 225.1.2.1), uptime 00:03:17
  Incoming vlan: Vlan 2
  Outgoing interface list:
    GigabitEthernet 4/11
    GigabitEthernet 4/13
    GigabitEthernet 4/20
```

Figure 34-15. PIM-SM snooping: show running-config

```
FTOS#show running-config pim
!
ip pim snooping enable
```

Figure 34-16. PIM-SM snooping: show configuration

```
FTOS(conf-if-vl-2)#show config
!
interface Vlan 2
 no ip address
 tagged GigabitEthernet 4/11-13,20-23
 no shutdown
FTOS(conf-if-vl-2)#
```

35

PIM Source-Specific Mode

PIM Source-Specific Mode is supported on platforms: C E S

PIM-SSM is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

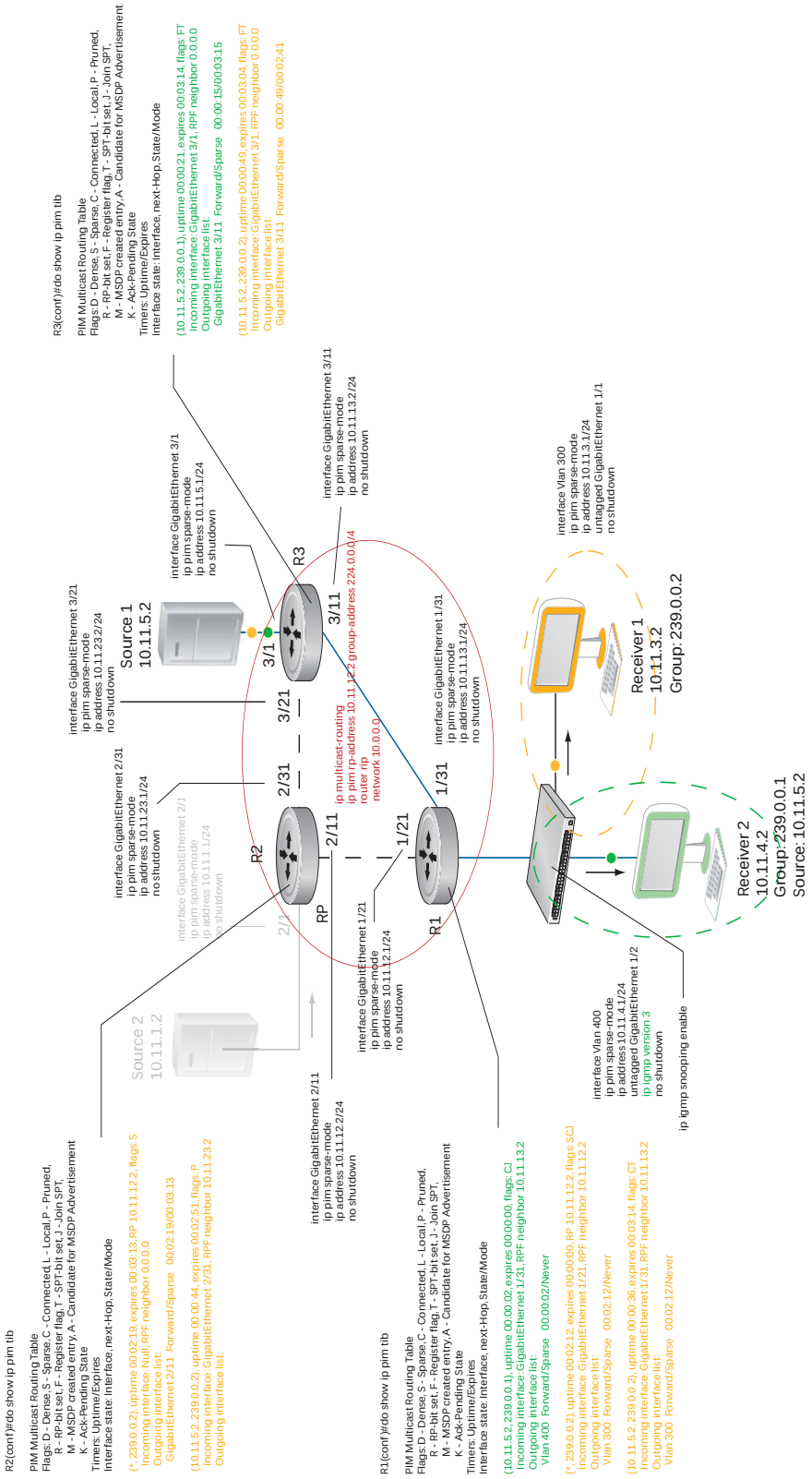
PIM-Source-Specific Mode (PIM-SSM) is a multicast protocol that forwards multicast traffic from a single source to a subnet. In the other versions of Protocol Independent Multicast (PIM), a receiver subscribes to a group only. The receiver receives traffic not just from the source in which it is interested but from all sources sending to that group. PIM-SSM requires that receivers specify the sources in which they are interested using IGMPv3 include messages to avoid receiving unwanted traffic.

PIM-SSM is more efficient than PIM-SM because it immediately creates shortest path trees (SPT) to the source rather than first using shared trees. PIM-SM requires a shared tree rooted at the RP because IGMPv2 receivers do not know about the source sending multicast data. Multicast traffic passes from the source to the receiver through the RP, until the receiver learns the source address, at which point it switches to the SPT. PIM-SSM uses IGMPv3. Since receivers subscribe to a source and group, the RP and shared tree is unnecessary, so only SPTs are used. On Dell Force10 systems, it is possible to use PIM-SM with IGMPv3 to achieve the same result, but PIM-SSM eliminates the unnecessary protocol overhead.

PIM-SSM also solves the multicast address allocation problem. Applications should use unique multicast addresses because if multiple applications use the same address, receivers receive unwanted traffic. However, global multicast address space is limited. Currently GLOP/EGLOP is used to statically assign Internet-routable multicast addresses, but each autonomous system number yields only 255 multicast addresses. For short-term applications, an address could be leased, but no global dynamic multicast address allocation scheme has been accepted yet. PIM-SSM eliminates the need for unique multicast addresses because routing decisions for (S1, G1) are independent from (S2, G1). As a result, subnets do not receive unwanted traffic when multiple applications use the same address.

In Figure 35-1, Receiver 1 is an IGMPv2 host. The packets for group 239.0.0.2 travel to it first via the RP, then by the SPT. Receiver 2 is an IGMPv3 host. The packets for group 239.0.0.1 travel only via the STP.

Figure 35-1. PIM-SM with IGMPv2 versus PIM-SM with IGMPv3



Implementation Information

- The Dell Force10 implementation of PIM-SSM is based on RFC 3569.
- C-Series supports a maximum of 31 PIM interfaces and 4K multicast entries including (*,G), and (S,G) entries. There is no limit on the number of PIM neighbors C-Series can have.
- S-Series supports a maximum of 31 PIM interfaces and 2K multicast entries including (*,G), and (S,G) entries. There is no limit on the number of PIM neighbors S-Series can have.
- E-Series supports a maximum of 511 PIM interfaces and 50K multicast entries including (*,G), (S,G), and (S,G,rpt) entries. There is no limit on the number of PIM neighbors E-Series can have.
- FTOS reduces the number of control messages sent between multicast routers by bundling Join and Prune requests in the same message.

Important Points to Remember

- The default SSM range is 232/8 always. Applying an SSM range does not overwrite the default range. Both the default range and SSM range are effective even when the default range is not added to the SSM ACL.
- Extended ACLs cannot be used for configuring SSM range. Be sure to create the ACL first and then apply it to the SSM range.
- The default range is always supported, so range can never be smaller than the default.

Configure PIM-SM

Configuring PIM-SSM is a one-step process:

1. Enable PIM-SM. See page 758.
2. Enable PIM-SSM for a range of addresses. See page 780.

Related Configuration Tasks

- Use PIM-SSM with IGMP version 2 Hosts on page 780

Enable PIM-SSM

To enable PIM-SSM:

Step	Task	Command Syntax	Command Mode
1	Create an ACL that uses permit rules to specify what range of addresses should use SSM. You must at least include one rule, **permit 232.0.0.0/8**, which is the default range for PIM-SSM.	[**ip** | **ipv6**] **access-list standard** *name*	CONFIGURATION
2	Enter the command **ip pim ssm-range** and specify the ACL you created.	[**ip** | **ipv6**] **pim ssm-range** *acl-name*	CONFIGURATION

Display address ranges in the PIM-SSM range using the command **show [ip | ipv6] pim ssm-range** from EXEC Privilege mode.

Figure 35-2. Enabling PIM-SSM

```
R1(conf)#do show run pim
!
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
ip pim ssm-range ssm
R1(conf)#do show run acl
!
ip access-list standard ssm
 seq 5 permit host 239.0.0.2
R1(conf)#do show ip pim ssm-range
Group Address   / MaskLen
239.0.0.2       / 32
```

Use PIM-SSM with IGMP version 2 Hosts

PIM-SSM requires receivers that support IGMP version 3. You can employ PIM-SSM even when receivers support only IGMP version 1 or version 2 by translating (*,G) entries to (S,G) entries.

Translate (*,G) entries to (S,G) entries using the command **ip igmp ssm-map** *acl source* from CONFIGURATION mode. In a standard access list, specify the groups or the group ranges that you want to map to a source. Then, specify the multicast source.

- When a SSM map is in place and FTOS cannot find any matching access lists for a group, it continues to create (*,G) entries because there is an implicit deny for unspecified groups in the ACL.
- When you remove the mapping configuration, FTOS removes the corresponding (S,G) states that it created and reestablishes the original (*,G) states.
- You may enter multiple **ssm-map** commands for different access lists. You may also enter multiple **ssm-map** commands for the same access list, as long as they use different source addresses.

- When an extended ACL is associated with this command, FTOS displays an error message. If you apply an extended ACL before you create it, FTOS accepts the configuration, but when the ACL is later defined, FTOS ignores the ACL and the stated mapping has no effect.

Display the source to which a group is mapped using the command **show ip igmp ssm-map** [*group*], as shown in Figure 35-4 on page 783. If use the *group* option, the command displays the group-to-source mapping even if the group is not currently in the IGMP group table. If you do not specify the *group* option, then the display is a list of groups currently in the IGMP group table that have a group-to-source mapping.

Display the list of sources mapped to a group currently in the IGMP group table using the command **show ip igmp groups** *group* **detail**, as shown in Figure 35-4 on page 783.

Figure 35-3. Using PIM-SM with IGMPv2 versus PIM-SSM with IGMPv2

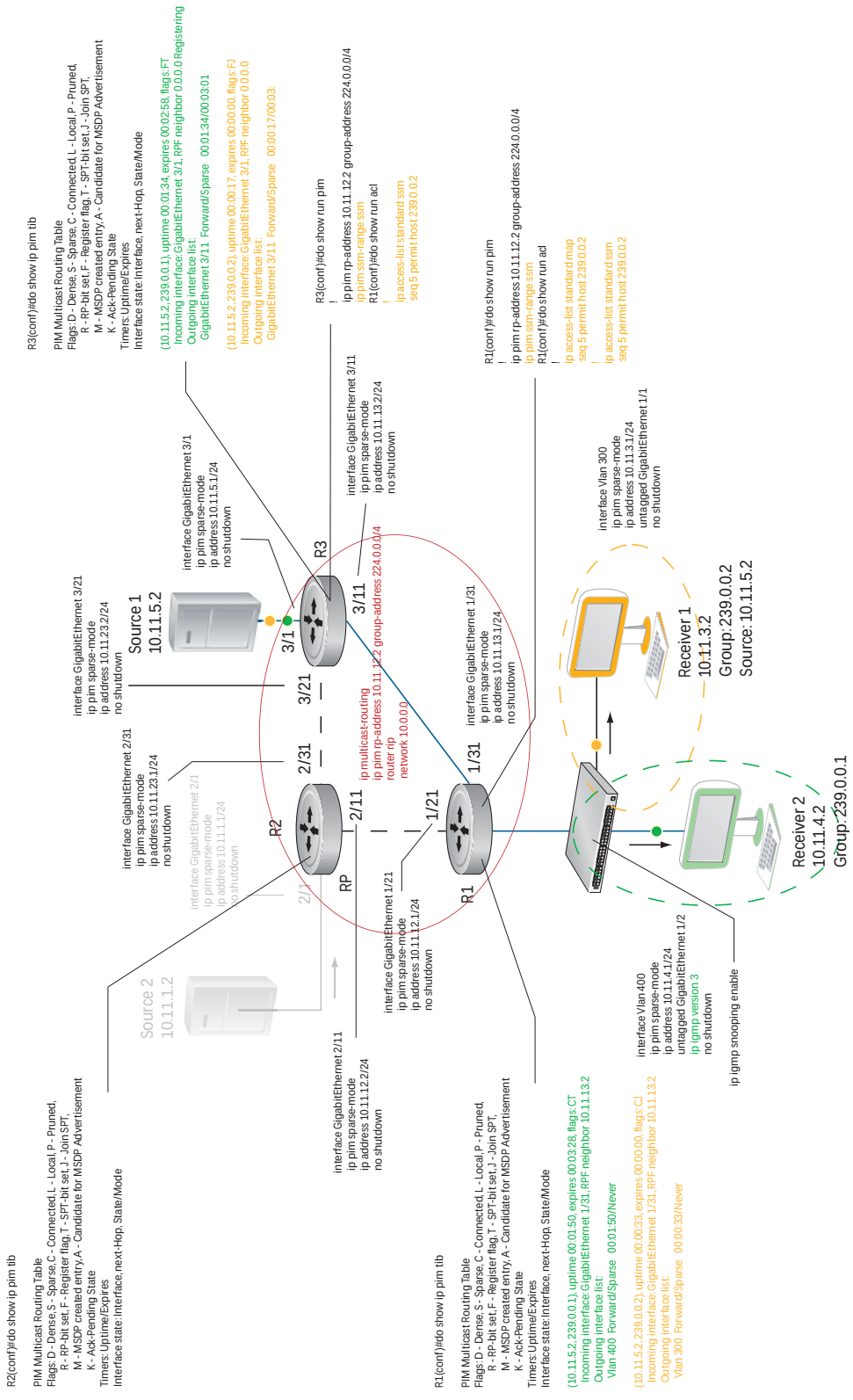


Figure 35-4. Configuring PIM-SSM with IGMPv2

```
R1(conf)#do show run pim
!
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
ip pim ssm-range ssm
R1(conf)#do show run acl
!
ip access-list standard map
 seq 5 permit host 239.0.0.2
!
ip access-list standard ssm
 seq 5 permit host 239.0.0.2
R1(conf)#ip igmp ssm-map map 10.11.5.2
R1(conf)#do show ip igmp groups
Total Number of Groups: 2
IGMP Connected Group Membership
Group Address    Interface                Mode          Uptime    Expires   Last Reporter
239.0.0.2        Vlan 300                 IGMPv2-Compat 00:00:07  Never     10.11.3.2
  Member Ports: Gi 1/1
239.0.0.1        Vlan 400                 INCLUDE       00:00:10  Never     10.11.4.2
R1(conf)#do show ip igmp ssm-map
IGMP Connected Group Membership
Group Address    Interface                Mode          Uptime    Expires   Last Reporter
239.0.0.2        Vlan 300                 IGMPv2-Compat 00:00:36  Never     10.11.3.2
  Member Ports: Gi 1/1
R1(conf)#do show ip igmp ssm-map 239.0.0.2
SSM Map Information
Group     : 239.0.0.2
Source(s) : 10.11.5.2
R1(conf)#do show ip igmp groups detail

Interface             Vlan 300
Group                 239.0.0.2
Uptime                00:00:01
Expires               Never
Router mode           IGMPv2-Compat
Last reporter         10.11.3.2
Last reporter mode    IGMPv2
Last report received  Join
Group source list
Source address                   Uptime      Expires
10.11.5.2                        00:00:01    Never

Interface             Vlan 400
Group                 239.0.0.1
Uptime                00:00:05
Expires               Never
Router mode           INCLUDE
Last reporter         10.11.4.2
Last reporter mode    INCLUDE
Last report received  ALLOW
Group source list
Source address                   Uptime      Expires
10.11.5.2                        00:00:05    00:02:04
  Member Ports: Gi 1/2
```

36

Power over Ethernet

Power over Ethernet (PoE) is supported only on platforms: C S

This chapter contains the following major sections:

- Configuring Power over Ethernet on page 786
- Power Additional PoE Ports on the S-Series on page 794
- Deploying VOIP on page 795

FTOS supports Power over Ethernet (PoE), as described by IEEE 802.3af . IEEE 802.3af specifies that a maximum of 15.4 Watts can be transmitted to Ethernet devices over the signal pairs of an unshielded twisted pair (UTP) cable. PoE is useful in networks with IP phones and wireless access points because separate power supplies for *powered devices* (PD) are not needed.

Table 36-2 describes the classes of powered devices defined by IEEE 802.3af:

Table 36-1. PoE Classes of Powered Devices

Class	Power Range (Watts)	Classification Current (mA)
0	0.44 to 12.95	< 5.0
1	0.44 to 3.84	10.5
2	3.84 to 6.49	18.5
3	6.49 to 12.95	28
4	Reserved	40

Note: FTOS treats Class 0, Class 3, and Class 4 powered devices the same. Class 4 is meant for IEEE802.3at compliant devices which require >12.95 Watts. Currently FTOS treats Class 4 devices as Class 3.

FTOS supports PoE on all copper ports on the C-Series and on the S25V and S50V models of the S-Series. The C-Series and S-Series transmit power to connected IEEE 802.3af-compliant powered devices through ports that have been configured to supply PoE. Those platforms also support the protocols LLDP and LLDP-MED, which help optimize power distribution to PoE devices. See Chapter 46, Link Layer Discovery Protocol, on page 861.

For the C-Series, FTOS requires that a minimum number of AC power supplies (PSU) be installed before PoE can be enabled, and some PSUs are reserved for PoE redundancy, as described in Table 36-2.

Note: The C-Series can provide PoE only through its AC power supplies.

Table 36-2. PoE Ports per Power Supply Unit in the C-Series*

Number of Power Supply Units	Max PoE Ports on C300	Max PoE Ports on C150
1	—	—
2	—	System Redundancy
3	System Redundancy	96
4	96	192
5	192	PoE Redundancy
6	288	PoE Redundancy
7	384	N/A
8	PoE redundancy	N/A

FTOS Behavior: Table 36-2 provides the maximum number of PoE ports per PSU, based on the assumption that each port deliver 15.4W. In many cases, the PD requires <15.4W. Typical IP Phones require only 3-10 Watts. So, if the ports are configured optimally, more PDs can be powered with fewer PSUs.

On the C-Series, though each PSU used for PoE (units 4-7 on the C300, and 3-4 on the C150) provides 1200 Watts of power, each actually makes available 1478.40 Watts for PoE. This is possible because each unit, once installed, borrows 278.40 Watts from the system redundancy power supply. If a power supply used for PoE is removed, PoE ports are shut down so that the system redundancy PSU retains is capability.

Note: The S25V and S50V models contain AC power supplies in order to support PoE. You can also add the external Dell Force10 470W Redundant Power Supply to power more PoE devices. For details, see Power Additional PoE Ports on the S-Series on page 794 and see the **power budget** command in the Power Over Ethernet (PoE) chapter of the *FTOS Command Reference for the S-Series*.

Configuring Power over Ethernet

Configuring PoE is a two-step process:

1. Connect the IEEE 802.3af-compliant powered device directly to a port.
2. Enable PoE on the port, as described next.

Related Configuration Tasks

- Manage Ports using Power Priority and the Power Budget on page 789
- Monitor the Power Budget on page 792
- Manage Power Priorities on page 792
- Recover from a Failed Power Supply on page 793
- Power Additional PoE Ports on the S-Series on page 794

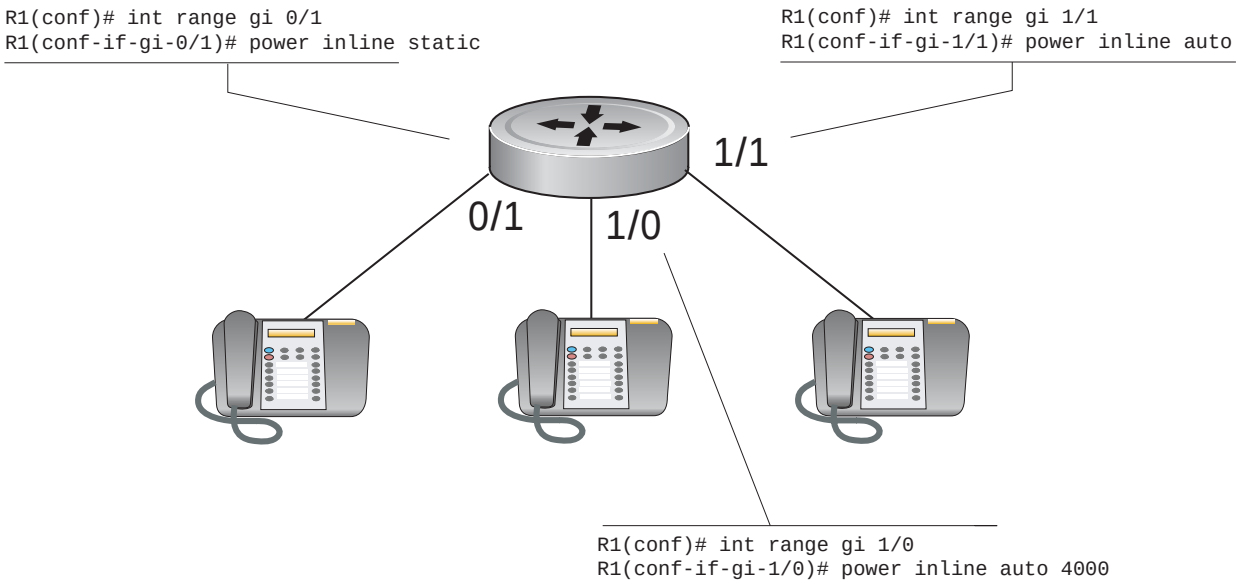
Enabling PoE on a Port

PoE is disabled by default. Enable PoE on a port from INTERFACE mode using the command **power inline** {**auto** [*max_milliwatts*] | **static** [*max_milliwatts*]}.

- The **power inline auto** command allows the port to determine the amount of power that a connected Class 1–4 powered device requires, and supply it. See Table 36-1 on page 785.
- The **power inline static** command without the qualifier guarantees 15.4W to the powered device.
- You can limit the maximum amount of power (in milliwatts) available to a powered device with the command **power inline auto** *max_milliwatts* or with **power inline static** *max_milliwatts*
- Disable PoE on a port using the **no power inline** command.

Ports configured with **power inline auto** have a lower priority for access to power than those configured with **power inline static**. As a second layer of priority setting, use the **[no] power inline priority** command. Use the **power inline static** *max_milliwatts* command to avoid allocating more power than necessary to a port because allocated power is made unavailable to other ports regardless of whether it is consumed. Typical IP phones use 3-10 Watts.

Figure 36-1. Enabling PoE



View the amount of power that a port is consuming using the **show power inline** command from EXEC privilege mode.

Figure 36-2. PoE Allocation Displayed with show power inline Command

```
FTOS(conf-if-range-gi-0/1-48)#do show power inline
Interface   Admin   Inline Power   Inline Power    Class        User
                     Allocated      Consumed                  Priority
                      (Watts)        (Watts)
---------   -----   ------------   ------------    -----     ----------
Gi 0/1      auto       0.00           0.00         NO_DEVICE    Low
Gi 0/2      auto       7.00           3.20              2       Low
```

Table 36-3 describes the fields that the **show power inline** command displays:

Table 36-3. show power inline Field Description

Field	Port Number
Interface	Displays all PoE-enabled ports.
Admin	Displays the administrative mode of the interface: • *auto* indicates that power is supplied according to the requirements of the powered device. • *static* indicates that the maximum configured amount of power is supplied to the powered device.
Inline Power Allocated	Displays the amount of power that is allocated to a port when sufficient power is available. When sufficient power is not available for particular port, then inline power is not supplied to that port. If you insert an additional power supply, or when the priority of the port is sufficiently increased, then the system supplies the allocated power to the port.
Inline Power Consumed	Displays the amount of power that a powered device is consuming.
Class	Displays the type of powered device: Class 0, Class 1, Class 2, Class 3, or Class 4. Displays NO_DEVICE if no PD is connected.

View the total power consumption of the chassis using the **show power detail** command from EXEC privilege mode.

Figure 36-3. PoE Consumed, Allocated, and Available with show power detail Command

```
FTOS(conf-if-range-gi-0/1-48)#do show power detail
Unit    Total      Logic     Inline     Inline     Inline    Inline
        Power      Power     Power      Power      Power     Power
        Available  Consumed  Available  Allocated  Consumed  Remaining
        (Watts)    (Watts)   (Watts)    (Watts)    (Watts)   (Watts)
-------------------------------------------------------------------------------
   0     470.00     150       320.00     308.00     190.00      12.00
```

Table 36-4 describes the fields that the **show power detail** command displays.

Table 36-4. show power detail Field Description

Field	Port Number
Unit	(S-Series only) The stack member unit ID.
Catalog Name	(C-Series only) Displays the component's Dell Force10 catalog number.
Slot ID	(C-Series only) Displays the slot number in which the line card or RPM is installed.
Total Power Available	The total power available in the stack member or chassis. **Note:** On the S-Series a maximum of 790W can be allocated for PoE, even if you add the 470W external power supply.
Logic Power Consumed	The power consumed by the system logic.
Inline Power Available	Power available for PoE (whatever was configured using **power-budget** command. Default: 320 watts
Inline Power Allocated	Total power allocated to the ports.
Inline Power Consumed	Total power consumed by connected devices.
Inline Power Remaining	Difference between power available and power allocated.

Manage Ports using Power Priority and the Power Budget

The allocation and return of power on ports depends on the total inline power available in the system and the power priority calculation.

Determine the Power Priority for a Port

FTOS uses a sophisticated port prioritization algorithm for determining which ports receive PoE so that PoE ports are powered up/down deterministically.

FTOS uses the following four parameters, in order, for defining the power priority for a port:

1. the **power-inline** mode: **static** or **auto**,
2. the **power-inline priority** configuration,
3. the LLDP-MED priority sent by the PD in the Extended Power-via-MDI TLV,
4. and the port's slot and port number.

FTOS maintains a sorted list of PoE ports based these four parameters. Static ports have a higher weight than auto mode ports, so all static ports always stay on top of all auto ports regardless of the other 3 parameters. Within the set of static ports, FTOS attempts to order them based on the second parameter **power-inline priority**, the default of which is "Low". If FTOS finds multiple ports with the same

power-inline priority, it breaks the tie using the third parameter, the LLDP-MED Priority advertised by the PD, which like **power-inline priority** could be “Critical,” “High,” or “Low”. After this, if FTOS still finds a tie, priority is based on the fourth parameter which is the ports position in the chassis; there cannot be a tie based on this parameter.

This sorted list is dynamically updated by FTOS when:

- a user changes the **power-inline** mode or priority
- the PD advertises a different LLDP-MED priority
- the PD is connected or disconnected

FTOS always uses this sorted list of ports for allocation. When an additional PSU is added, additional ports are powered based on this list, and PSU is removed, this same list is used to remove power from the lowest priority ports.

power-inline mode

FTOS allows ports to be configured in one of two modes: **auto** and **static**.

auto: Ports configured for auto mode manage the power allocation by themselves. There is no prior reservation of power made on these ports. When no PD is connected on this port, the power allocated is zero. Once a PD is connected, FTOS detects its PoE class dynamically and the maximum power for its class is allocated to the port. The PD then boots using this allocated power. After bootup, if the PD is LLDP-MED capable, it might send in Extended Power via MDI TLV to the system. In this case, the Dell Force10 switch revises the power allocation to the value that the PD requests via LLDP-MED. The advertised Power Requirement from the PD could be less than or greater than the currently allocated value.

Ports configured for auto mode with the *max_milliwatts* option allocate power the same way, but the allocation never exceeds the specified maximum. If *max_milliwatts* is greater than the PoE class maximum the system allocates only the class maximum. Note that if a PD has class maximum that is greater than *max_milliwatts*, the system allocates no power, and the PD does not power up.

static: Ports configured in static mode reserve a fixed power allocation whether a device is connected or not. By default 15.4W is allocated, but this is user-configurable with the *max_milliwatts* option. No dynamic PoE class detection is performed on static ports, and Extended Power via MDI TLVs have no effect.

Extended Power-via-MDI TLV

The PD sends three pieces of information in the LLDP-MED Extended Power-via-MDI TLV:

1. Power Requirement: FTOS honors this and uses it for power allocation.
2. Power Priority—Critical, High, or Low: FTOS honors this information and uses it for power priority calculation.
3. External Power Source: FTOS does not use this information.

Determine the Affect of a Port on the Power Budget

The PoE power budget is affected differently depending on how PoE is enabled and whether a device is connected:

1. When you configure a port with **power inline auto** *without* the *max_milliwatts* power limit option, power is only allocated after you connect a device to the port.
   - When you connect a device, the maximum power for the device class is allocated if there is sufficient power in the budget. See Table 36-1 on page 785.
   - If there is not enough power in the budget, the configuration is maintained and the port waits for power to become available.
   - If the device advertises its power requirement through LLDP-MED, then FTOS allocates the required amount and returns the remaining amount to the budget.

Note: LLDP-MED TLVs are only honored if the port is configured with **power inline auto** (with or without the *max_milliwatts* option).

2. When you configure a port with **power inline auto** *with* the power limit option *max_milliwatts*, power is only allocated after you connect a device to the port.
   - If the maximum power for the device class is *less* than the power limit you specified, FTOS allocates the required amount and returns the remaining amount to the budget.
   - If there is not enough power in the budget, the configuration is maintained and the port waits for power to become available.
   - If the maximum power for the device class is *more than* than the power limit you specified, FTOS does not allocate any power.

Note: When a port is configured with **power inline auto** (with or without the *max_milliwatts* option) and the PoE device is disconnected, the allocated power is returned to the power budget.

3. When you configure a port with **power inline static** *without* the power limit option (*max_milliwatts*), FTOS allocates 15.4W (subject to availability and priority) to the port whether or not a device is connected.
4. When you configure a port with **power inline static** *with* the power limit option (*max_milliwatts*), FTOS allocates the specified number of Watts.
   - If there is not enough power in the budget, the configuration is maintained and port waits for power to become available.
   - If the maximum power for the device class is *more than* than the power limit you specified, FTOS does not allocate any power.

Monitor the Power Budget

The power budget is the amount of power available from the installed PSUs minus the power required to operate the chassis. Use the **show power inline** (Figure 36-2 on page 788) and **show power detail** (Figure 36-3 on page 788) commands to help you determine if power is available for additional PoE ports (1478.40 Watts are supplied per C-Series PSU; max of 790W on S-Series with load-sharing external DC PSU).

Enabling PoE on more ports than is supported by the power budget produces one of these results:

- If the newly PoE-enabled port has a lower priority, then the command is accepted, but power is not allocated to the port. In this case, the following message is displayed.

Message 1 Insufficient Power to Enable PoE

```
%Warning: Insufficient power to enable. POE oper-status set to OFF for port <linecard/
portnumber>
```

- If the newly PoE-enabled port has a higher priority, then the CLI is accepted, and power is terminated on the lowest priority port in the chassis. If another power supply is added to the system at a later time, both ports receive power.
- If all of the lower priority ports combined cannot meet the power requirements of the newly enabled port, then the CLI is accepted, but power on the lower priority ports is not terminated, and no power is supplied to the port.

The second result in this scenario is true even if a powered device is not connected to the port. Power can be allocated to a port, thus subtracting it from the power budget and making it unavailable to other ports, but that power does not have to be consumed.

Manage Power Priorities

PoE-enabled ports have power access priorities based first on their configuration and then by line card and port number. The default prioritization is presented in Table 36-5.

Note: For S-Series, where Table 36-5 refers to “line cards with the lowest slot number”, substitute “S-Series stack members with the lowest unit ID”.)

Table 36-5. PoE Ports Priorities

Configuration	Port Number	Priority
Ports configured with **power inline static**	Ports with the lowest port numbers in line cards with the lowest slot number	1
	Ports with the lowest port numbers	2
Ports configured with **power inline auto**	Ports with the lowest port numbers in line cards with the lowest slot number	3
	Ports with the lowest port numbers	4

You can augment the default prioritization using the command **[no] power inline priority** {**critical** | **high** | **low**}, where **critical** is the highest priority, and **low** is the lowest. FTOS ignores any LLDP-MED priority on this port if you configure a priority with this command. If you do not configure a port priority with this command, FTOS honors any LLDP-MED priority.

In general, priority is assigned in this order:

1. **power inline** [**static | auto**] setting: **power inline static** ports have a higher priority than **power inline auto** ports
2. **power inline priority** {**critical | high | low**} setting or LLDP-MED TLV, if **power inline priority** is not configured
3. slot ID
4. port ID

Recover from a Failed Power Supply

If ports are PoE-enabled, and a PSU fails, power might be terminated on some ports to compensate for the power loss. This does not affect PoE individual port configurations.

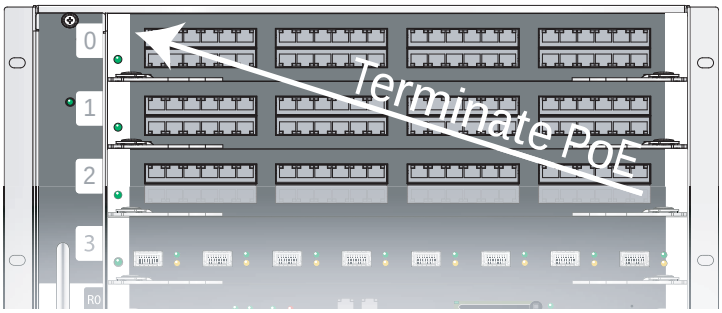
For C-Series, use the **show power supply** command to display PSU status (Figure 36-4).
For S-Series, see the Power over Ethernet (PoE) chapter in the *FTOS Command Reference for the S-Series* for an example of the output of the **show power inline** output and its field descriptions.

Figure 36-4. show power supply Command Example

```
R1#show power supply

Power         Model
Supply        Number             Type      Status
-----------------------------------------------
PS0           --                 --         Absent
PS1           CC-C300-PWR-AC     AC          Active
PS2           CC-C300-PWR-AC     AC         Fail
PS3           CC-C300-PWR-AC     AC         Remote Off
PS4           --                 --         Absent
PS5           --                 --         Absent
PS6           --                 --         Absent
PS7           --                  --        Absent
```

If power must be terminated for some ports, the order in which ports are affected is based on priority. Ports with the lowest priority are terminated first (see Manage Power Priorities on page 792).

Figure 36-5. Order of PoE Termination

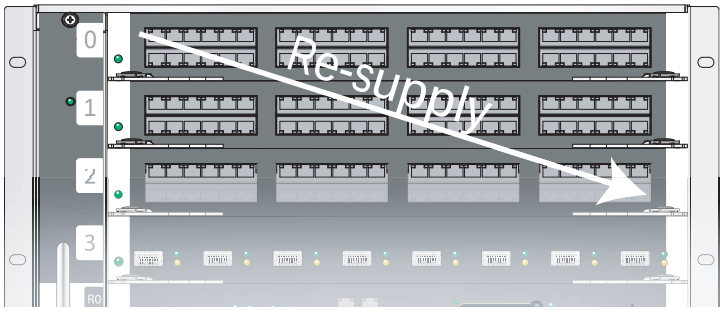


For the configuration in Figure 36-2:

- Power for ports 7/1 and 7/2 is terminated first because it is configured with **inline power auto**.
- Power for port 7/2 is terminated before PoE for port 7/1 because port 7/1 has a lower port number.
- Power for port 7/0 is terminated last because it is configured with **inline power static**.

When a failed PSU is replaced and there is sufficient power for PoE, power is automatically re-supplied for previously configured PoE ports, and power is supplied first to ports with the highest priority.

Figure 36-6. Order of PoE Re-Supply



Power Additional PoE Ports on the S-Series

By default, 320 Watts is available for PoE on the S50V and S25V models of the S-Series. You have the option of enabling more power by connecting the external Dell Force10 DC 470W Redundant Power Supply to the Current Sharing terminal of the S50V and S25V. This power supply is in backup mode by default, but you can use the **power budget stack-unit** command to allow that external power supply to be used for powering PoE ports. 790W is the maximum that you can allocate to PoE, although the combined output of the internal AC power supply and the external DC power supply is 940W. That external power supply is the only power supply that you can use to add power for PoE with the S-Series.

Message 28 appears if you attempt to use the **power budget** command when an external power supply is not connected.

Message 2 External Power Supply Not Found

```
% Error: External power supply not found or incompatible external power supply.
```

Deploying VOIP

VoIP phones on the market today follow the same basic boot and operations process:

1. Wait for an LLDP from the Ethernet switch.
2. Obtain an IP address from a DHCP server.
3. Send an LLDP-MED frame to the switch.
4. Wait for an LLDP-MED frame from the switch and read the Network Policy TLV to get the VLAN ID, Layer 2 Priority, and DSCP value.
5. Download applications and software from the call manager.
6. After configuration, send voice packets as tagged frames and data packets as untagged frames.

Figure 36-7 shows a basic configuration for a deployment in which the end workstation plugs into an IP phone for its Ethernet connection.

Figure 36-7. Office VOIP Deployment

Create VLANs for an Office VOIP Deployment

The phone requires one tagged VLAN for VOIP service and one untagged VLAN for PC data, as shown in Figure 36-7. You may configure voice signaling on the voice VLAN, but some implementations might need an additional tagged VLAN for this traffic; Figure 36-8 adds an additional tagged VLAN for voice signaling. The example is from a C-Series, but an S-Series would be configured in the same way.

Figure 36-8. Creating VLANs for an Office VOIP Deployment

```
FTOS#show running-config interface configured
!
interface GigabitEthernet 6/0
 no ip address
 no shutdown
!
interface GigabitEthernet 6/10
 no ip address
 portmode hybrid
 switchport!
 power inline auto
 no shutdown
!
interface Vlan 100
 description "Data VLAN"
 no ip address
 untagged GigabitEthernet 6/10-11,22-23,46-47
 shutdown
!
interface Vlan 200
 description "Voice VLAN"
 no ip address
 tagged GigabitEthernet 6/10-11,22-23,46-47
 shutdown
!
interface Vlan 300
 description "Voice Signaling VLAN"
 no ip address
 tagged GigabitEthernet 6/10-11,22-23,46-47
 shutdown
```

Configure LLDP-MED for an Office VOIP Deployment

VOIP deployments may optionally use LLDP-MED. LLDP-MED advertises VLAN, dot1P, and DSCP configurations on the switch so that you do not need to manually configure every phone with this information. See Chapter 26, Link Layer Discovery Protocol. Based on the configuration in Figure 36-9, the phone will initiate a DHCP request on the advertised voice VLAN, VLAN 200.

Figure 36-9. LLDP Configuration for Office VOIP Deployment

```
FTOS#show running-config lldp
protocol lldp
 advertise med
 advertise med voice 200 6 46
 advertise med voice-signaling 300 5 28
 no disable
FTOS#show lldp neighbors
 Loc PortID        Rem Chassis Id              Rem Port Id
 ---------------------------------------------------------------------------

 Gi 6/10           0.0.0.0                     001B0CDBA109:P1
 Gi 6/11           0.0.0.0                     001AA2197992:P1
 Gi 6/22           0.0.0.0                     08:00:0f:22:7f:83
 Gi 6/23           0.0.0.0                     08:00:0f:23:de:a9
```

Configure Quality of Service for an Office VOIP Deployment

There are multiple ways you can use QoS to map ingress phone and PC traffic so that you can give them each a different quality of service. See Chapter 41, Quality of Service.

Honor the incoming DSCP value

On both the C-Series or S-Series, if you know traffic originating from the phone is tagged with the DSCP value of 46 (EF), you might make the associated queue a strict priority queue, as shown in Figure 36-10; on the C-Series and S-Series, FTOS maps DSCP 46 to queue 2 (see Table 41-5 on page 865 in the QoS chapter.)

Figure 36-10. Honoring the DSCP Value on Incoming Voice Data

```
FTOS#sh run policy-map-input
!
policy-map-input HonorDSCP
 trust diffserv
FTOS#sh run int gigabitethernet 6/11
!
interface GigabitEthernet 6/11
 description "IP Phone X"
 no ip address
 portmode hybrid
 switchport
 service-policy input HonorDSCP
 power inline auto
 no shutdown
FTOS#sh run | grep strict-priority
strict-priority unicast 2
```

Honor the incoming dot1p value

On the C-Series, if you know traffic originating from the phone is tagged with a dot1p value of 5, you might make the associated queue a strict priority queue, as shown in Figure 36-11; on the C-Series, FTOS maps dot1p priority 5 to queue 2.

Figure 36-11. Honoring the Dot1P Value on Incoming Voice Traffic

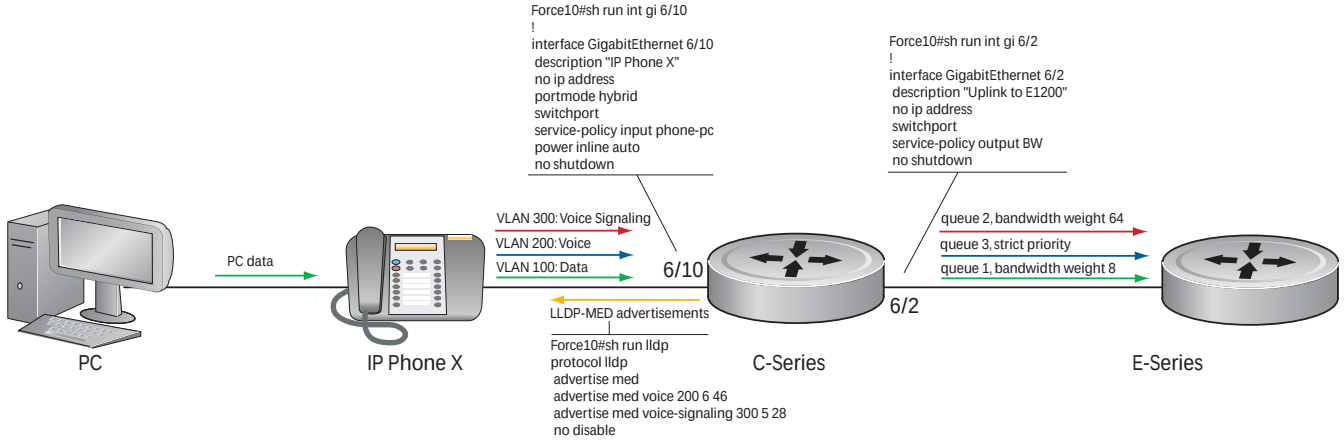
```
FTOS#sh run int gi 6/10
!
interface GigabitEthernet 6/10
 description "IP Phone X"
 no ip address
 portmode hybrid
 switchport
 service-class dynamic dot1p
 power inline auto
 no shutdown
FTOS#sh run | grep strict-priority
strict-priority unicast 2
```

Classifying VOIP traffic and applying QoS policies

Avoid congestion and give precedence to voice and signaling traffic by classifying traffic based on subnet and using strict priority and bandwidth weights on egress, as outlined in the steps below.

Figure 36-12 depicts the topology and shows the configuration for a C-Series. The steps are the same on an S-Series. Figure 36-13 on page 799 is a screenshot showing some of the steps and the resulting running-config.

Figure 36-12. Classifying VOIP Traffic and Applying QoS Policies for an Office VOIP Deployment



Step	Task	Command	Command Mode
1	Create three standard or extended access-lists, one each for voice, voice signaling, and PC data, and place each in its own match-any class-map.	**ip access-list** **class-map match-any**	CONFIGURATION CLASS-MAP
2	Create an input policy-map containing all three class-maps, and assign each class-map a different service queue.	**policy-map-input** **service-queue**	CONFIGURATION POLICY-MAP-IN
3	Create two input QoS policies, one each for PC data and voice signaling. Assign a different bandwidth weight to each policy.	**qos-policy-out** **bandwidth-weight**	CONFIGURATION QOS-POLICY-IN
4	Create an output policy map containing both QoS policies, and assign them to different service queues.	**policy-map-out** **service-queue**	CONFIGURATION POLICY-MAP-OUT
5	Assign a strict priority to unicast traffic in queue 3.	**strict-priority**	CONFIGURATION
6	Apply the input policy map you created in Step 2 to the interface connected to the phone, and apply the output policy map you created in Step 4 to the interface connected your desired next-hop router.	**service-policy**	INTERFACE

Figure 36-13 on page 799 is a screenshot showing some of the steps, above, and the resulting running-config.

Figure 36-13. Classifying VOIP Traffic and Applying QoS Policies for an Office VOIP Deployment

```
FTOS#sh run acl
!
ip access-list extended pc-subnet
 seq 5 permit ip 201.1.1.0/24 any
!
ip access-list extended phone-signalling
 seq 5 permit ip 192.1.1.0/24 host 192.1.1.1
!
ip access-list extended phone-subnet
 seq 5 permit ip 192.1.1.0/24 any
FTOS#sh run class-map
!
class-map match-any pc-subnet
 match ip access-group pc-subnet
!
class-map match-any phone-signalling
 match ip access-group phone-signalling
!
class-map match-any phone-subnet
 match ip access-group phone-subnet
FTOS#sh run policy-map-input
!
policy-map-input phone-pc
 service-queue 1 class-map pc-subnet
 service-queue 2 class-map phone-signalling
 service-queue 3 class-map phone-subnet
FTOS#sh run qos-policy-output
!
qos-policy-output data
 bandwidth-weight 8
!
qos-policy-output signalling
 bandwidth-weight 64
FTOS#sh run policy-map-output
!
policy-map-output BW
 service-queue 1 qos-policy data
 service-queue 2 qos-policy signalling
FTOS#sh run | grep strict-p
strict-priority unicast 3
FTOS#sh run int gi 6/10
!
interface GigabitEthernet 6/10
 description "IP Phone X”
 no ip address
 portmode hybrid
 switchport
 service-policy input phone-pc
 power inline auto
 no shutdown
FTOS#sh run int gi 6/2
!
interface GigabitEthernet 6/2
 description "Uplink to E1200"
 no ip address
 switchport
 service-policy output BW
 no shutdown
```

37

Policy-based Routing

Policy-based Routing is supported on platforms: C E S

PBR is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

PBR is supported on the E-Series TeraScale, C-Series, and S-Series platforms in FTOS 8.4.2.0 and later.

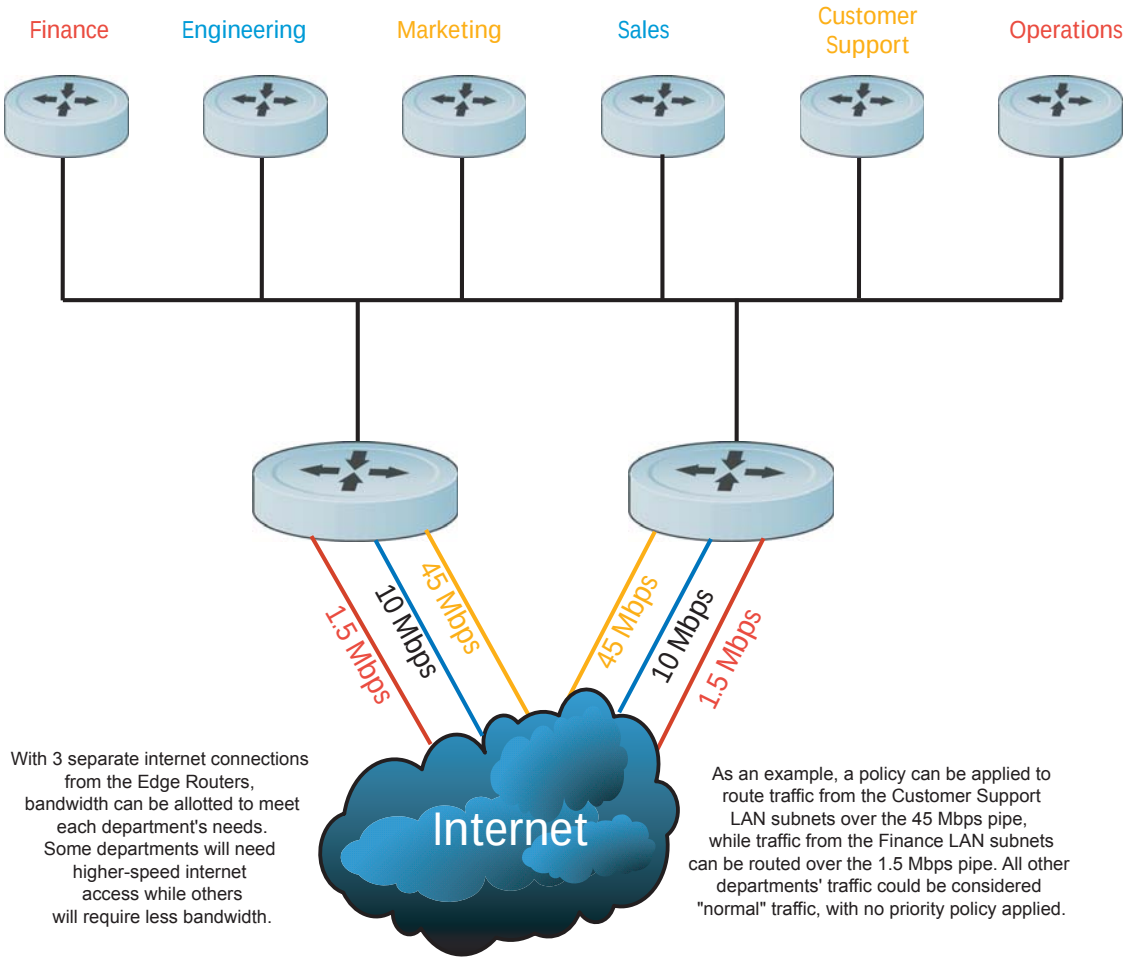
This chapter covers the following topics:

Overview

Policy-based Routing (PBR) enables you to make routing decisions based on policies applied to a specific interface. When a router receives a packet it normally decides where to forward it based on the destination address in the packet, which is used to look up an entry in a routing table. However, in some cases, there may be a need to forward the packet based on other criteria: size, source, protocol type, destination, etc. For example, a network administrator might want to forward a packet that uses TCP across a different next-hop than packets using ICMP.

Rules for PBR can also be a combination of things: when the packet comes from this source and wants to go to that destination then route it to this next-hop or onto that specific interface. This permits routing over different links or towards different networks even while the destination is the same but depending on where the packet originates.

Figure 37-1. PBR Example



To enable a PBR, you create a Redirect List. Redirect lists are defined by rules, or routing policies. The following parameters can be defined in the routing policies or rules:

- IP address of the forwarding router (next-hop IP address)
- Protocol as defined in the header
- Source IP address and mask
- Destination IP address and mask
- Source port
- Destination port
- TCP Flags

Once a redirect-list is applied to an interface, all traffic passing through it is subjected to the rules defined in the redirect-list.

The traffic is forwarded based on the following:

1. Next-hop addresses are verified. If the specified next hop is reachable, then the traffic is forwarded to the specified next-hop.

2. If the specified next-hops are not reachable, then the normal routing table is used to forward the traffic.
3. FTOS supports multiple next-hop entries in the redirect lists.
4. Redirect-Lists are applied at Ingress.

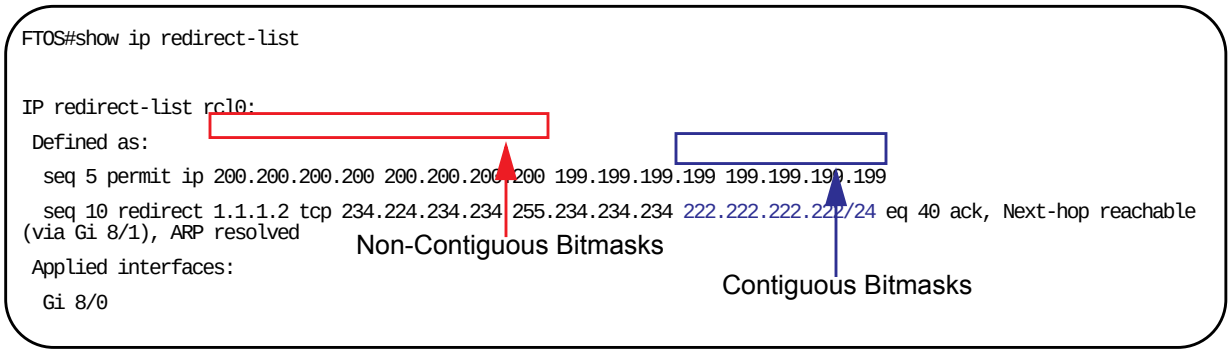
Implementing Policy-based Routing with FTOS

Non-contiguous bitmasks for PBR

Non-contiguous bitmasks for PBR allows more granular and flexible control over routing policies. Network addresses that are in the middle of a subnet can be included or excluded.

Specific bitmasks can be entered using the dotted decimal format.

Figure 37-2. Non-contiguous bitmask example



Hot-Lock PBR

Ingress and egress Hot Lock PBR allow you to add or delete new rules into an existing policy (already written into CAM) without disruption to traffic flow. Existing entries in CAM are adjusted to accommodate the new entries. Hot Lock PBR is enabled by default.

Configuration Task List for Policy-based Routing

To enable the PBR:

1. Create a Redirect List
2. Create a Rule for a Redirect-list
3. Apply a Redirect-list to an Interface using a Redirect-group

Create a Redirect List

Use the following command in CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip redirect-list *redirect-list-name*	CONFIGURATION	Create a redirect list by entering the list name. Format: 16 characters
		Delete the redirect list with the **no ip redirect-list** command.

The following example creates a redirect list by the name of “xyz.”

Figure 37-3. Creating a Redirect List Example

```
FTOS(conf)#ip redirect-list ?
WORD                    Redirect-list name (max 16 chars)
FTOS(conf)#ip redirect-list xyz
```

Create a Rule for a Redirect-list

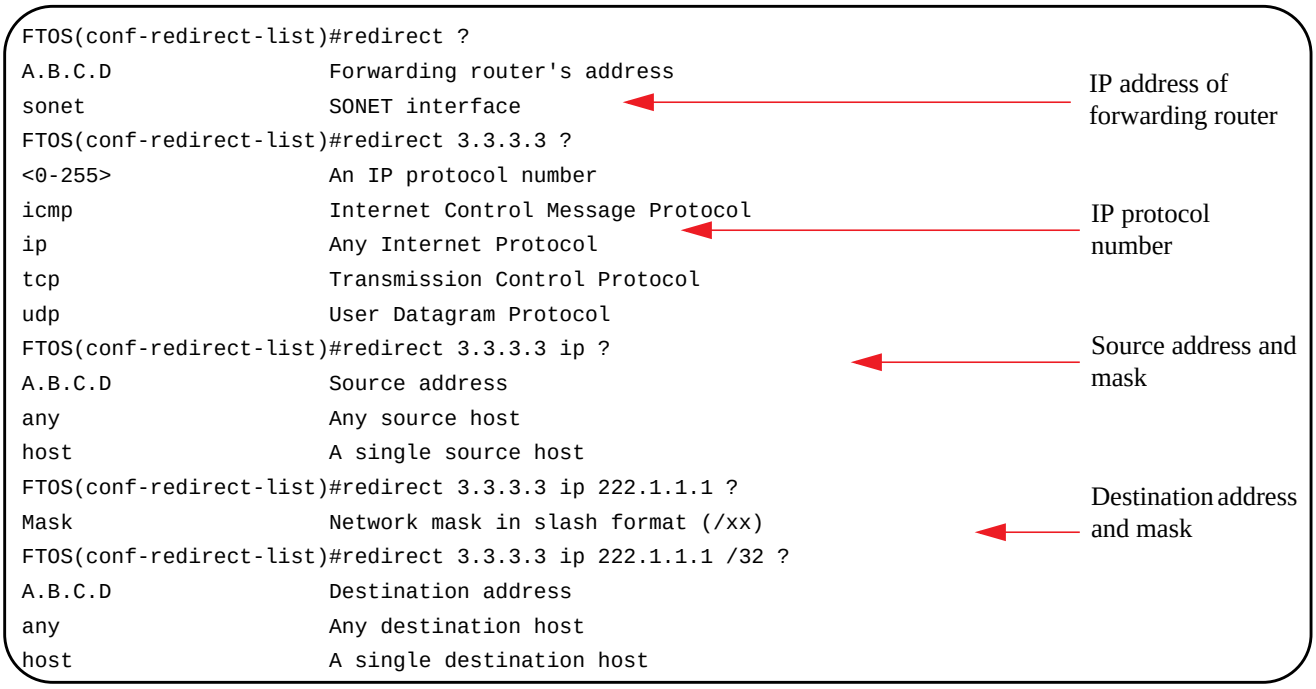
Use the following command in CONFIGURATION REDIRECT-LIST mode to set the rules for the redirect list. You can enter the command multiple times and create a sequence of redirect rules. Use the **seq** *nn* **redirect** version of the command to organize your rules**.**

Command Syntax	Command Mode	Purpose
seq *{number}* **redirect** {*ip-address* | *sonet*} {*ip-protocol-number* | *protocol-type* [*bit*]} {*source mask* | **any** | **host** *ip-address*} {*destination mask* | **any** | **host** *ip-address*}	CONF-REDIRECT-LIST	Configure a rule for the redirect list. *number* is the number in sequence to initiate this rule *ip-address* is the Forwarding router's address FORMAT: A.B.C.D *sonet* is a sonet interface FORMAT: **sonet** slot/port *ip-protocol-number* or *protocol-type* is the type of protocol to be redirected FORMAT: 0-255 for IP protocol number, or enter protocol type *source ip-address* or *any* or *host ip-address* is the Source's IP address FORMAT: A.B.C.D/NN, or ANY or HOST IP address *destination ip-address* or *any* or *host ip-address* is the Destination's IP address FORMAT: A.B.C.D/NN, or ANY or HOST IP address Delete a rule with the **no redirect** command. The redirect rule supports Non-contiguous bitmasks for PBR in the Destination router IP address

Figure 37-4 shows a step-by-step example of how to create a rule for a redirect list by configuring:

- IP address of the next-hop router in the forwarding route
- IP protocol number
- Source address with mask information
- Destination address with mask information.

Figure 37-4. Creating a Rule Example



Multiple rules can be applied to a single redirect-list. The rules are applied in ascending order, starting with the rule that has the lowest sequence number in a redirect-list. Figure 37-5 displays the correct method for applying multiple rules to one list.

Figure 37-5. Creating multiple rules for a redirect-list

```
FTOS(conf)#ip redirect-list test
FTOS(conf-redirect-list)#seq 10 redirect 10.1.1.2 ip 20.1.1.0/24 any
FTOS(conf-redirect-list)#seq 15 redirect 10.1.1.3 ip 20.1.1.0/25 any
FTOS(conf-redirect-list)#seq 20 redirect 10.1.1.3 ip 20.1.1.128/24 any
FTOS(conf-redirect-list)#show config
!
ip redirect-list test
 seq 10 redirect 10.1.1.2 ip 20.1.1.0/24 any
 seq 15 redirect 10.1.1.3 ip 20.1.1.0/25 any
 seq 20 redirect 10.1.1.3 ip 20.1.1.0/24 any
FTOS(conf-redirect-list)#
```

Note: Starting in release 8.4.1.2, FTOS supports the use of multiple recursive routes with the same source-address and destination-address combination in a redirect policy on an E-Series ExaScale router.

A recursive route is a route for which the immediate next-hop address is learned dynamically through a routing protocol and acquired through a route lookup in the routing table.You can configure multiple recursive routes in a redirect list by entering multiple **seq redirect** statements with the same source and destination address and specify a different next-hop IP address. In this way, the recursive routes are used as different forwarding routes for dynamic failover. If the primary path goes down and the recursive route is removed from the routing table, the **seq redirect** statement is ignored and the next statement in the list with a different route is used.

PBR Exceptions (Permit)

Use the command **permit** to create an exception to a redirect list. Exceptions are used when a forwarding decision should be based on the routing table rather than a routing policy.

FTOS assigns the first available sequence number to a rule configured without a sequence number and inserts the rule into the PBR CAM region next to the existing entries. Since the order of rules is important, ensure that you configure any necessary sequence numbers.

In Figure 37-6, the permit statement is never applied because the redirect list covers all source and destination IP addresses.

Figure 37-6. Ineffective PBR Exception due to Low Sequence Number

```
ip redirect-list rcl0
seq 5 redirect 2.2.2.2 ip any any
seq 10 permit ip host 3.3.3.3 any
```

To ensure that the permit statement or PBR exception is effective, use a lower sequence number, as shown in Figure 37-7.

Figure 37-7. Effective PBR Exception due to Proper Sequencing

```
ip redirect-list rcl0
seq 10 permit ip host 3.3.3.3 any
seq 15 redirect 2.2.2.2 ip any any
```

Apply a Redirect-list to an Interface using a Redirect-group

IP redirect lists are supported on physical interfaces as well as VLAN and port-channel interfaces.

Note: When you apply a redirect-list on a port-channel on the E-Series, when traffic is redirected to the next hop and the destination port-channel is shut down, the traffic is dropped. However, on the C-Series, the traffic redirected to the destination port-channel is sometimes switched.

Use the following command in INTERFACE mode to apply a redirect list to an interface. Multiple redirect-lists can be applied to a redirect-group. It is also possible to create two or more redirect-groups on one interface for backup purposes.

Command Syntax	Command Mode	Purpose
ip redirect-group *redirect-list-name*	INTERFACE	Apply a redirect list (policy-based routing) to an interface. *redirect-list-name* is the name of a redirect list to apply to this interface. FORMAT: up to 16 characters Delete the redirect list from this interface with the **[no] ip redirect-group** command.

In this example, the list “xyz” is applied to the GigabitEthernet 4/0 interface.

Figure 37-8. Applying a Redirect-list to an Interface Example

```
FTOS(conf-if-gi-4/0)#ip redirect-group xyz
FTOS(conf-if-gi-4/0)#
```

Figure 37-9. Applying multiple Redirect-groups to an interface

```
FTOS(conf-if-gi-1/0)#ip redirect-group test
FTOS(conf-if-gi-1/0)#ip redirect-group xyz
FTOS(conf-if-gi-1/0)#show config
!
interface GigabitEthernet 1/0
 no ip address
 ip redirect-group test
 ip redirect-group xyz
 shutdown
FTOS(conf-if-gi-1/0)#
```

In addition to supporting multiple redirect-lists in a redirect-group, multiple redirect-groups are supported on a single interface. FTOS has the capability to support multiple groups on an interface for backup purposes.

Show Redirect List Configuration

To view the configuration redirect list configuration, use the following command in EXEC mode:

Command Syntax	Command Mode	Purpose
show ip redirect-list *redirect-list-name*	EXEC	View the redirect list configuration and the associated interfaces.
show cam pbr **show cam-usage**	EXEC	View the redirect list entries programmed in the CAM.

List the redirect list configuration using the **show ip redirect-list redirect-list-name** command.

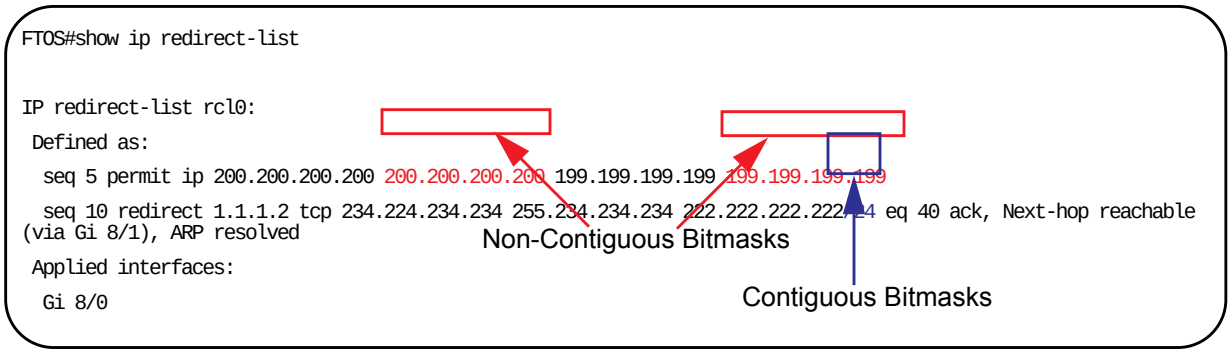
The non-contiguous mask is displayed in dotted format (x.x.x.x). The contiguous mask is displayed in /x format. Some sample outputs are shown in Figure 37-10, Figure 37-11, and Figure 37-12.

Figure 37-10. Showing Redirect List Configuration Example 1

```
FTOS#show ip redirect-list xyz

IP redirect-list xyz:
 Defined as:
  seq 5 redirect 3.3.3.3 ip host 222.1.1.1 host 77.1.1.1
```

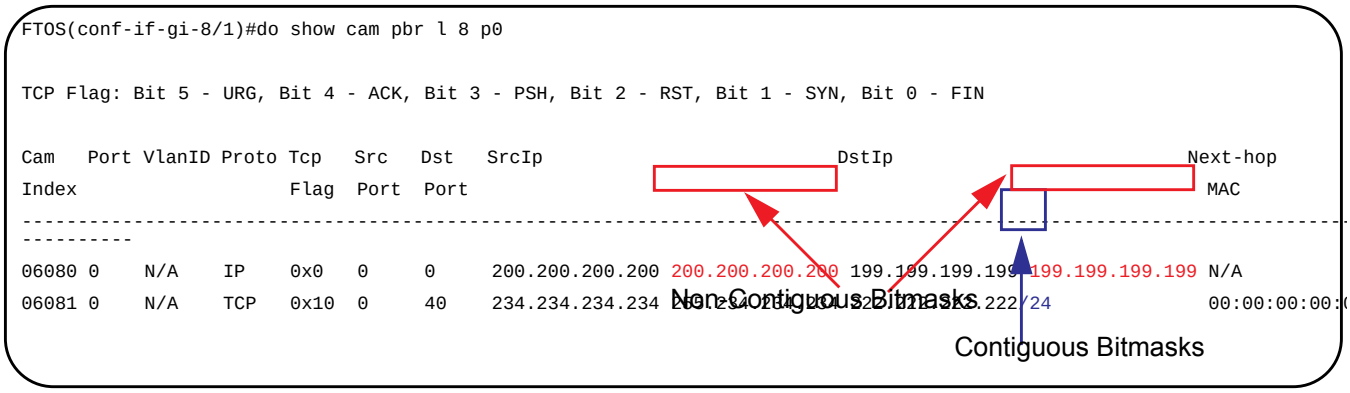
Use the **show ip redirect-list** (without the list name) to display all the redirect-lists configured on the device.

Figure 37-11. Showing Redirect List Configuration Example 2



Note: If, the redirect-list is applied to an interface, the output of **show ip redirect-list redirect-list-name** command displays reachability and ARP status for the specified next-hop.

Figure 37-12. Showing CAM PBR Configuration Example



Sample Configuration

The following configuration is an example for setting up a PBR. These are not comprehensive directions. They are intended to give you a some guidance with typical configurations.

You can copy and paste from these examples to your CLI. Be sure you make the necessary changes to support your own IP Addresses, Interfaces, Names, etc.

Figure 37-13 is a graphic illustration of the configuration shown in Figure 37-14. The Redirect-List GOLD defined in this example, creates the following rules:

- description Route Gold traffic to the DS3.
- seq 5 redirect 10.99.99.254 ip 192.168.1.0/24 any
  - " Redirect to next-hop router IP 10.99.99.254 any traffic originating in 192.168.1.0/24."
- seq 10 redirect 10.99.99.254 ip 192.168.2.0/24 any
  - " Redirect to next-hop router IP 10.99.99.254 any traffic originating in 192.168.2.0/24."
- seq 15 permit ip any any

Figure 37-13. PBR Sample Illustration

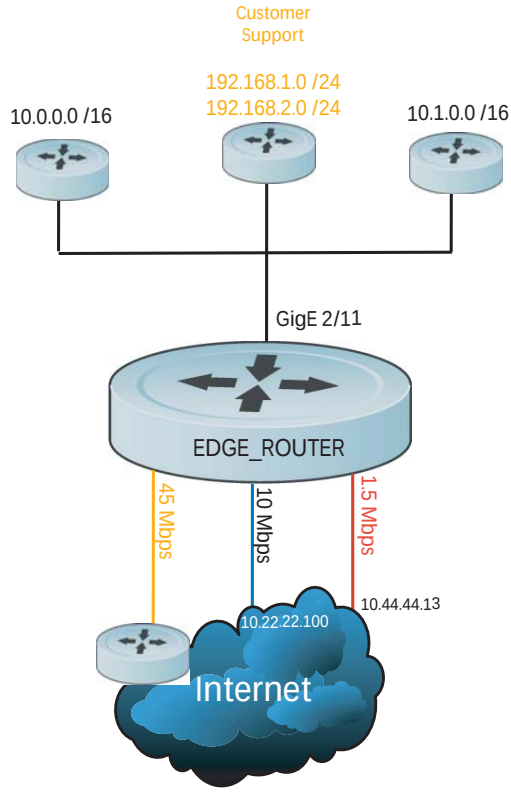


Figure 37-14. PBR Sample Configuration

Create the Redirect-List GOLD.

```
EDGE_ROUTER(conf-if-gi-3/23)#ip redirect-list GOLD
EDGE_ROUTER(conf-redirect-list)#description Route GOLD traffic to ISP_GOLD.
EDGE_ROUTER(conf-redirect-list)#$direct 10.99.99.254 ip 192.168.1.0/24 any
EDGE_ROUTER(conf-redirect-list)#$redirect 10.99.99.254 ip 192.168.2.0/24 any
EDGE_ROUTER(conf-redirect-list)# seq 15 permit ip any any
EDGE_ROUTER(conf-redirect-list)#sho config
!
ip redirect-list GOLD
 description Route GOLD traffic to ISP_GOLD.
 seq 5 redirect 10.99.99.254 ip 192.168.1.0/24 any
 seq 10 redirect 10.99.99.254 ip 192.168.2.0/24 any
 seq 15 permit ip any any
```

Assign Redirect-List GOLD to Interface 2/11.

```
EDGE_ROUTER(conf)#int gig 2/11
EDGE_ROUTER(conf-if-gi-2/11)#ip add 192.168.3.2/24
EDGE_ROUTER(conf-if-gi-2/11)#no shut
00:09:47: %RPM0-P:CP %IFMGR-5-ASTATE_UP: Changed interface Admin state to up:
Gi 2/11
EDGE_ROUTER(conf-if-gi-2/11)#
EDGE_ROUTER(conf-if-gi-2/11)#
EDGE_ROUTER(conf-if-gi-2/11)#ip redirect-group GOLD
EDGE_ROUTER(conf-if-gi-2/11)#no shut
EDGE_ROUTER(conf-if-gi-2/11)#end
EDGE_ROUTER(conf-redirect-list)#end
EDGE_ROUTER#00:08:23: %RPM0-P:CP %SYS-5-CONFIG_I: Configured from console by
console

EDGE_ROUTER#
```

View Redirect-List GOLD.

```
EDGE_ROUTER#show ip redirect-list

IP redirect-list GOLD:
 Defined as:
  seq 5 redirect 10.99.99.254 ip 192.168.1.0/24 any, Next-hop reachable (via
Gi 3/23), ARP resolved
  seq 10 redirect 10.99.99.254 ip 192.168.2.0/24 any, Next-hop reachable (via
Gi 3/23), ARP resolved
  seq 15 permit ip any any
 Applied interfaces:
  Gi 2/11
EDGE_ROUTER#
```

38

Port Monitoring

Port Monitoring is supported on platforms: C E S

Port Monitoring is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

Port Monitoring is a feature that copies all incoming or outgoing packets on one port and forwards (mirrors) them to another port. The source port is the monitored port (MD) and the destination port is the monitoring port (MG). Port Monitoring functionality is different between platforms, but the behavior is the same, with highlighted exceptions.

This chapter is divided into the following sections:

Important Points to Remember

- Port Monitoring is not supported on EtherScale versions of the E-Series platform.
- Port Monitoring is supported on physical ports only; VLAN and port-channel interfaces do not support port monitoring.
- A SONET port may only be a monitored port.
- The Monitored (source, “MD”) and Monitoring ports (destination, “MG”) must be on the same switch.
- In general, a monitoring port should have **no ip address** and **no shutdown** as the only configuration; FTOS permits a limited set of commands for monitoring ports; display them using the command **?**. A monitoring port also may not be a member of a VLAN.
- There may only be one destination port in a monitoring session.

- A source port (MD) can only be monitored by one destination port (MG). The following error is displayed if you try to assign a monitored port to more than one monitoring port.

```
FTOS(conf)#mon ses 1
FTOS(conf-mon-sess-1)#$gig 0/0 destination gig 0/60 direction both
FTOS(conf-mon-sess-1)#do show mon ses
     SessionID       Source       Destination     Direction      Mode        Type
     ---------       ------       -----------     ---------      ----        ----
             1       Gi 0/0       Gi 0/60         both           interface   Port-based
FTOS(conf-mon-sess-1)#mon ses 2
FTOS(conf-mon-sess-2)#source gig 0/0 destination gig 0/61 direction both
% Error: MD port is already being monitored.
```

- The C-Series and S-Series may only have four destination ports per port-pipe. There is no limitation on the total number of monitoring sessions.

Table 38-1 lists the maximum number of monitoring sessions per system. For the C-Series and S-Series, the total number of sessions is derived by consuming a unique destination port in each session, in each port-pipe.

Table 38-1. Maximum Number of Monitoring Sessions per System

System	Maximum Sessions	System	Maximum Sessions
C150	∞ (Note)	E1200/E1200i (TeraScale)	28
C300	∞ (Note)	E1200i (ExaScale)	∞
S50V, S50N	∞ (Note)	E600/E600i (TeraScale)	14
S25P	∞ (Note)	E600i (ExaScale)	∞
		E300	6

Note: On the C-Series and S-Series, there is no limit to the number of monitoring sessions per system, provided that there are only 4 destination ports per port-pipe. If each monitoring session has a unique destination port, then the maximum number of session is 4 per port-pipe.

Port Monitoring on E-Series

Both the E-Series TeraScale and E-Series ExaScale support the following.

- FTOS supports one destination (MG) port per monitoring session. The same destination port (MG) can be used in another monitoring session.
- One destination (MG) port can monitor up to 28 source (MD) ports.
- A port cannot be defined as both a source (MD) and a destination (MG) port (Message 1).

Message 1 Cannot define source (MD) and destination (MG) on same port

```
% Error: MD port is already being monitored.
```

E-Series TeraScale

The E-Series TeraScale system supports 1 monitoring session per port-pipe. E-Series TeraScale supports a maximum of 28 port pipes.

On the E-Series TeraScale, FTOS supports a single source-destination statement in a monitor session (Message 2). E-Series TeraScale supports only one source and one destination port per port-pipe (Message 3). Therefore, the E-Series TeraScale supports as many monitoring sessions as there are port-pipes in the system.

Message 2 Multiple Source-Destination Statements Error Message on E-Series TeraScale

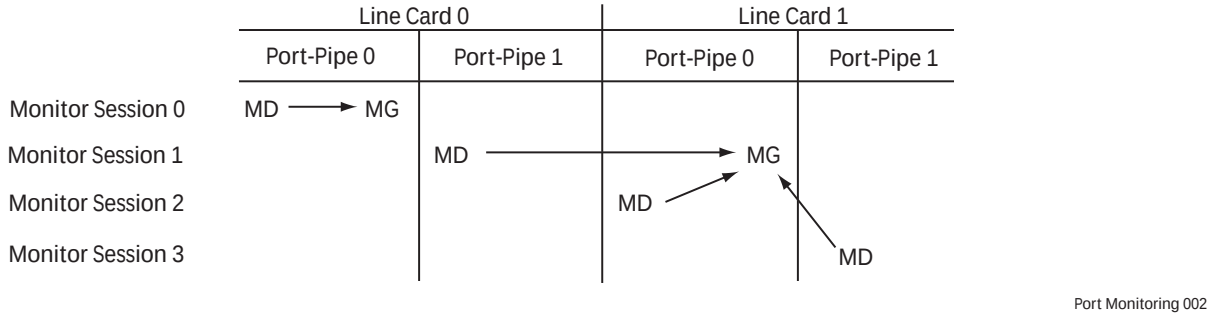
```
% Error: Remove existing monitor configuration.
```

Message 3 One Source/Destination Port per Port-pipe Error Message on E-Series TeraScale

```
% Error: Some port from this port pipe is already configured as MD.
% Error: Some port from this port pipe is already configured as MG.
```

Figure 38-1 illustrates a possible port monitoring configuration on the E-Series.

Figure 38-1. Port Monitoring Configurations on the E-Series



E-Series ExaScale

FTOS on E-Series ExaScale supports a single destination (MG) port monitoring multiple multiple source (MD) ports in one monitor session. One monitor session can have only one destination (MG) port. The same destination (MG) port can be uses with multiple monitoring sessions.

There is no restriction on the number of source (MD) or destination (MG) ports on the chassis because there is no port-pipe restriction on the E-Series ExaScale system.

There is no restriction to the number of monitoring sessions supported on the E-Series ExaScale system.

Port Monitoring on C-Series and S-Series

The C-Series and S-Series support multiple source-destination statements in a monitor session, but there may only be one destination port in a monitoring session (Message 4).

Message 4 One Destination Port in a Monitoring Session Error Message on C-Series and S-Series

```
% Error: Only one MG port is allowed in a session.
```

The number of source ports FTOS allows within a port-pipe is equal to the number of physical ports in the port-pipe (n). However, n number of ports may only have four different destination ports (Message 5).

Figure 38-2. Number of Monitoring Ports on the C-Series and S-Series

```
FTOS#show mon session
     SessionID       Source       Destination     Direction      Mode        Type
     ---------       ------       -----------     ---------      ----        ----
             0       Gi 0/13      Gi 0/1          rx             interface   Port-based
            10       Gi 0/14      Gi 0/2          rx             interface   Port-based
            20       Gi 0/15      Gi 0/3          rx             interface   Port-based
            30       Gi 0/16      Gi 0/37         rx             interface   Port-based
FTOS(conf)#mon ses 300
FTOS(conf-mon-sess-300)#source gig 0/17 destination gig 0/4 direction tx
% Error: Exceeding max MG ports for this MD port pipe.
FTOS(conf-mon-sess-300)#
FTOS(conf-mon-sess-300)#source gig 0/17 destination gig 0/1 direction tx
FTOS(conf-mon-sess-300)#do show mon session
     SessionID       Source       Destination     Direction      Mode        Type
     ---------       ------       -----------     ---------      ----        ----
             0       Gi 0/13      Gi 0/1          rx             interface   Port-based
            10       Gi 0/14      Gi 0/2          rx             interface   Port-based
            20       Gi 0/15      Gi 0/3          rx             interface   Port-based
            30       Gi 0/16      Gi 0/37         rx             interface   Port-based
           300       Gi 0/17      Gi 0/1          tx             interface   Port-based
FTOS(conf-mon-sess-300)#
```

In Figure 38-2, ports 0/13, 0/14, 0/15, and 0/16 all belong to the same port-pipe. They are pointing to four different destinations (0/1, 0/2, 0/3, and 0/37). Now it is not possible for another source port from the same port-pipe (for example, 0/17) to point to another new destination (for example, 0/4). If you attempt to configure another destination, Message 5 appears. However, you can configure another monitoring session that uses one of previously used destination ports, as shown in Figure 38-3.

Figure 38-3. Number of Monitoring Ports on the C-Series and S-Series

```
FTOS(conf)#mon ses 300
FTOS(conf-mon-sess-300)#source gig 0/17 destination gig 0/4 direction tx
% Error: Exceeding max MG ports for this MD port pipe.
FTOS(conf-mon-sess-300)#
FTOS(conf-mon-sess-300)#source gig 0/17 destination gig 0/1 direction tx
FTOS(conf-mon-sess-300)#do show mon session
     SessionID       Source       Destination     Direction      Mode        Type
     ---------       ------       -----------     ---------      ----        ----
             0       Gi 0/13      Gi 0/1          rx             interface   Port-based
            10       Gi 0/14      Gi 0/2          rx             interface   Port-based
            20       Gi 0/15      Gi 0/3          rx             interface   Port-based
            30       Gi 0/16      Gi 0/37         rx             interface   Port-based
           300       Gi 0/17      Gi 0/1          tx             interface   Port-based
```

In Figure 38-4, 0/25 and 0/26 belong to Port-pipe 1. This port-pipe again has the same restriction of only four destination ports, new or used.

Figure 38-4. Number of Monitoring Ports on the C-Series and S-Series

```
FTOS(conf-mon-sess-300)#do show mon session
     SessionID       Source       Destination     Direction      Mode        Type
     ---------       ------       -----------     ---------      ----        ----
             0       Gi 0/13      Gi 0/1          rx             interface   Port-based
            10       Gi 0/14      Gi 0/2          rx             interface   Port-based
            20       Gi 0/15      Gi 0/3          rx             interface   Port-based
            30       Gi 0/16      Gi 0/37         rx             interface   Port-based
           100       Gi 0/25      Gi 0/38         tx             interface   Port-based
           110       Gi 0/26      Gi 0/39         tx             interface   Port-based
           300       Gi 0/17      Gi 0/1          tx             interface   Port-based
FTOS(conf-mon-sess-300)#
```

A source port may only be monitored by one destination port (Message 6), but a destination port may monitor more than one source port. Given these parameters, Figure 38-1 illustrates conceptually the possible port monitoring configurations on the C-Series and S-Series.

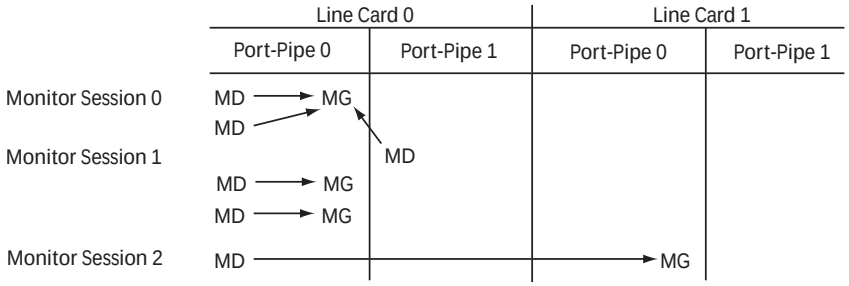
Message 5 One Destination Port in a Monitoring Session Error Message on C-Series and S-Series

```
% Error: Exceeding max MG ports for this MD port pipe.
```

Message 6 One Destination Port per Source Port Error Message

```
% Error: MD port is already being monitored.
```

Figure 38-5. Port Monitoring Configurations on the C-Series and S-Series



FTOS Behavior: On the C-Series and S-Series, all monitored frames are tagged if the configured monitoring direction is transmit (TX), regardless of whether the monitored port (MD) is a Layer 2 or Layer 3 port. If the MD port is a Layer 2 port, the frames are tagged with the VLAN ID of the VLAN to which the MD belongs. If the MD port is a Layer 3 port, the frames are tagged with VLAN ID 4095. If the MD port is in a Layer 3 VLAN, the frames are tagged with the respective Layer 3 VLAN ID. For example, in the configuration *source gig 6/0 destination gig 6/1 direction tx*, if the MD port gigabitethernet 6/0 is an untagged member of any VLAN, all monitored frames that the MG port gigabitethernet 6/1 receives are tagged with the VLAN ID of the MD port. Similarly, if BPDUs are transmitted, the MG port receives them tagged with the VLAN ID 4095. This behavior might result in a difference between the number of egress packets on the MD port and monitored packets on the MG port.

FTOS Behavior: The C-Series and S-Series continue to mirror outgoing traffic even after an MD participating in Spanning Tree Protocol transitions from the forwarding to blocking.

Configuring Port Monitoring

To configure port monitoring:

Step	Command Syntax	Command Mode	Task
1	**show interface**	EXEC Privilege	Verify that the intended monitoring port has no configuration other than **no shutdown**, as shown in Figure 38-6.
2	**monitor session**	CONFIGURATION	Create a monitoring session using the command monitor session from CONFIGURATION mode, as shown in Figure 38-6.
3	**source**	MONITOR SESSION	Specify the source and destination port and direction of traffic, as shown in Figure 38-6.

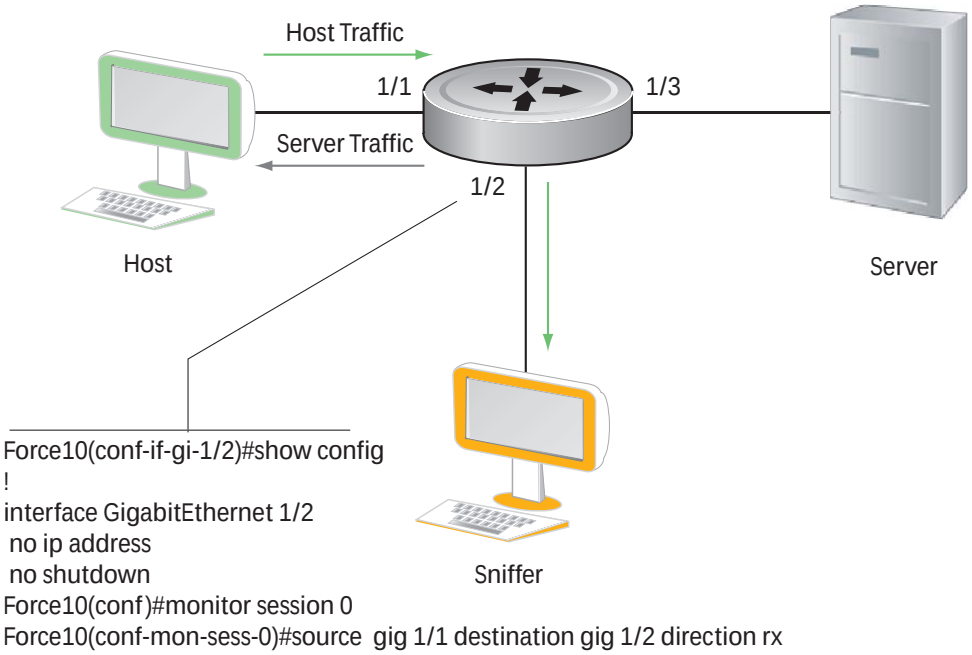
Display monitor sessions using the command **show monitor session** from EXEC Privilege mode, as shown in Figure 38-6.

Figure 38-6. Configuring Port-based Monitoring

```
FTOS(conf-if-gi-1/2)#show config
!
interface GigabitEthernet 1/2
 no ip address
 no shutdown
FTOS(conf-if-gi-1/2)#exit
FTOS(conf)#monitor session 0
FTOS(conf-mon-sess-0)#source gig 1/1 dest gig 1/2 direction rx
FTOS(conf-mon-sess-0)#exit
FTOS(conf)#do show monitor session 0
     SessionID       Source       Destination     Direction      Mode        Type
     ---------       ------       -----------     ---------      ----        ----
             0       Gi 1/1       Gi 1/2          rx             interface   Port-based
FTOS(conf)#
```

In Figure 38-7, the host and server are exchanging traffic which passes through interface gigabitethernet 1/1. Interface gigabitethernet 1/1 is the monitored port and gigabitethernet 1/2 is the monitoring port, which is configured to only monitor traffic received on gigabitethernet 1/1 (host-originated traffic).

Figure 38-7. Port Monitoring Example



Port Monitoring 001

Flow-based Monitoring

Flow-based Monitoring is supported only on platform E

Flow-based monitoring conserves bandwidth by monitoring only specified traffic instead all traffic on the interface. This feature is particularly useful when looking for malicious traffic. It is available for Layer 2 and Layer 3 ingress and egress traffic. You may specify traffic using standard or extended access-lists.

To configure flow-based monitoring:

Step	Command Syntax	Command Mode	Task
1	**flow-based enable**	MONITOR SESSION	Enable flow-based monitoring for a monitoring session.
2	**ip access-list**	CONFIGURATION	Define in an access-list rules that include the keyword **monitor**. FTOS only considers for port monitoring traffic matching rules with the keyword **monitor**. See Chapter 8, IP Access Control Lists (ACL), Prefix Lists, and Route-maps.
3	**ip access-group access-list**	INTERFACE	Apply the ACL to the monitored port. See Chapter 8, IP Access Control Lists (ACL), Prefix Lists, and Route-maps.

View an access-list that you applied to an interface using the command **show ip accounting access-list** from EXEC Privilege mode, as shown in Figure 38-8.

Figure 38-8. Configuring Flow-based Monitoring

```
FTOS(conf)#monitor session 0
FTOS(conf-mon-sess-0)#flow-based enable
FTOS(conf)#ip access-list ext testflow
FTOS(config-ext-nacl)#seq 5 permit icmp any any count bytes monitor
FTOS(config-ext-nacl)#seq 10 permit ip 102.1.1.0/24 any count bytes monitor
FTOS(config-ext-nacl)#seq 15 deny udp any any count bytes
FTOS(config-ext-nacl)#seq 20 deny tcp any any count bytes
FTOS(config-ext-nacl)#exit
FTOS(conf)#interface gig 1/1
FTOS(conf-if-gi-1/1)#ip access-group testflow in
FTOS(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
 ip address 10.11.1.254/24
 ip access-group testflow in
 shutdown
FTOS(conf-if-gi-1/1)#exit
FTOS(conf)#do show ip accounting access-list testflow
!
Extended Ingress IP access list testflow on GigabitEthernet 1/1
Total cam count 4
 seq 5 permit icmp any any monitor count bytes (0 packets 0 bytes)
 seq 10 permit ip 102.1.1.0/24 any monitor count bytes (0 packets 0 bytes)
 seq 15 deny udp any any count bytes (0 packets 0 bytes)
 seq 20 deny tcp any any count bytes (0 packets 0 bytes)
FTOS(conf)#do show monitor session 0
     SessionID       Source       Destination     Direction      Mode        Type
     ---------       ------       -----------     ---------      ----        ----
             0       Gi 1/1       Gi 1/2          rx             interface   Flow-based
```

Remote Port Mirroring

Remote Port Mirroring is supported on platforms: E(X).

While local port monitoring allows you to monitor traffic from one or more source ports by directing it to a destination port on the same switch/router, remote port mirroring allows you to monitor Layer 2 and Layer 3 ingress traffic on multiple source ports on different switches and forward the mirrored traffic to multiple destination ports on different switches. Remote port mirroring helps network administrators monitor and analyze traffic to troubleshoot network problems in a time-saving and efficient way.

In a remote-port mirroring session, monitored traffic is tagged with a VLAN ID and switched on a user-defined, non-routable L2 VLAN. The VLAN is reserved in the network to carry only mirrored traffic, which is forwarded on all egress ports of the VLAN. Each intermediate switch that participates in the transport of mirrored traffic must be configured with the reserved L2 VLAN. Remote port mirroring supports mirroring sessions in which multiple source and destination ports are distributed across multiple switches.

Remote Port Mirroring Example

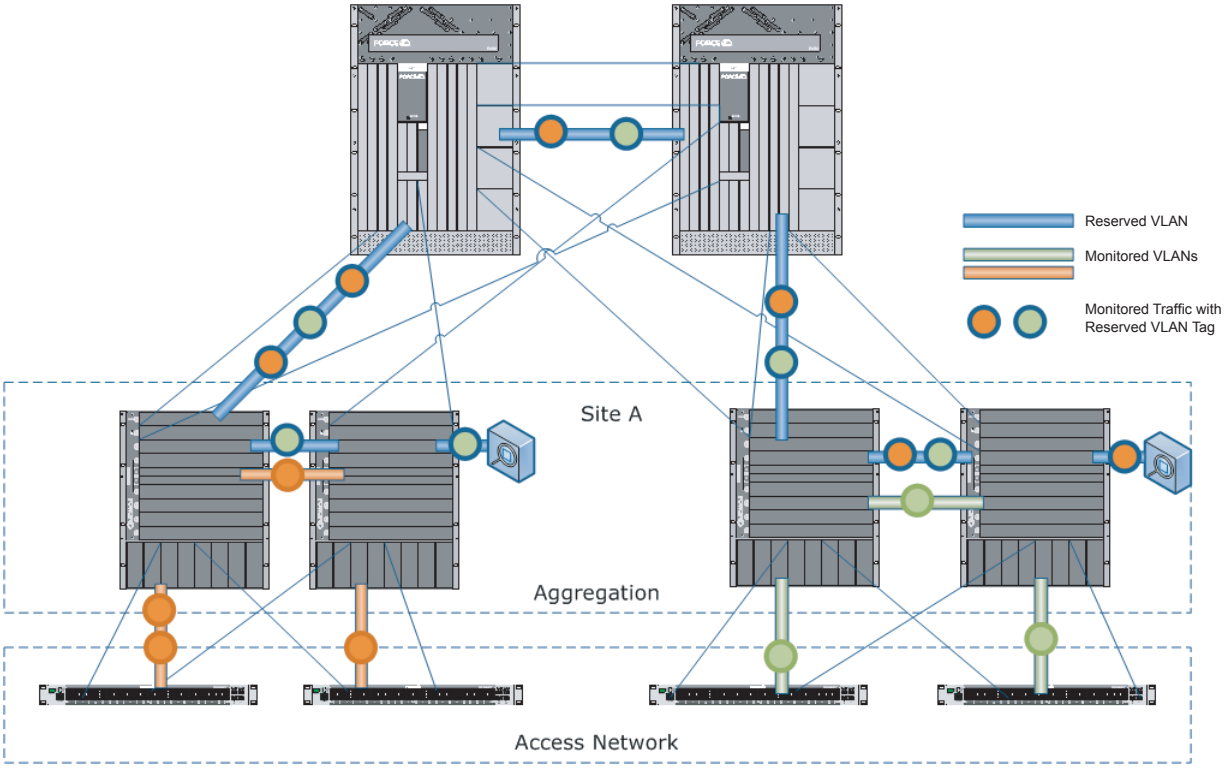
Figure 38-9 shows an example of how remote port mirroring works.

Remote port mirroring uses the analyzers shown in the aggregation network in Site A.

The VLAN traffic on monitored links from the access network is tagged and assigned to a dedicated L2 VLAN. Monitored links are configured in two source sessions shown with orange and green circles. Each source session uses a separate reserved VLAN to transmit mirrored packets (mirrored source-session traffic is shown with an orange or green circle with a blue border).

The reserved VLANs transport the mirrored traffic in sessions (blue pipes) to the destination analyzers in the local network. Two destination sessions are shown: one for the reserved VLAN that transports orange-circle traffic; one for the reserved VLAN that transports green-circle traffic.

Figure 38-9. Remote Port Mirroring



Configuring Remote Port Mirroring

Remote port mirroring requires a source session (monitored ports on different source switches), a reserved tagged VLAN for transporting mirrored traffic (configured on source, intermediate, and destination switches), and a destination session (destination ports connected to analyzers on destination switches).

Configuration Notes

When you configure remote port mirroring, the following conditions apply:

- You can configure any switch in the network with source ports and destination ports, and allow it to function in an intermediate transport session for a reserved VLAN at the same time for multiple remote-port mirroring sessions. You can enable and disable individual mirroring sessions.
- BPDU monitoring is not required to use remote port mirroring.
- A remote port mirroring session mirrors monitored traffic by prefixing the reserved VLAN tag to monitored packets so that they are copied to the reserve VLAN.

  Mirrored traffic is transported across the network using 802.1Q-in-802.1Q tunneling. The source address, destination address and original VLAN ID of the mirrored packet are preserved with the tagged VLAN header. Untagged source packets are tagged with the reserve VLAN ID.
- In the reserved L2 VLAN used for remote port mirroring:
  - MAC address learning in the reserved VLAN is automatically disabled.
  - The reserved VLAN for remote port mirroring can be automatically configured in intermediate switches by using GVRP.
  - There is no restriction on the VLAN IDs used for the reserved remote-mirroring VLAN. Valid VLAN IDs are from 1 to 4094. The default VLAN ID is not supported.
- In a source session used for remote port mirroring:
  - Maximum number of source sessions supported on a switch: 4

    Important: FTOS supports a maximum of four source sessions for remote port mirroring and a maximum of two source sessions per datapath. To use a third or fourth source session on a switch, you must configure the session on a different datapath. Datapaths on different line cards are defined as follows:

    On a 50-port 1G line card, the datapath size is 5 ports; for example, ports 1/0 to 1/4 are the first datapath; 1/5-1/9 are the second datapath, and so on.

    On a 90-port 1G line card, the datapath size is 10 ports; for example, ports 1/0 to 1/9 are the first datapath; 1/10-1/19 are the second datapath, and so on.

    On a 10-port 10G line card, no datapath size is implemented; four source sessions are supported on the line card.

    On a 40-port 10G line card, the datapath size is 4 ports; for example, ports 1/0-1/3 are the first datapath; 1/4-1/7 are the second datapath, and so on.
  - Maximum number of source ports supported on a switch: 128

    You can configure physical ports, port-channels, and VLANs as sources in remote port mirroring and use them in the same source session. You can use both Layer 2 (configured with the **switchport** command) and Layer 3 ports as source ports.

    When you configure a port channel or VLAN in a source session, all ports in the port channel or VLAN are used as source ports, up to a maximum of 128 source ports.

    You can configure trunk ports and access ports as source ports.

    You can configure trunk ports and non-trunk ports as source ports in a remote-port mirroring session.

    You can use the default VLAN and native VLANs as a source VLAN. You cannot configure the dedicated VLAN used to transport mirrored traffic as a source VLAN.

    A destination port for remote port mirroring cannot be used as a source port, including the session in which the port functions as the destination port. A source port channel or source VLAN, which has a member port that is configured as a destination port, cannot be used as a source port channel or source VLAN.

    A VLAN cannot be used as a source VLAN for remote port mirroring if:

- The VLAN consists of more than 128 ports.
- You add a port to a VLAN, which has already been configured in a source session, and the newly added port exceeds the 128-port limit.
- You configure a range of VLANs in a source session and the combined number of ports in the VLANs exceeds 128.

  - You can use ACLs on a source port. In a flow-based source session, packets sent from the RPM are not monitored.
  - Rate-limiting tagged-VLAN egress traffic on a source port is supported.
- In a destination session used for remote port mirroring:
  - Maximum number of destination sessions supported on a switch: 64
    Maximum number ports supported in a destination session: 64
  - You can configure any port as a destination port. A port-channel interface is not supported as a destination port.
  - You can configure additional destination ports in an active session.
  - You can tunnel the mirrored traffic from multiple remote-port source sessions to the same destination port.
  - You can configure a destination port to send only tagged or untagged traffic to the analyzer. By default, the port sends untagged packets so that the reserved VLAN ID is removed and the original monitored packet is analyzed.
  - By default, ingress traffic on a destination port is dropped.

Restrictions

When you configure remote port mirroring, the following restrictions apply:

- You cannot configure the same source port to be used in multiple source sessions.
- You cannot configure a source port channel or source VLAN in a source session if the port channel or VLAN has a member port that is configured as a destination port in a remote-port mirroring session.
- A destination port for remote port mirroring cannot be used as a source port, including the session in which the port functions as the destination port.
- A destination port cannot be used in any spanning tree instance.
- The reserved VLAN used to transport mirrored traffic must be a L2 VLAN. L3 VLANs are not supported.

Configuration Procedure

To configure remote port mirroring, you must configure:

1. A reserved L2 VLAN used to transport (switched) mirrored packets on source, intermediate, and destination switches
2. A source session that consists of multiple source ports, port channels, and VLANs which are associated with the dedicated VLAN and located on different source switches
3. A destination session that consists of multiple destination ports associated with the dedicated VLAN and located on different destination switches

Configure a dedicated L2 VLAN for Remote Port Mirroring

Step	Command Syntax	Command Mode	Task
1	**interface vlan** *vlan-id*	CONFIGURATION	Create a VLAN to transport mirrored traffic in remote port mirroring. Valid *vlan-id* values are 1 to 4094. The default VLAN ID is not supported.
2	**mode remote-port-mirroring**	VLAN INTERFACE	Configure the dedicated L2 VLAN to be used to transport mirrored traffic in remote port mirroring.
3	**tagged** *interface*	VLAN INTERFACE	Configure a tagged port to carry mirrored traffic in the reserved VLAN. Repeat this command to configure additional tagged ports for the VLAN.
4	Repeat Steps 1 to 3 on source, intermediate, and destination switches on which mirrored traffic in the reserved L2 VLAN is transmitted (see Figure 38-11 and Figure 38-12).		

To remove the remote-port mirroring assignment from a VLAN, enter the **no mode remote-port-mirroring** command.

Configure a Source Session on Multiple Switches

Step	Command Syntax	Command Mode	Task
1	**monitor session** *session-id*	CONFIGURATION	Configure a new remote-port mirroring session or add or delete source ports from an existing session, and enter Monitor Session configuration mode. Up to 4 source sessions are supported on a switch. Refer to Configuration Notes for information on datapath limitations. *session-id*: Session number used to identify the mirroring session. Range: 0 - 65535.
2	**source {***single-interface* | **range {***interface-list* | *interface-range* | *mixed-interface-list***}** | **vlan** *vlan-id* | **range {***vlan-list* | *vlan-range* | *mixed-vlan-list***}} destination remote-vlan** *vlan-id* **direction {rx** | **tx** | **both}**	MONITOR SESSION	Configure the source ports, ingress/egress traffic to be mirrored, and the reserved VLAN used to transport mirrored traffic. *single-interface* specifies one of the following interface types: 1-Gigabit Ethernet: Enter **gigabitethernet** *slot*/*port*. 10-Gigabit Ethernet: Enter **tengigabitethernet** *slot*/*port*. Port channel: Enter **port-channel** {1-511}. **range** *interface-list* specifies multiple interfaces separated by a comma and space: *single-interface*, *single-interface*, *single-interface*...

Configure a Source Session on Multiple Switches

Step	Command Syntax	Command Mode	Task
2	**source {***single-interface* | **range {***interface-list* | *interface-range* | *mixed-interface-list***}** | **vlan** *vlan-id* | **range {***vlan-list* | *vlan-range* | *mixed-vlan-list***}} destination remote-vlan** *vlan-id* **direction {rx** | **tx** | **both}**	MONITOR SESSION	**range** *interface-range* specifies one of the following interface ranges: **gigabitethernet** *slot/first_port - last_port* **tengigabitethernet** *slot/first_port - last_port* **port-channel** *first_number - last_number* A space is required before and after the dash (-). **range** *mixed-interface-list* specifies single interfaces and interface ranges in any order: *single-interface*, *interface-range*, *single-interface*... **vlan** *vlan-id* specifies a single VLAN ID. Range: 1-4094. **range** *vlan-list* specifies multiple VLAN IDs separated by a comma and space: **vlan** *vlan-id*, **vlan** *vlan-id*, **vlan** *vlan-id*... **range** *vlan-range* specifies a range of VLANs in the format: **vlan** *first_vlanID - last_vlanID.* A space is required before and after the dash (-). **range** *mixed-vlan-list* specifies single VLAN IDs and VLAN ranges in any order: **vlan** *vlan-id*, **vlan** *first_vlanID - last_vlanID*, **vlan** *vlan-id*... **destination remote-vlan** *vlan-id* associates the reserved VLAN with the source ports used in this source session. **direction** specifies the incoming/outgoing traffic on the source port/ to be mirrored: ingress (**rx**), egress (**tx**), or ingress and egress (**both**).
3	**no disable**	MONITOR SESSION	Activate a remote-port mirroring session.
4	**flow-based enable**	MONITOR SESSION	(Optional) Enable flow-based mirroring for this source session to monitor only specified traffic. See Flow-based Monitoring on page 820 for more information.
5	Repeat Steps 1 to 4 on other source switches to configure additional source ports for this session.		

To delete one or more source ports or source VLANs from a mirroring session, enter the **no source destination remote-vlan** *vlan-id* command, specifying the ports to be deleted in the command syntax.

To change the reserved L2 VLAN used in a source session, you can delete the session (**no monitor session** command) or remove the current VLAN by entering the complete **no source** command. Then re-enter the complete **source** command as described above to configure a new reserved VLAN for the session.

Configure a Destination Session on Multiple Switches

Step	Command Syntax	Command Mode	Task
1	**monitor session** *session-id*	CONFIGURATION	Configure the destination session for remote port mirroring and enter Monitor Session configuration mode.
2	**source remote-vlan** *vlan-id* **destination {***single-interface* | **range {***interface-list* | *interface-range* | *mixed-interface-list***}}**	MONITOR SESSION	Associate the reserved L2 VLAN used to transport mirrored traffic with this destination session and configure the destination ports to which an analyzer is a connected. *single-interface* is one of the following values: **gigabitethernet** *slot*/*port* **tengigabitethernet** *slot*/*port* **range** *interface-list* specifies multiple interfaces separated by a comma and space: *single-interface*, *single-interface*, *single-interface*... **range** *interface-range* specifies one of the following interface ranges: **gigabitethernet** *slot*/*first_port-last_port* **tengigabitethernet** *slot*/*first_port-last_port* A space is required before and after the dash (-). **range** *mixed-interface-list* specifies single interfaces and interface ranges in any order: *single-interface*, *interface-range*, *single-interface*...
3	**tagged destination {***single-interface* | **range** *interface-range***}**	MONITOR SESSION	(Optional) Configure destination ports so that the reserved VLAN tag is added to the monitored traffic sent to an analyzer. Default: Destination ports send untagged packets so that the reserved VLAN ID is removed and the original monitored packet is analyzed. *single-interface* is one of the following values: **gigabitethernet** *slot*/*port* **tengigabitethernet** *slot*/*port* **range** *interface-range* specifies one of the following interface ranges: **gigabitethernet** *slot*/*first_port-last_port* **tengigabitethernet** *slot*/*first_port-last_port*
4	Repeat Steps 1 to 3 on other destination switches to configure additional destination ports for this session.		

To delete one or more destination ports from a destination session, enter the **no source remote-vlan** *vlan-id* **destination** command.

To change the reserved L2 VLAN used in the destination session, you must first remove all destination ports. Then delete the current VLAN by entering the **no monitor session** *session-id* **source remote-vlan** *vlan-id* command and re-enter the **monitor session** *session-id* **source remote-vlan** command to configure the new VLAN ID.

To reconfigure destination ports as untagged ports, enter the **untagged destination** command.

Displaying Remote-Port Mirroring Configurations

To display the current configuration of remote port mirroring for a specified session, enter the **show config** command in MONITOR SESSION configuration mode.

```
FTOS(conf-mon-sess-50)# show config
!
monitor session 50
source GigabitEthernet 1/2 destination remote-vlan 50 direction both mode Remote-Port-Mirroring
source vlan 4, vlan 11 - 12, destination remote-vlan 50 direction both mode Remote-Port-Mirroring
no disable
```

To display the currently configured source and destination sessions for remote port mirroring on a switch, enter the **show monitor session** command in EXEC Privilege mode.

```
FTOS# show monitor session

SessionID     Source          Destination      Direction      Mode                    Type
---------     ------          -----------      ---------      ----                    ----
       50     Gi 1/2          remote-vlan 50   both           Remote-Port-Mirroring   Port-based
       50     Po 1            remote-vlan 50   both           Remote-Port-Mirroring   Port-based
```

Session 50: Source session monitors a source port and source port-channel using VLAN 50

```
       20    Gi 1/3           remote-vlan 60   both           Remote-Port-Mirroring   Flow-based
```

Session 20: Source session monitors a source port configured for flow-based monitoring using VLAN 50 at a destination port-channel

```
       10   remote-vlan 101   Gi 1/4          N/A           Remote-Port-Mirroring    N/A
```

Session 10: Destination session receives mirrored traffic on VLAN 101 at a destination port

To display the current configuration of the reserved VLAN, enter the **show vlan** command.

```
FTOS(conf-if-gi-1/2)# show vlan

Codes: * - Default VLAN, G - GVRP VLANs, R - Remote Port Mirroring VLANs,
P - Primary, C - Community, I - Isolated
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged, M - Vlan-stack

NUM            Status        Q Ports
*       1      Inactive
R,G     2      Active        T Gi 13/7
```

Sample Configuration: Remote Port Mirroring

Remote port mirroring requires a source session (monitored ports on different source switches), a reserved tagged VLAN for transporting mirrored traffic (configured on source, intermediate, and destination switches), and a destination session (destination ports connected to analyzers on destination switches).

Figure 38-10 shows a sample configuration of remote port mirroring on a source switch. Note that in the **show monitor session** output of a source session, the source is a source port/port-channel (for example, Gi 2/2) and the destination is the reserved VLAN (for example, remote-vlan 22).

Figure 38-10. Configuring Remote Port Mirroring: Source Switch

```
FTOS(conf)#interface gigabitethernet 2/22
FTOS(conf-if-gi-2/22)#switchport
FTOS(conf-if-gi-2/22)#no shutdown
FTOS(conf-if-gi-2/22)#interface vlan 22
FTOS(conf-if-vl-22)#mode remote-port-mirroring
FTOS(conf-if-vl-22)#tagged gigabitethernet 2/22
FTOS(conf-if-vl-22)#exit
FTOS(conf)#monitor session 100
FTOS(conf-mon-sess-100)#source gi 2/2 destination remote-vlan 22 direction both
FTOS(conf-mon-sess-100)#no disable
FTOS(conf-mon-sess-100)#show config

monitor session 100
source GigabitEthernet 2/2 destination remote-vlan 22 direction both
no disable
FTOS(conf-mon-sess-100)#end
FTOS#show monitor session 100
   SessionID     Source       Destination     Direction   Mode                    Type
   ---------     ------       -----------     ---------   ----                    ----
     100         Gi 2/2       remote-vlan 22  both        Remote-Port-Mirroring   Port-based
```

Figure 38-11 shows a sample configuration of remote port mirroring on an intermediate (transport) switch.

Figure 38-11. Configuring Remote Port Mirroring: Intermediate Switch

```
FTOS(conf)#interface gigabitethernet 4/74
FTOS(conf-if-gi-4/74)#switchport
FTOS(conf-if-gi-4/74)#no shutdown
FTOS(conf-if-gi-4/74)#exit
FTOS(conf)#interface gigabitethernet 4/18
FTOS(conf-if-gi-4/18)#switchport
FTOS(conf-if-gi-4/18)#no shutdown
FTOS(conf-if-gi-4/18)#exit
FTOS(conf)#interface vlan 22
FTOS(conf-if-vl-22)#mode remote-port-mirroring
FTOS(conf-if-vl-22)#tagged gi 4/74
FTOS(conf-if-vl-22)#tagged gi 4/18
FTOS(conf-if-vl-22)#no shut
FTOS(conf-if-vl-22)#end
```

Figure 38-12 shows a sample configuration of remote port mirroring on a destination switch. Note that in the **show monitor session** output of a destination session, the source is the reserved VLAN (for example, remote-vlan 22) and the destination is the destination port (for example, Gi 4/73) to which an analyzer is attached.

Figure 38-12. Configuring Remote Port Mirroring: Destination Switch

```
FTOS(conf)#interface vlan 22
FTOS(conf-if-vl-22)#mode remote-port-mirroring
FTOS(conf-if-vl-22)#tagged gi 4/48
FTOS(conf-if-vl-22)#no shutdown
FTOS(conf-if-vl-22)#exit
FTOS(conf)#monitor session 100
FTOS(conf-mon-sess-100)#source remote-vlan 22 destination gi 4/73
FTOS(conf-mon-sess-100)#tagged destination gi 4/73
FTOS(conf-mon-sess-100)#show config

monitor session 100
source remote-vlan 22 destination GigabitEthernet 4/73
FTOS(conf-mon-sess-100)#end

FTOS#show monitor session 100
   SessionID     Source             Destination    Direction    Mode                   Type
   ---------     ------             -----------    ---------    ----                   ----
      100        remote-vlan 22     Gi 4/73        N/A          Remote-Port-Mirroring  N/A
```

39

Private VLANs

Private VLANs is available on platforms: C S

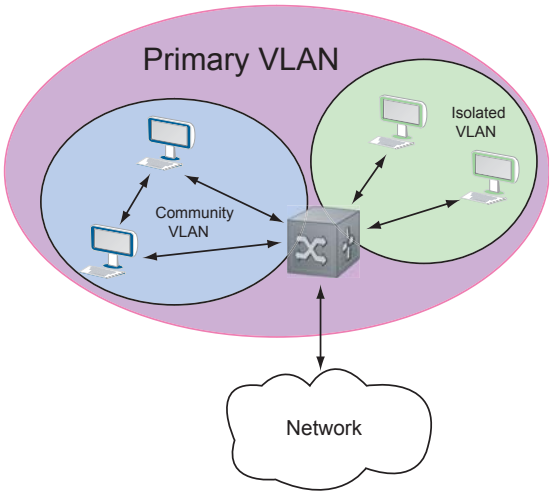
Private VLANs (PVLANs) provide Layer 2 isolation between ports within the same VLAN. That is, peer-to-peer communication is restricted or blocked. This is done by dividing the VLAN, into subdomains, and then restricting or blocking traffic flow between them.

Note: While conceptually, the primary VLAN is divided into secondary VLANs, when configuring PVLAN in FTOS, you explicitly define the secondary VLANs, and then make them members of the primary VLAN.

The VLAN that is divided into subdomains is called the *Primary VLAN*; the subdomains are called *secondary VLANs*. There are two types of secondary VLANs:

- **Community VLAN** — a group of ports in which ports may communicate with each other and promiscuous ports, but not to ports outside of their own secondary VLAN. A service provider can provide Layer 2 security for customers and use the IP addresses more efficiently, by using a separate community VLAN per customer, while at the same time using the same IP subnet address space for all community and isolated VLANs mapped to the same primary VLAN.
- **Isolated VLAN** — a group of ports in which ports may communicate with promiscuous ports only; they may not communicate with each other, or to other ports outside of their own secondary VLAN. An enterprise, such as a hotel, can use an isolated VLAN in a private VLAN to provide Internet access for its guests, while stopping direct access between the guest ports.

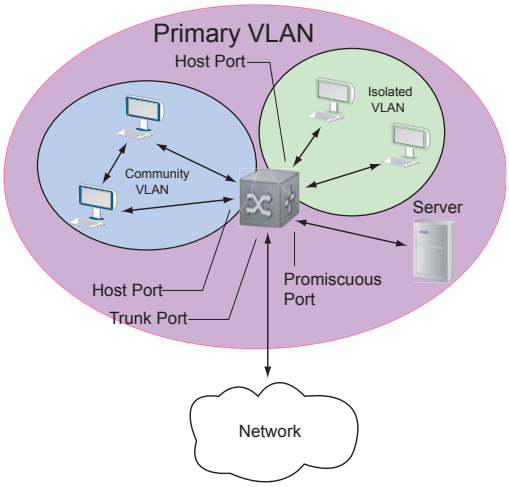
Figure 39-1. PVLAN: Primary and Secondary VLANs



There are three types of ports in PVLAN:

- **Host Ports**—these ports are the ones that Private VLAN aims to isolate. They are connected to end-stations.
- **Promiscuous Ports**—these ports are members of the primary VLAN, and function as gateways to the primary and secondary VLANs.
- **Trunk Ports**—trunk ports carry tagged traffic between switches. They have promiscuous and trunk ports as members.

Figure 39-2. PVLAN: Primary and Secondary VLANs



Important Points to Remember

- Even if secondary VLANs are operationally down, if the primary VLAN is operationally up, Layer 3 traffic is still be transmitted across the secondary VLANs.
- PVLAN ports cannot be added to regular VLANs. Conversely, regular VLAN ports cannot be added to PVLANs.
- If a promiscuous or host port is untagged in a VLAN, and it receives a tagged packet in the same VLAN, the packet will NOT be dropped.
- A primary VLAN and each of its secondary VLANs decrement the available number of VLAN IDs in the switch.

Configure Private VLANs

Configuring Private VLANs is a 3-step process:

1. Configure PVLAN Ports
2. Place PVLAN Ports in a Secondary VLAN
3. Place the Secondary VLANs in a Primary VLAN

Related Configuration Tasks

- Private VLAN show Commands on page 834

Configure PVLAN Ports

You must assign switchports a PVLAN Port role—host, promiscuous, or trunk—before you can add them to a primary or secondary VLAN.

- **Host ports** may not be a part of a non-private (regular) VLAN.
- **Promiscuous ports** may be a member of more than one primary VLAN, but may not be a member of a regular VLAN.
- **Trunk ports** may be a member of a regular VLAN.

Task	Command Syntax	Command Mode
Assign a PVLAN port role to a switchport.	**switchport mode private-vlan {host | promiscuous | trunk}**	INTERFACE

Place PVLAN Ports in a Secondary VLAN

PVLAN has two types of secondary VLANs:

Community VLANs:

- Can only have host ports.
- Host ports can communicate with each other and to promiscuous ports.

Isolated VLANs:

- Can only have host ports.
- Host ports cannot communicate with each other; they can only communicate with promiscuous ports.

Step	Task	Command Syntax	Command Mode
1	Access the INTERFACE VLAN mode for the VLAN that you want to make a community VLAN.	**interface vlan** *vlan-id*	CONFIGURATION
2	Designate the VLAN as a community or isolated VLAN.	**private-vlan mode {community | isolated}**	INTERFACE VLAN
3	Add one or more host ports to the VLAN.	{**tagged** | **untagged**} *interface*	INTERFACE VLAN

Place the Secondary VLANs in a Primary VLAN

A *primary VLAN* is a port-based VLAN that is specifically designated as a private VLAN. Doing so enables the VLAN to be divided into secondary VLANs.

Step	Task	Command Syntax	Command Mode
1	Access INTERFACE VLAN mode for the VLAN that you want to make the primary VLAN.	**interface vlan** *vlan-id*	CONFIGURATION
2	Designate a VLAN as a primary VLAN.	**private-vlan mode primary**	INTERFACE VLAN
3	Map secondary VLANs to the primary VLAN.	**private-vlan mapping secondary-vlan** *vlan-list*	INTERFACE VLAN
4	Add promiscuous ports as tagged or untagged interfaces. Add trunk ports as tagged.	{**tagged** | **untagged**} *interface*	INTERFACE VLAN
5	Enable Proxy ARPing on the primary VLAN to enable Layer 3 communication between hosts on different secondary VLANs.	ip local-proxy-arp	INTERFACE VLAN

Private VLAN show Commands

Table 39-1. Private VLAN Commands

Task	Command Syntax	Command Mode
Display type and status of PVLAN interfaces.	**show interfaces private-vlan** [**interface** *interface*]	EXEC Privilege
Display PVLANs and/or interfaces that are part of a PVLAN.	**show vlan private-vlan** [**community** | *interface* | **isolated** | **primary |** *primary_vlan* | **interface** *interface*]	EXEC Privilege
Display primary-secondary VLAN mapping.	**show vlan private-vlan mapping**	EXEC Privilege

40

Per-VLAN Spanning Tree Plus

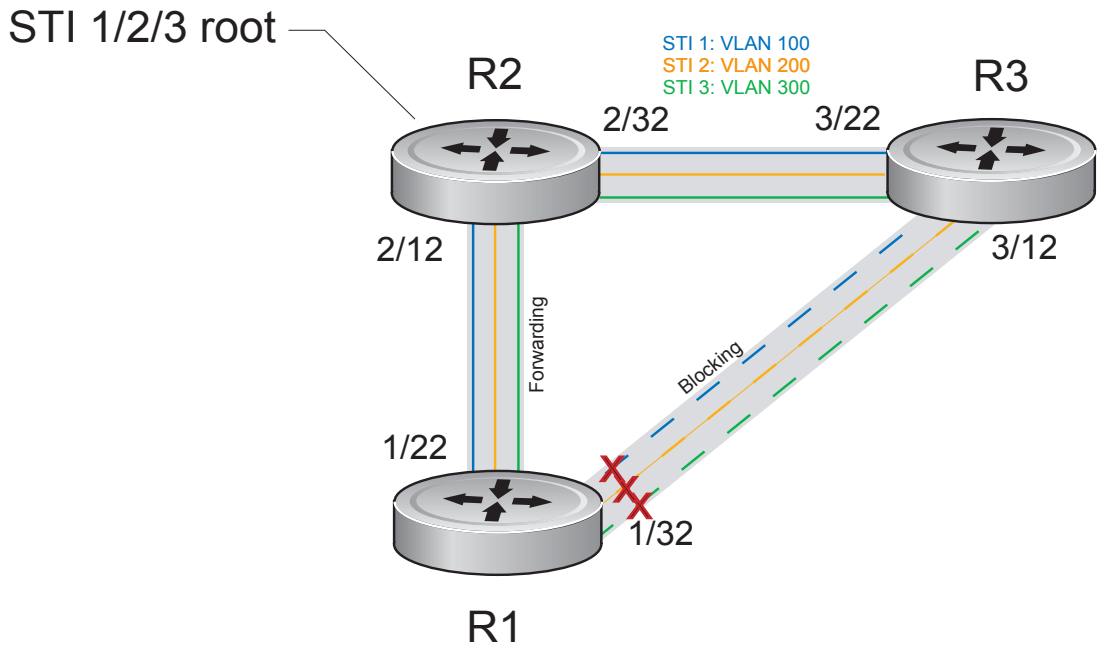
Per-VLAN Spanning Tree Plus is supported platforms: C E S

Port Monitoring is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

Protocol Overview

Per-VLAN Spanning Tree Plus (PVST+) is a variation of Spanning Tree—developed by a third party—that allows you to configure a separate Spanning Tree instance for each VLAN. For more information on Spanning Tree, see Chapter 52, Spanning Tree Protocol.

Figure 40-1. Per-VLAN Spanning Tree



FTOS supports three other variations of Spanning Tree, as shown in Table 40-1.

Table 40-1. FTOS Supported Spanning Tree Protocols

Dell Force10 Term	IEEE Specification
Spanning Tree Protocol	802.1d
Rapid Spanning Tree Protocol	802.1w
Multiple Spanning Tree Protocol	802.1s
Per-VLAN Spanning Tree Plus	Third Party

Implementation Information

- The FTOS implementation of PVST+ is based on IEEE Standard 802.1d.
- The FTOS implementation of PVST+ uses IEEE 802.1s costs as the default costs (Table 40-2). Other implementations use IEEE 802.1d costs as the default costs if you are using Dell Force10 systems in a multi-vendor network, verify that the costs are values you intended.
- You must allocate at least the default minimum amount of Layer 2 ACL CAM space when employing PVST+ on the E-Series. See Configure Ingress Layer 2 ACL Sub-partitions on page 295.
- On the C-Series and S-Series, you can enable PVST+ on 254 VLANs.

Configure Per-VLAN Spanning Tree Plus

Configuring PVST+ is a four-step process:

1. Configure interfaces for Layer 2.
2. Place the interfaces in VLANs.
3. Enable PVST+. See page 837.
4. Optionally, for load balancing, select a non-default bridge-priority for a VLAN. See page 837.

Related Configuration Tasks

- Modify Global PVST+ Parameters on page 840
- Modify Interface PVST+ Parameters on page 840
- Configure an EdgePort on page 841
- Flush MAC Addresses after a Topology Change on page 654
- Preventing Network Disruptions with BPDU Guard on page 1057
- Configuring Spanning Trees as Hitless on page 1064
- PVST+ in Multi-vendor Networks on page 845
- PVST+ Extended System ID on page 845
- PVST+ Sample Configurations on page 847

Enable PVST+

When you enable PVST+, FTOS instantiates STP on each active VLAN. To enable PVST+ globally:

Step	Task	Command Syntax	Command Mode
1	Enter PVST context.	**protocol spanning-tree pvst**	PROTOCOL PVST
2	Enable PVST+.	**no disable**	PROTOCOL PVST

Disable PVST+

Task	Command Syntax	Command Mode
Disable PVST+ globally.	**disable**	PROTOCOL PVST
Disable PVST+ on an interface, or remove a PVST+ parameter configuration.	**no spanning-tree pvst**	INTERFACE

Display your PVST+ configuration by entering the command **show config** from PROTOCOL PVST context, as shown in fig.

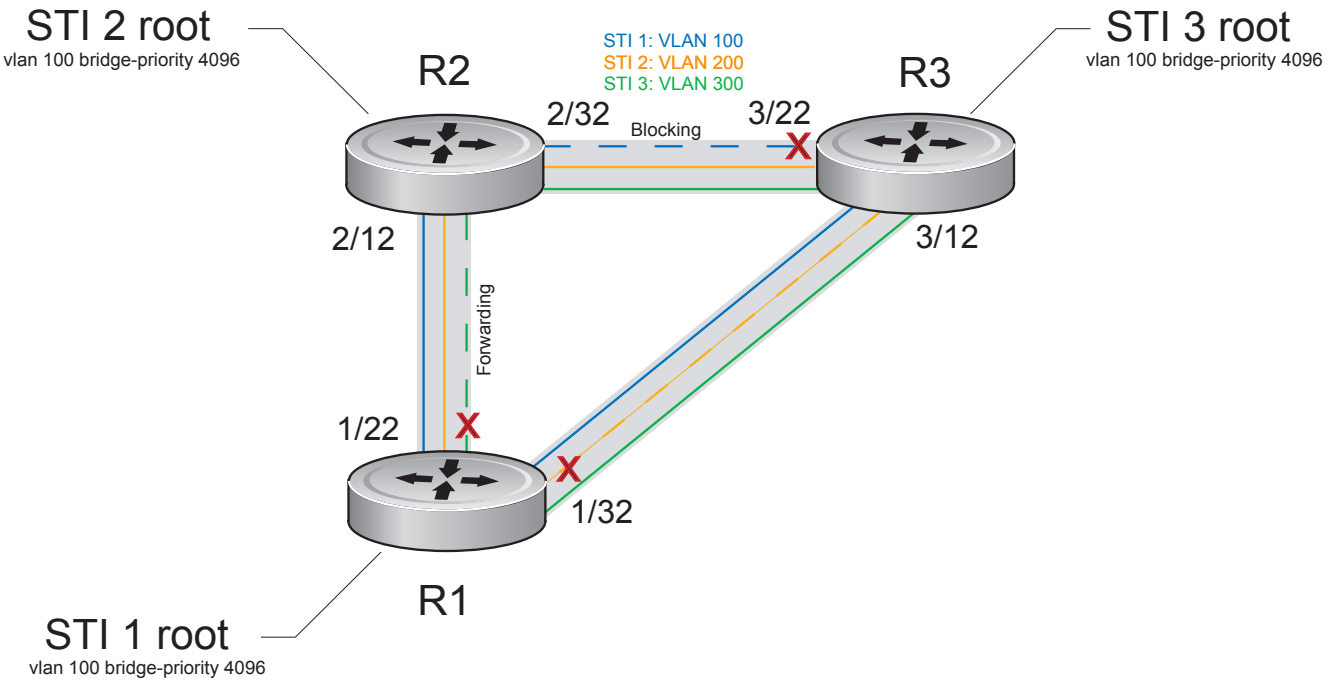
Figure 40-2. Display the PVST+ Configuration

```
Force10_E600(conf-pvst)#show config verbose
!
protocol spanning-tree pvst
 no disable
 vlan 100 bridge-priority 4096
```

Influence PVST+ Root Selection

In Figure 40-1, all VLANs use the same forwarding topology because R2 is elected the root, and all GigabitEthernet ports have the same cost. Figure 40-3 changes the bridge priority of each bridge so that a different forwarding topology is generated for each VLAN. This behavior demonstrates how you can use PVST+ to achieve load balancing.

Figure 40-3. Load Balancing with PVST+



The bridge with the bridge value for bridge priority is elected root. Since all bridges use the default priority (until configured otherwise), lowest MAC address is used as a tie-breaker. Assign bridges a low non-default value for bridge priority to increase the likelihood that it will be selected as the STP root.

Task	Command Syntax	Command Mode
Assign a bridge priority. Range: 0 to 61440 Default: 32768	**vlan bridge-priority**	PROTOCOL PVST

Display the PVST+ forwarding topology by entering the command **show spanning-tree pvst** [**vlan** *vlan-id*] from EXEC Privilege mode, as shown in Figure 40-4.

Figure 40-4. Display the PVST+ Forwarding Topology

```
Force10_E600(conf)#do show spanning-tree pvst vlan 100
VLAN 100
Root Identifier has priority 4096, Address 0001.e80d.b6d6
Root Bridge hello time 2, max age 20, forward delay 15
Bridge Identifier has priority 4096, Address 0001.e80d.b6d6
Configured hello time 2, max age 20, forward delay 15
We are the root of VLAN 100
Current root has priority 4096, Address 0001.e80d.b6d6
Number of topology changes 5, last change occurred 00:34:37 ago on Gi 1/32

Port 375 (GigabitEthernet 1/22) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.375
Designated root has priority 4096, address 0001.e80d.b6:d6
Designated bridge has priority 4096, address 0001.e80d.b6:d6
Designated port id is 128.375 , designated path cost 0
Number of transitions to forwarding state 2
BPDU sent 1159, received 632
The port is not in the Edge port mode

Port 385 (GigabitEthernet 1/32) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.385
Designated root has priority 4096, address 0001.e80d.b6:d6
Designated bridge has priority 4096, address 0001.e80d.b6:d6
Designated port id is 128.385 , designated path cost 0
```

Modify Global PVST+ Parameters

The root bridge sets the values for forward-delay, and hello-time and overwrites the values set on other PVST+ bridges.

- **Forward-delay** is the amount of time an interface waits in the Listening State and the Learning State before it transitions to the Forwarding State.
- **Hello-time** is the time interval in which the bridge sends Bridge Protocol Data Units (BPDUs).
- **Max-age** is the length of time the bridge maintains configuration information before it refreshes that information by recomputing the PVST+ topology.

To change PVST+ parameters, use the following commands on the root bridge:

Task	Command Syntax	Command Mode
Change the forward-delay parameter. • Range: 4 to 30 • Default: 15 seconds	**vlan forward-delay**	PROTOCOL PVST
Change the hello-time parameter. **Note:** With large configurations (especially those with more ports) Dell Force10 recommends that you increase the hello-time. Range: 1 to 10 Default: 2 seconds	**vlan hello-time**	PROTOCOL PVST
Change the max-age parameter. Range: 6 to 40 Default: 20 seconds	**vlan max-age**	PROTOCOL PVST

The values for global PVST+ parameters are given in the output of the command **show spanning-tree pvst**, as shown in Figure 40-4.

Modify Interface PVST+ Parameters

You can adjust two interface parameters to increase or decrease the probability that a port becomes a forwarding port:

- **Port cost** is a value that is based on the interface type. The greater the port cost, the less likely the port will be selected to be a forwarding port.
- **Port priority** influences the likelihood that a port will be selected to be a forwarding port in case that several ports have the same port cost.

Table 40-2 lists the default values for port cost by interface.

Table 40-2. PVST+ Default Port Cost Values

Port Cost	Default Value
100-Mb/s Ethernet interfaces	200000
1-Gigabit Ethernet interfaces	20000
10-Gigabit Ethernet interfaces	2000
Port Channel with 100 Mb/s Ethernet interfaces	180000
Port Channel with 1-Gigabit Ethernet interfaces	18000
Port Channel with 10-Gigabit Ethernet interfaces	1800

Note: The FTOS implementation of PVST+ uses IEEE 802.1s costs as the default costs. Other implementations use IEEE 802.1d costs as the default costs if you are using Dell Force10 systems in a multi-vendor network, verify that the costs are values you intended.

To change the port cost or priority of an interface:

Task	Command Syntax	Command Mode
Change the port cost of an interface. Range: 0 to 200000 Default: see Table 40-2.	**spanning-tree pvst vlan cost**	INTERFACE
Change the port priority of an interface. Range: 0 to 240, in increments of 16 Default: 128	**spanning-tree pvst vlan priority**	INTERFACE

The values for interface PVST+ parameters are given in the output of the command **show spanning-tree pvst**, as shown in Figure 40-4.

Configure an EdgePort

The EdgePort feature enables interfaces to begin forwarding traffic approximately 30 seconds sooner. In this mode an interface forwards frames by default until it receives a BPDU that indicates that it should behave otherwise; it does not go through the Learning and Listening states. The **bpduguard shutdown-on-violation** option causes the interface hardware to be shutdown when it receives a BPDU. When only **bpduguard** is implemented, although the interface is placed in an Error Disabled state when receiving the BPDU, the physical interface remains up and spanning-tree will drop packets in the hardware after a BPDU violation. BPDUs are dropped in the software after receiving the BPDU violation. This feature is the same as PortFast mode in Spanning Tree.

Caution: Configure EdgePort only on links connecting to an end station. EdgePort can cause loops if it is enabled on an interface connected to a network.

To enable EdgePort on an interface, use the following command:

Task	Command Syntax	Command Mode
Enable EdgePort on an interface.	**spanning-tree pvst edge-port [bpduguard | shutdown-on-violation]**	INTERFACE

The EdgePort status of each interface is given in the output of the command **show spanning-tree pvst**, as shown in Figure 40-4.

FTOS Behavior: Regarding **bpduguard shutdown-on-violation** behavior:

1. If the interface to be shutdown is a port channel then all the member ports are disabled in the hardware.
2. When a physical port is added to a port channel already in error disable state, the new member port will also be disabled in the hardware.
3. When a physical port is removed from a port channel in error disable state, the error disabled state is cleared on this physical port (the physical port will be enabled in the hardware).
4. The **reset linecard** command does not clear the error disabled state of the port or the hardware disabled state. The interface continues to be disables in the hardware.

   The error disabled state can be cleared with any of the following methods:

   - Perform an **shutdown** command on the interface.
   - Disable the **shutdown-on-violation** command on the interface (**no spanning-tree** *stp-id* **portfast** [**bpduguard |** [**shutdown-on-violation**]]).
   - Disable spanning tree on the interface (**no spanning-tree** in INTERFACE mode).
   - Disabling global spanning tree (**no spanning-tree** in CONFIGURATION mode).

Configure a Root Guard

Use the Root Guard feature in a Layer 2 PVST+ network to avoid bridging loops.

You enable root guard on a per-port or per-port-channel basis.

FTOS Behavior: The following conditions apply to a port enabled with root guard:

- Root guard is supported on any PVST-enabled port or port-channel interface except when used as a stacking port.
- Root guard is supported on a port in any Spanning Tree mode:
  - Spanning Tree Protocol (STP)
  - Rapid Spanning Tree Protocol (RSTP)
  - Multiple Spanning Tree Protocol (MSTP)
  - Per-VLAN Spanning Tree Plus (PVST+)
- When enabled on a port, root guard applies to all VLANs configured on the port.
- Root guard and loop guard cannot be enabled at the same time on a PVST+ port. For example, if you configure loop guard on a port on which root guard is already configured, the following error message is displayed:
  `% Error: RootGuard is configured. Cannot configure LoopGuard.`

To enable a root guard on a PVST-enabled port or port-channel interface, enter the **spanning-tree pvst rootguard** command. Refer to STP Root Guard on page 1060 for more information on how to use the root guard feature.

Task	Command Syntax	Command Mode
Enable root guard on a port or port-channel interface.	**spanning-tree pvst rootguard**	INTERFACE INTERFACE PORT-CHANNEL

To disable PVST+ root guard on a port or port-channel interface, enter the **no spanning-tree pvst rootguard** command in an interface configuration mode.

To verify the PVST+ root guard configuration on a port or port-channel interface, enter the **show spanning-tree pvst** [**vlan** *vlan-id*] **guard** command in global configuration mode.

Configure a Loop Guard

The Loop Guard feature provides protection against Layer 2 forwarding loops (STP loops) caused by a hardware failure, such as a cable failure or an interface fault. When a cable or interface fails, a participating STP link may become unidirectional (STP requires links to be bidirectional) and an STP port does not receive BPDUs. When an STP blocking port does not receive BPDUs, it transitions to a forwarding state. This condition can create a loop in the network.

You enable loop guard on a per-port or per-port channel basis.

FTOS Behavior: The following conditions apply to a port enabled with loop guard:

- Loop guard is supported on any PVST-enabled port or port-channel interface.
- Loop guard is supported on a port or port-channel in any Spanning Tree mode:
  - Spanning Tree Protocol (STP)
  - Rapid Spanning Tree Protocol (RSTP)
  - Multiple Spanning Tree Protocol (MSTP)
  - Per-VLAN Spanning Tree Plus (PVST+)
- Root guard and loop guard cannot be enabled at the same time on a PVST+ port. For example, if you configure root guard on a port on which loop guard is already configured, the following error message is displayed:
  `% Error: LoopGuard is configured. Cannot configure RootGuard.`
- Enabling Portfast BPDU guard and loop guard at the same time on a port results in a port that remains in a blocking state and prevents traffic from flowing through it. For example, when Portfast BPDU guard and loop guard are both configured:

  - If a BPDU is received from a remote device, BPDU guard places the port in an err-disabled blocking state and no traffic is forwarded on the port.

  - If no BPDU is received from a remote device, loop guard places the port in a loop-inconsistent blocking state and no traffic is forwarded on the port.
- When used in a PVST+ network, loop guard is performed per-port or per-port channel at a VLAN level. If no BPDUs are received on a VLAN interface, the port or port-channel transitions to a loop-inconsistent (blocking) state only for this VLAN.

To enable a loop guard on a PVST-enabled port or port-channel interface, enter the **spanning-tree mstp loopguard** command. Refer to STP Loop Guard on page 1064 for more information on how to use the loop guard feature.

Task	Command Syntax	Command Mode
Enable loop guard on an PVST-enabled port or port-channel interface.	**spanning-tree pvst loopguard**	INTERFACE INTERFACE PORT-CHANNEL

To disable PVST+ loop guard on a port or port-channel interface, enter the **no spanning-tree pvst loopguard** command in an INTERFACE configuration mode.

To verify the PVST+ loop guard configuration on a port or port-channel interface, enter the **show spanning-tree pvst** [**vlan** *vlan-id*] command in global configuration mode.

PVST+ in Multi-vendor Networks

Some non-Dell Force10 systems which have hybrid ports participating in PVST+ transmit two kinds of BPDUs: an 802.1D BPDU and an untagged PVST+ BPDU.

Dell Force10 systems do not expect PVST+ BPDU (tagged or untagged) on an untagged port. If this happens, FTOS places the port in error-disable state. This behavior might result in the network not converging. To prevent FTOS from executing this action, use the command **no spanning-tree pvst err-disable cause invalid-pvst-bpdu**. After you configure this command, if the port receives a PVST+ BPDU, the BPDU is dropped, and the port remains operational.

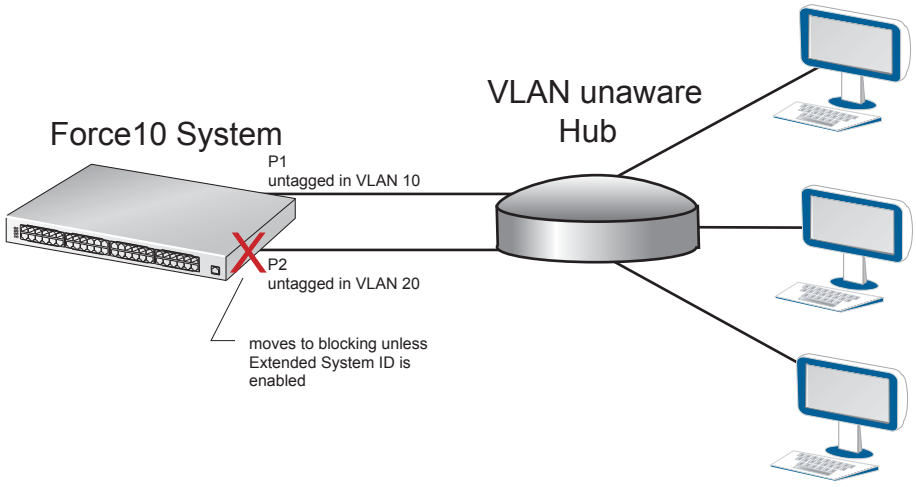
PVST+ Extended System ID

In Figure 40-5, ports P1 and P2 are untagged members of different VLANs. These ports are untagged because the hub is VLAN unaware. There is no data loop in the above scenario, however, PVST+ can be employed to avoid potential mis-configurations.

If PVST+ is enabled on the Dell Force10 switch in this network, P1 and P2 receive BPDUs from each other. Ordinarily, the Bridge ID in the frame matches the Root ID, a loop is detected, and the rules of convergence require that P2 move to blocking state because it has the lowest port ID.

To keep both ports in forwarding state, use Extend System ID. Extend System ID augments the Bridge ID with a VLAN ID to differentiate BPDUs on each VLAN so that PVST+ does not detect a loop, and both ports can remain in forwarding state.

Figure 40-5. PVST+ with Extend System ID



Task	Command Syntax	Command Mode
Augment the Bridge ID with the VLAN ID.	**extend system-id**	PROTOCOL PVST

```
FTOS(conf-pvst)#do show spanning-tree pvst vlan 5 brief

VLAN 5
Executing IEEE compatible Spanning Tree Protocol
Root ID    Priority 32773, Address 0001.e832.73f7
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID    Priority 32773  (priority 32768 sys-id-ext 5), Address 0001.e832.73f7
We are the root of Vlan 5
Configured hello time 2, max age 20, forward delay 15
...
```

Displaying STP Guard Configuration

To verify the STP guard configured on PVST interfaces, enter the **show spanning-tree pvst** [**vlan** *vlan-id*] **guard** command. Refer to Chapter 52, “Spanning Tree Protocol,” on page 1049 for information on how to configure and use the STP root guard, loop guard, and BPDU guard features.

Figure 40-6 shows an example for VLAN 5 in a PVST network in which:

- Root guard is enabled on a port that is in a root-inconsistent state.
- Loop guard is enabled on a port that is in a forwarding state.
- BPDU guard is enabled on a port that is shut down.

Figure 40-6. Displaying STP Guard Configuration

```
FTOS#show spanning-tree pvst vlan 5 guard
Interface
Name       Instance    Sts          Guard type
---------  --------    ---------    ----------
Gi 0/1     5           INCON(Root)  Root
Gi 0/2     5           FWD          Loop
Gi 0/3     5           EDS(shut)    Bpdu
```

PVST+ Sample Configurations

Figure 40-7, Figure 40-8, and Figure 40-9 provide the running configurations for the topology shown in Figure 40-3.

Figure 40-7. PVST+ Sample Configuration: R1 Running-configuration

```
interface GigabitEthernet 1/22
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/32
 no ip address
 switchport
 no shutdown
!
protocol spanning-tree pvst
 no disable
 vlan 100 bridge-priority 4096
interface Vlan 100
 no ip address
 tagged GigabitEthernet 1/22,32
 no shutdown
!
interface Vlan 200
 no ip address
 tagged GigabitEthernet 1/22,32
 no shutdown
!
interface Vlan 300
 no ip address
 tagged GigabitEthernet 1/22,32
 no shutdown
!
protocol spanning-tree pvst
 no disable
 vlan 100 bridge-priority 4096
```

Figure 40-8. PVST+ Sample Configuration: R2 Running-configuration

```
interface GigabitEthernet 2/12
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 2/32
 no ip address
 switchport
 no shutdown
!
interface Vlan 100
 no ip address
 tagged GigabitEthernet 2/12,32
 no shutdown
!
interface Vlan 200
 no ip address
 tagged GigabitEthernet 2/12,32
 no shutdown
!
interface Vlan 300
 no ip address
 tagged GigabitEthernet 2/12,32
 no shutdown
!
protocol spanning-tree pvst
 no disable
 vlan 200 bridge-priority 4096
```

Figure 40-9. PVST+ Sample Configuration: R3 Running-configuration

```
interface GigabitEthernet 3/12
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 3/22
 no ip address
 switchport
 no shutdown
!
interface Vlan 100
 no ip address
 tagged GigabitEthernet 3/12,22
 no shutdown
!
interface Vlan 200
 no ip address
 tagged GigabitEthernet 3/12,22
 no shutdown
!
interface Vlan 300
 no ip address
 tagged GigabitEthernet 3/12,22
 no shutdown
!
protocol spanning-tree pvst
 no disable
 vlan 300 bridge-priority 4096
```

41

Quality of Service

Quality of Service (QoS) is supported on platforms: C E S

Differentiated service is accomplished by classifying and queuing traffic, and assigning priorities to those queues.

The E-Series has eight unicast queues per port and 128 multicast queues per-port pipe. Traffic is queued on ingress and egress. By default, on ingress, all data traffic is mapped to Queue 0, and all control traffic is mapped to Queue 7. On egress control traffic is mapped across all eight queues. All queues are serviced using the Weighted Fair Queuing scheduling algorithm. You can only manage queuing prioritization on egress.

The C-Series traffic has eight queues per port. Four queues are for data traffic and four are for control traffic. All queues are serviced using the Deficit Round Robin scheduling algorithm. You can only manage queuing prioritization on egress.

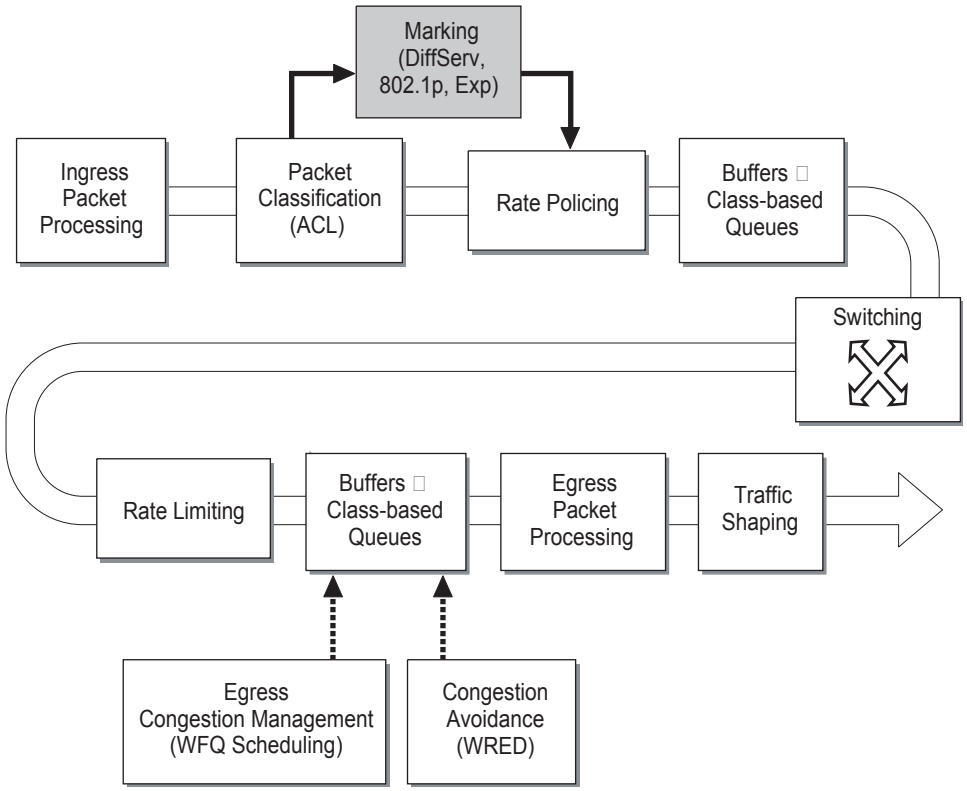
Table 41-1. FTOS Support for Port-based, Policy-based, and Multicast QoS Features

Feature	Platform	Direction
Port-based QoS Configurations	C E S	Ingress + Egress
Set dot1p Priorities for Incoming Traffic	C E S	Ingress
Honor dot1p Priorities on Ingress Traffic	C E S	
Configure Port-based Rate Policing	C E S	
Configure Port-based Rate Limiting	E	Egress
Configure Port-based Rate Shaping	C E S	
Policy-based QoS Configurations	C E S	Ingress + Egress
Classify Traffic	C E S	Ingress
Create a Layer 3 class map	C E S	
Set DSCP values for egress packets based on flow	C E S	
Create a Layer 2 class map	C E S	
Create a QoS Policy	C E S	Ingress + Egress

Table 41-1. FTOS Support for Port-based, Policy-based, and Multicast QoS Features (continued)

Feature	Platform	Direction
Create an input QoS policy	C E S	Ingress
Configure policy-based rate policing	C E S	
Set a DSCP value for egress packets	C E S	
Set a dot1p value for egress packets	C E S	
Create an output QoS policy	C E S	Egress
Configure policy-based rate limiting	E	
Configure policy-based rate shaping	C E S	
Allocate bandwidth to queue	C E S	
Specify WRED drop precedence	E	
Create Policy Maps	C E S	Ingress + Egress
Create Input Policy Maps	C E S	Ingress
Honor DSCP values on ingress packets	C E S	
Honoring dot1p values on ingress packets	E C S	
Create Output Policy Maps	C E S	Egress
Specify an aggregate QoS policy	C E S	
QoS Rate Adjustment	C E S	
Strict-priority Queueing	C E S	—
Weighted Random Early Detection	E	Egress
Create WRED Profiles	E	
Configure WRED for Storm Control	E	
Allocating Bandwidth to Multicast Queues	E	Egress
Pre-calculating Available QoS CAM Space	C E S	—
Viewing QoS CAM Entries	E	—

Figure 41-1. Dell Force10 QoS Architecture



Implementation Information

Dell Force10 QoS implementation complies with IEEE 802.1p *User Priority Bits for QoS Indication*. It also implements these Internet Engineering Task Force (IETF) documents:

- RFC 2474, *Definition of the Differentiated Services Field (DS Field) in the IPv4 Headers*
- RFC 2475, *An Architecture for Differentiated Services*
- RFC 2597, *Assured Forwarding PHB Group*
- RFC 2598, *An Expedited Forwarding PHB*

You cannot configure port-based and policy-based QoS on the same interface, and SONET line cards support only port-based QoS.

Port-based QoS Configurations

You can configure the following QoS features on an interface:

Set dot1p Priorities for Incoming Traffic

Change the priority of incoming traffic on the interface using the command **dot1p-priority** from INTERFACE mode, as shown in Figure 41-2. FTOS places traffic marked with a priority in a queue based on Table 41-2. If you set a dot1p priority for a port-channel, all port-channel members are configured with the same value. You cannot assign a dot1p value to an individual interfaces in a port-channel.

FTOS Behavior: The C-Series and S-Series distribute eight dot1p priorities across four data queues. This is different from the E-Series, which distributes eight dot1p priorities across eight queues (Table 41-2).

Table 41-2. dot1p-priority values and queue numbers

dot1p	E-Series Queue Number	C-Series Queue Number	S-Series Queue Number
0	2	1	1
1	0	0	0
2	1	0	0
3	3	1	1
4	4	2	2
5	5	2	2
6	6	3	3
7	7	3	3

Figure 41-2. Configuring dot1p Priority on an Interface

```
FTOS#config
FTOS(conf)#interface gigabitethernet 1/0
FTOS(conf-if)#switchport
FTOS(conf-if)#dot1p-priority 1
FTOS(conf-if)#end
FTOS#
```

Honor dot1p Priorities on Ingress Traffic

By default FTOS does not honor dot1p priorities on ingress traffic. Use the command **service-class dynamic dot1p** from INTERFACE mode to honor dot1p priorities on ingress traffic, as shown in Figure 41-3. You can configure this feature on physical interfaces and port-channels, but you cannot configure it on individual interfaces in a port channel.

On the C-Series and S-Series you can configure **service-class dynamic dot1p** from CONFIGURATION mode, which applies the configuration to all interfaces. A CONFIGURATION mode **service-class dynamic dot1p** entry supersedes any INTERFACE entries. See Mapping dot1p values to service queues on page 867.

Note: You cannot configure **service-policy input** and **service-class dynamic dot1p** on the same interface.

Figure 41-3. service-class dynamic dot1p Command Example

```
FTOS#config t
FTOS(conf)#interface gigabitethernet 1/0
FTOS(conf-if)#service-class dynamic dot1p
FTOS(conf-if)#end
FTOS#
```

Priority-tagged Frames on the Default VLAN

Priority-tagged Frames on the Default VLAN is available only on platforms: E(X) C S

Priority-tagged frames are 802.1Q tagged frames with VLAN ID 0. For VLAN classification these packets are treated as untagged. However, the dot1p value is still honored when **service-class dynamic dot1p** or **trust dot1p** is configured.

When priority-tagged frames ingress an untagged port or hybrid port the frames are classified to the default VLAN of the port, and to a queue according to their dot1p priority if **service-class dynamic dotp** or **trust dot1p** are configured. When priority-tagged frames ingress a tagged port, the frames are dropped because for a tagged port the default VLAN is 0.

FTOS Behavior: Hybrid ports can receive untagged, tagged, and priority tagged frames. The rate metering calculation might be inaccurate for untagged ports, since an internal assumption is made that all frames are treated as tagged. Internally the ASIC adds a 4-bytes tag to received untagged frames. Though these 4-bytes are not part of the untagged frame received on the wire, they are included in the rate metering calculation resulting in metering inaccuracy.

Configure Port-based Rate Policing

Rate policing ingress traffic on an interface using the command **rate police** from INTERACE mode, as shown in Figure 41-4. If the interface is a member of a VLAN, you may specify the VLAN for which ingress packets are policed.

FTOS Behavior:
On the C-Series and S-Series, rate shaping is effectively rate limiting because of its smaller buffer size.
On the E-Series:
— 802.1Q-priority tagged frames are sometimes not rate-limited according to the configured rate-limit value. Only hybrid ports reliably apply the configured rate limit to priority-tagged frames
— Rate-limiting may not be applied according to the configured rate-limit value on an interface on which the dot.1p priority is changed on incoming traffic using the **dot1p-priority** command

Figure 41-4. Rate Policing Ingress Traffic

```
FTOS#config t
FTOS(conf)#interface gigabitethernet 1/0
FTOS(conf-if)#rate police 100 40 peak 150 50
FTOS(conf-if)#end
FTOS#
```

Figure 41-5. Displaying your Rate Policing Configuration

```
FTOS#show interfaces gigabitEthernet 1/2 rate police
  Rate police 300 (50) peak 800 (50)
    Traffic Monitor 0: normal 300 (50) peak 800 (50)
      Out of profile yellow 23386960 red 320605113
    Traffic Monitor 1: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 2: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 3: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 4: normal NA peak NA
      Out of profile yellow 0 red 0
```

Configure Port-based Rate Limiting

Configure Port-based Rate Limiting is supported only on platform E

FTOS Behavior:
On the C-Series and S-Series, rate shaping is effectively rate limiting because of its smaller buffer size.
On the E-Series:
— 802.1Q-priority tagged frames are sometimes not rate-limited according to the configured rate-limit value. Only hybrid ports reliably apply the configured rate limit to priority-tagged frames
— Rate-limiting may not be applied according to the configured rate-limit value on an interface on which the dot.1p priority is changed on incoming traffic using the **dot1p-priority** command

Rate limit egress traffic on an interface using the command rate limit from INTERFACE mode, as shown in Figure 41-6. If the interface is a member of a VLAN, you may specify the VLAN for which egress packets are rate limited.

Figure 41-6. Rate Limiting Egress Traffic

```
FTOS#config t
FTOS(conf)#interface gigabitethernet 1/0
FTOS(conf-if)#rate limit 100 40 peak 150 50
FTOS(conf-if)#end
FTOS#
```

Display how your rate limiting configuration affects traffic using the keyword **rate limit** with the command **show interfaces**, as shown in Figure 41-7.

Figure 41-7. Displaying How Your Rate Limiting Configuration Affects Traffic

```
FTOS#show interfaces gigabitEthernet 1/1 rate limit
  Rate limit 300 (50) peak 800 (50)
    Traffic Monitor 0: normal 300 (50) peak 800 (50)
      Out of profile yellow 23386960 red 320605113
    Traffic Monitor 1: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 2: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 3: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 4: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 5: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 6: normal NA peak NA
      Out of profile yellow 0 red 0
    Traffic Monitor 7: normal NA peak NA
      Out of profile yellow 0 red 0
    Total: yellow 23386960 red 320605113
```

Configure Port-based Rate Shaping

Rate shaping buffers, rather than drops, traffic exceeding the specified rate until the buffer is exhausted. If any stream exceeds the configured bandwidth on a continuous basis, it can consume all of the buffer space that is allocated to the port.

Apply rate shaping to outgoing traffic on a port using the command **rate shape** from INTERFACE mode, as shown in Figure 41-8.

FTOS Behavior: On ExaScale, when rate shaping is configured on an interface, the “Dropped Packets” counter in the outputs of **show queue statistics egress** and **show qos statistics wred-profile** does not increment. This is because, while TeraScale systems maintain QoS counters per interface, ExaScale systems maintain QoS counters per port-pipe. The matched packets counter, however, increments as expected.

FTOS Behavior: On Exascale 10G line cards, the granularity for rate shaping is 10Mbps so traffic is not always rate shaped according to the configured value. Specifically, if the configured value is below 5Mbps or a multiple of 5: for values less than 5Mbps, 0Mbps is received at remote end, and for values greater than or equal to 5Mbps, the remote end receives the next highest increment of 10; 15Mbps, for example, is rate shaped to 20Mbps.

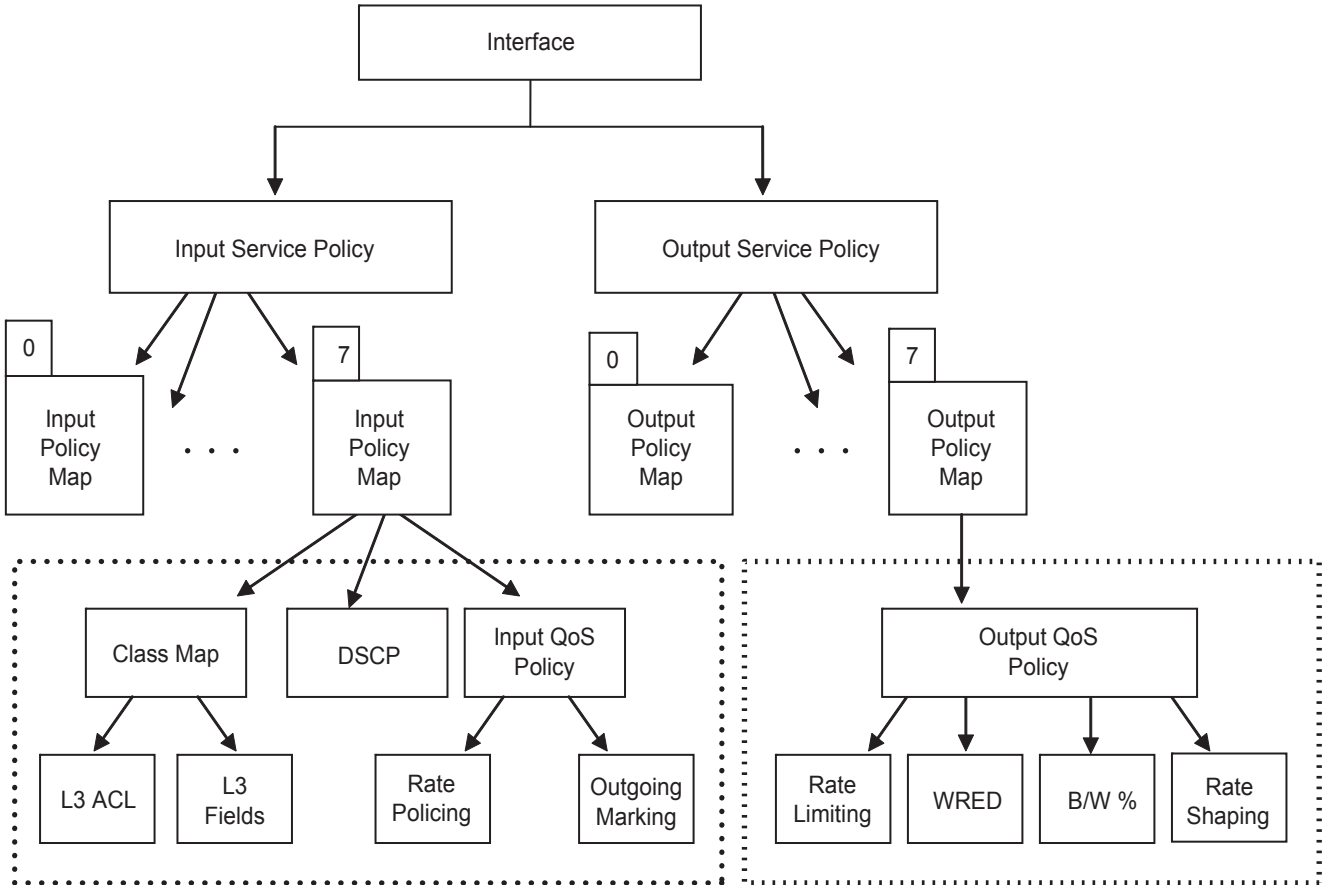
Figure 41-8. Applying Rate Shaping to Outgoing Traffic

```
FTOS#config
FTOS(conf)#interface gigabitethernet 1/0
FTOS(conf-if)#rate shape 500 50
FTOS(conf-if)#end
FTOS#
```

Policy-based QoS Configurations

Policy-based QoS configurations consist of the components shown in Figure 41-9.

Figure 41-9. Constructing Policy-based QoS Configurations



Classify Traffic

Class maps differentiate traffic so that you can apply separate quality of service policies to each class. For both class maps, Layer 2 and Layer 3, FTOS matches packets against match criteria in the order that you configure them.

Create a Layer 3 class map

A Layer 3 class map differentiates ingress packets based on DSCP value or IP precedence, and characteristics defined in an IP ACL. You may specify more than one DSCP and IP precedence value, but only one value must match to trigger a positive match for the class map.

1. Create a match-any class map using the command **class-map match-any** or a match-all class map using the command **class-map match-all** from CONFIGURATION mode, as shown in Figure 41-10.

2. Once you create a class-map, FTOS places you in CLASS MAP mode. From this mode, specify your match criteria using the command **match ip**, as shown in Figure 41-10. Match-any class maps allow up to five ACLs, and match-all class-maps allow only one ACL.
3. After you specify your match criteria, link the class-map to a queue using the command **service-queue** from POLICY MAP mode, as shown in Figure 41-10.

Figure 41-10. Using the Order Keyword in ACLs

```
FTOS(conf)#ip access-list standard acl1
FTOS(config-std-nacl)#permit 20.0.0.0/8
FTOS(config-std-nacl)#exit
FTOS(conf)#ip access-list standard acl2
FTOS(config-std-nacl)#permit 20.1.1.0/24 order 0
FTOS(config-std-nacl)#exit
FTOS(conf)#class-map match-all cmap1
FTOS(conf-class-map)#match ip access-group acl1
FTOS(conf-class-map)#exit
FTOS(conf)#class-map match-all cmap2
FTOS(conf-class-map)#match ip access-group acl2
FTOS(conf-class-map)#exit
FTOS(conf)#policy-map-input pmap
FTOS(conf-policy-map-in)#service-queue 7 class-map cmap1
FTOS(conf-policy-map-in)#service-queue 4 class-map cmap2
FTOS(conf-policy-map-in)#exit
FTOS(conf)#interface gig 1/0
FTOS(conf-if-gi-1/0)#service-policy input pmap
```

Create a Layer 2 class map

All class maps are Layer 3 by default; you can create a Layer 2 class map by specifying the option **layer2** with the **class-map** command. A Layer 2 class map differentiates traffic according to 802.1p value and/or characteristics defined in a MAC ACL.

1. Create a match-any class map using the command **class-map match-any** or a match-all class map using the command **class-map match-all** from CONFIGURATION mode, and enter the keyword **layer2**.
2. Once you create a class-map, FTOS places you in CLASS MAP mode. From this mode, specify your match criteria using the command **match mac**. Match-any class maps allow up to five access-lists, and match-all class-maps allow only one. You can match against only one VLAN ID.
3. After you specify your match criteria, link the class-map to a queue using the command **service-queue** from POLICY MAP mode.

The following configuration maps each queue to a VLAN (you can map 8 VLAN to 8 queues on the E-Series, and 4 VLANs to 4 queues on the C-Series and S-Series).

```
class-map match-any c0 layer2
match mac vlan 3
class-map match-any c1 layer2
match mac vlan 4
policy-map-input p0 layer2
service-queue 0 class-map c0
service-queue 1 class-map c1
```

Determine the order in which ACLs are used to classify traffic

When you link class-maps to queues using the command **service-queue**, FTOS matches the class-maps according to queue priority (queue numbers closer to 0 have lower priorities). For example, in Figure 41-10, class-map *cmap2* is matched against ingress packets before *cmap1*.

ACLs *acl1* and *acl2* have overlapping rules because the address range 20.1.1.0/24 is within 20.0.0.0/8. Therefore, (without the keyword **order**) packets within the range 20.1.1.0/24 match positive against *cmap1* and are buffered in queue 7, though you intended for these packets to match positive against *cmap2* and be buffered in queue 4.

In cases such as these, where class-maps with overlapping ACL rules are applied to different queues, use the **order** keyword to specify the order in which you want to apply ACL rules, as shown in Figure 41-10. The order can range from 0 to 254. FTOS writes to the CAM ACL rules with lower order numbers (order numbers closer to 0) before rules with higher order numbers so that packets are matched as you intended. By default, all ACL rules have an order of 254.

Set DSCP values for egress packets based on flow

Match-any Layer 3 flows may have several match criteria. All flows that match at least one of the match criteria are mapped to the same queue since they are in the same class map. Setting a DSCP value from QOS-POLICY-IN mode (see Set a DSCP value for egress packets on page 861) assigns the *same* DSCP value to all of the matching flows in the class-map. The Flow-based DSCP Marking feature allows you to assign *different* DSCP to each match criteria CLASS-MAP mode using the option **set-ip-dscp** with the **match** command so that matching flows within a class map can have *different* DSCP values, as shown in Figure 41-11. The values you set from CLASS-MAP mode override the value you QoS input policy DSCP value, and packets matching the rule are marked with the specified value.

Figure 41-11. Marking Flows in the Same Queue with Different DSCP Values

```
FTOS#show run class-map
!
class-map match-any example-flowbased-dscp
 match ip access-group test set-ip-dscp 2
 match ip access-group test1 set-ip-dscp 4
 match ip precedence 7 set-ip-dscp 1

FTOS#show run qos-policy-input
!
qos-policy-input flowbased
 set ip-dscp 3

FTOS# show cam layer3 linecard 2 port-set 0
Cam   Port Dscp Proto Tcp   Src   Dst   SrcIp             DstIp             DSCP    Queue
Index                 Flag  Port  Port                                      Marking
-------------------------------------------------------------------------------------------
-----
16260 1     0   TCP   0x0   0     0     1.1.1.0/24        0.0.0.0/0         2       0
16261 1     0   UDP   0x0   0     0     2.2.2.2/32        0.0.0.0/0         4       0
16262 1    56   0     0x0   0     0     0.0.0.0/0         0.0.0.0/0         1       0
24451 1     0   0     0x0   0     0     0.0.0.0/0         0.0.0.0/0         -       0
```

Display configured class maps and match criteria

Display all class-maps or a specific class map using the command **show qos class-map** from EXEC Privilege mode.

FTOS Behavior: An explicit “deny any" rule in a Layer 3 ACL used in a (match any or match all) class-map creates a "default to Queue 0" entry in the CAM, which causes unintended traffic classification. Below, traffic is classified in two Queues, 1 and 2. Class-map ClassAF1 is “match any,” and ClassAF2 is “match all”.

```
FTOS#show running-config policy-map-input
!
policy-map-input PolicyMapIn
 service-queue 1 class-map ClassAF1 qos-policy QosPolicyIn-1
 service-queue 2 class-map ClassAF2 qos-policy QosPolicyIn-2
FTOS#show running-config class-map
!
class-map match-any ClassAF1
 match ip access-group AF1-FB1 set-ip-dscp 10
 match ip access-group AF1-FB2 set-ip-dscp 12
 match ip dscp 10 set-ip-dscp 14
!
class-map match-all ClassAF2
 match ip access-group AF2
 match ip dscp 18
FTOS#show running-config ACL
!
ip access-list extended AF1-FB1
 seq 5 permit ip host 23.64.0.2 any
 seq 10 deny ip any any
!
ip access-list extended AF1-FB2
 seq 5 permit ip host 23.64.0.3 any
 seq 10 deny ip any any
!
ip access-list extended AF2
 seq 5 permit ip host 23.64.0.5 any
 seq 10 deny ip any any
FTOS#show cam layer3-qos interface gigabitethernet 4/49
Cam   Port Dscp Proto Tcp   Src   Dst   SrcIp              DstIp              DSCP    Queue
Index                 Flag  Port  Port                                        Marking
-------------------------------------------------------------------------------------------
20416 1    18   IP    0x0   0     0     23.64.0.5/32       0.0.0.0/0          20      2
20417 1    18   IP    0x0   0     0     0.0.0.0/0          0.0.0.0/0          -       0
20418 1     0   IP    0x0   0     0     23.64.0.2/32       0.0.0.0/0          10      1
20419 1     0   IP    0x0   0     0     0.0.0.0/0          0.0.0.0/0          -       0
20420 1     0   IP    0x0   0     0     23.64.0.3/32       0.0.0.0/0          12      1
20421 1     0   IP    0x0   0     0     0.0.0.0/0          0.0.0.0/0          -       0
20422 1    10   0     0x0   0     0     0.0.0.0/0          0.0.0.0/0          14      1
24511 1     0   0     0x0   0     0     0.0.0.0/0          0.0.0.0/0          -       0
```

Above, the ClassAF1 does not classify traffic as intended. Traffic matching the first match criteria is classified to Queue 1, but all other traffic is classified to Queue 0 as a result of CAM entry 20419.

When the explicit “deny any” rule is removed from all three ACLs, the CAM reflects exactly the desired classification.

```
FTOS#show cam layer3-qos interface gigabitethernet 4/49
Cam   Port Dscp Proto Tcp   Src   Dst   SrcIp              DstIp              DSCP    Queue
Index                 Flag  Port  Port                                        Marking
-------------------------------------------------------------------------------------------
20416 1    18   IP    0x0   0     0     23.64.0.5/32       0.0.0.0/0          20      2
20417 1     0   IP    0x0   0     0     23.64.0.2/32       0.0.0.0/0          10      1
20418 1     0   IP    0x0   0     0     23.64.0.3/32       0.0.0.0/0          12      1
20419 1    10   0     0x0   0     0     0.0.0.0/0          0.0.0.0/0          14      1
24511 1     0   0     0x0   0     0     0.0.0.0/0          0.0.0.0/0          -       0
```

Create a QoS Policy

There are two types of QoS policies: input and output.

Input QoS policies regulate Layer 3 and Layer 2 ingress traffic. The regulation mechanisms for input QoS policies are rate policing and setting priority values. There are two types of input QoS policies: Layer 3 and Layer 2.

- Layer 3 QoS input policies allow you to rate police and set a DSCP or dot1p value.
- Layer 2 QoS input policies allow you to rate police and set a dot1p value.

Output QoS policies regulate Layer 3 egress traffic. The regulation mechanisms for output QoS policies are rate limiting, rate shaping, and WRED.

Note: When changing a "service-queue" configuration in a QoS policy map, all QoS rules are deleted and re-added automatically to ensure that the order of the rules is maintained. As a result, the Matched Packets value shown in the "show qos statistics" command is reset.

Create an input QoS policy

To create an input QoS policy:

1. Create a Layer 3 input QoS policy using the command **qos-policy-input** from CONFIGURATION mode. Create a Layer 2 input QoS policy by specifying the keyword **layer2** after the command **qos-policy-input**.
2. Once you create an input QoS policy, do one or more of the following:
   - Configure policy-based rate policing
   - Set a DSCP value for egress packets
   - Set a dot1p value for egress packets

Configure policy-based rate policing

Rate police ingress traffic using the command rate-police from QOS-POLICY-IN mode.

Set a DSCP value for egress packets

Set a DSCP value for egress packets based on ingress QoS classification, as shown in Figure 41-2. The 6 bits that are used for DSCP are also used to identify the queue in which traffic is buffered. When you set a DSCP value, FTOS displays an informational message advising you of the queue to which you should apply the QoS policy (using the command **service-queue** from POLICY-MAP-IN mode). If you apply the QoS policy to a queue *other than* the one specified in the informational message, FTOS replaces the first 3 bits in the DSCP field with the queue ID you specified.

Figure 41-12. Marking DSCP Values for Egress Packets

```
FTOS#config
FTOS(conf)#qos-policy-input my-input-qos-policy
FTOS(conf-qos-policy-in)#set ip-dscp 34
% Info: To set the specified DSCP value 34 (100-010 b) the QoS policy must be mapped to queue
4 (100 b).
FTOS(conf-qos-policy-in)#show config
!
qos-policy-input my-input-qos-policy
 set ip-dscp 34
FTOS(conf-qos-policy-in)#end

FTOS#
```

Set a dot1p value for egress packets

Set a dot1p value for egress packets using the command **set mac-dot1p** from QOS-POLICY-IN mode.

Create an output QoS policy

To create an output QoS policy:

1. Create an output QoS policy using the command **qos-policy-output** from CONFIGURATION mode.
2. Once you configure an output QoS policy, do one or more of the following
   - Configure policy-based rate limiting
   - Configure policy-based rate shaping
   - Allocate bandwidth to queue
   - Specify WRED drop precedence

Configure policy-based rate limiting

Configure policy-based rate limiting is supported only on platform E

Policy-based rate limiting is configured the same way as port-based rate limiting except that the command from QOS-POLICY-OUT mode is **rate-limit** rather than **rate limit** as it is in INTERFACE mode.

Configure policy-based rate shaping

Rate shape egress traffic using the command **rate-shape** from QOS-POLICY-OUT mode. Output QoS policy can be applied to an output policy map with a policy aggregate or to an specific queue. Per queue rate shaping is supported on C-Series and S-Series only; see Create Output Policy Maps on page 868.

```
FTOS#conf t
FTOS(conf)#qos-policy-output QosShape
FTOS(conf-qos-policy-out)# rate-shape 4 10
FTOS(conf-qos-policy-out)#show config
!
qos-policy-output
QosShape rate-shape 4 10
FTOS(conf-qos-policy-out)#exit
```

Allocate bandwidth to queue

The E-Series schedules unicast, multicast, and replication traffic for egress based on the Weighted Fair Queuing algorithm. The C-Series and S-Series schedule packets for egress based on Deficit Round Robin (DRR). These strategies both offer a guaranteed data rate.

To allocate an amount bandwidth to a queue using the command **bandwidth-percentage** on the E-Series.

To allocate bandwidth to queues on the C-Series and S-Series, assign each queue a weight ranging from 1 to 1024, in increments of 2^n, using the command **bandwidth-weight**. Table 41-3 shows the default bandwidth weights for each queue, and their equivalent percentage which is derived by dividing the bandwidth weight by the sum of all queue weights.

Table 41-3. Default Bandwidth Weights for C-Series and S-Series

Queue	Default Weight	Equivalent Percentage
0	1	6.67%
1	2	13.33%
2	4	26.67%
3	8	53.33%

There are two key differences between allocating bandwidth by weight on the C-Series and S-Series and allocating bandwidth by percentage on the E-Series:

1. Assigning a weight to one queue affects the amount of bandwidth that is allocated to other queues. Therefore, whenever you are allocating bandwidth to one queue, Dell Force10 recommends that you evaluate your bandwidth requirements for all other queues as well.
2. Because you are required to choose a bandwidth weight in increments of 2^n you may not be able to achieve exactly a target bandwidth allocation.

Table 41-4 shows an example of choosing bandwidth weights for all four queues to achieve a target bandwidth allocation.

Table 41-4. Assigning Bandwidth Weights for the C-Series and S-Series

Queue	Weight	Equivalent Percentage	Target Allocation
0	1	0.44%	1%
1	64	28.44%	25%
2	128	56.89%	60%
3	32	14.22%	14%

Specify WRED drop precedence

Specify WRED drop precedence is supported only on platform E

Specify a WRED profile to yellow and/or green traffic using the command **wred** from QOS-POLICY-OUT mode. See Apply a WRED profile to traffic on page 871.

Create Policy Maps

There are two types of policy maps: input and output.

Create Input Policy Maps

There are two types of input policy-maps: Layer 3 and Layer 2.

1. Create a Layer 3 input policy map using the command **policy-map-input** from CONFIGURATION mode. Create a Layer 2 input policy map by specifying the keyword **layer2** with the **policy-map-input** command.
2. Once you create an input policy map, do one or more of the following:
   - Apply a class-map or input QoS policy to a queue
   - Apply an input QoS policy to an input policy map
   - Honor DSCP values on ingress packets
   - Honoring dot1p values on ingress packets
3. Apply the input policy map to an interface. See page 868.

FTOS Behavior: On ExaScale, FTOS cannot classify protocol traffic on a Layer 2 interface using Layer 3 policy map. The packets always take the default queue, Queue 0, and cannot be rate-policed.

Apply a class-map or input QoS policy to a queue

Assign an input QoS policy to a queue using the command **service-queue** from POLICY-MAP-IN mode.

Apply an input QoS policy to an input policy map

Apply an input QoS policy to an input policy map using the command **policy-aggregate** from POLICY-MAP-IN mode.

Honor DSCP values on ingress packets

FTOS provides the ability to honor DSCP values on ingress packets using Trust DSCP feature. Enable this feature using the command **trust diffserv** from POLICY-MAP-IN mode. Table 41-5 lists the standard DSCP definitions, and indicates to which queues FTOS maps DSCP values. When Trust DSCP is configured the matched packets and matched bytes counters are not incremented in **show qos statistics**.

Table 41-5. Default DSCP to Queue Mapping

DSCP/CP hex range (XXX)	**DSCP Definition**	**Traditional IP Precedence**	**E-Series Internal Queue ID**	**C-Series Internal Queue ID**	**S-Series Internal Queue ID**	**DSCP/CP** decimal
111XXX		Network Control	7	3	3	48–63
110XXX		Internetwork Control	6	3	3	
101XXX	EF (Expedited Forwarding)	CRITIC/ECP	5	2	2	32–47
100XXX	AF4 (Assured Forwarding)	Flash Override	4	2	2	
011XXX	AF3	Flash	3	1	1	16–31
010XXX	AF2	Immediate	2	1	1	
001XXX	AF1	Priority	1	0	0	0–15
000XXX	BE (Best Effort)	Best Effort	0	0	0	

Honoring dot1p values on ingress packets

FTOS provides the ability to honor dot1p values on ingress packets with the Trust dot1p feature. Enable Trust dot1p using the command **trust dot1p** from POLICY-MAP-IN mode. Table 41-6 specifies the queue to which the classified traffic is sent based on the dot1p value.

Table 41-6. Default dot1p to Queue Mapping

dot1p	E-Series Queue ID	C-Series Queue ID	S-Series Queue ID
0	2	1	1
1	0	0	0
2	1	0	0
3	3	1	1
4	4	2	2
5	5	2	2
6	6	3	3
7	7	3	3

The dot1p value is also honored for frames on the default VLAN; see Priority-tagged Frames on the Default VLAN.

Fall Back to trust diffserve or dot1p

Fall Back to trust diffserve or dot1p is available only on platforms: E

When using QoS service policies with multiple class maps, you can configure FTOS to use the incoming DSCP or dot1p marking as a secondary option for packet queuing in the event that no match occurs in the class maps.

When class-maps are used, traffic is matched against each class-map sequentially from first to last. The sequence is based on the priority of the rules, as follows:

1. rules with lowest priority, or in the absence of a priority configuration,
2. rules of the next numerically higher queue

By default, if no match occurs, the packet is queued to the default queue, Queue 0.

In the following configuration, packets are classified to queues using the three class maps:

```
!
policy-map-input input-policy
 service-queue 1 class-map qos-BE1
 service-queue 3 class-map qos-AF3
 service-queue 4 class-map qos-AF4
!
class-map match-any qos-AF3
 match ip dscp 24
 match ip access-group qos-AF3-ACL
!
class-map match-any qos-AF4
 match ip dscp 32
 match ip access-group qos-AF4-ACL
!
class-map match-all qos-BE1
 match ip dscp 0
 match ip access-group qos-BE1-ACL
```

The packet classification logic for the above configuration is as follows:

1. Match packets against match-any qos-AF4. If a match exists, queue the packet as AF4 in Queue 4, and if no match exists, go to the next class map.
2. Match packets against match-any qos-AF3. If a match exists, queue the packet as AF3 in Queue 3, and if no match exists, go to the next class map.
3. Match packets against match-all qos-BE1. If a match exists, queue the packet as BE1, and if no match exists, queue the packets to the default queue, Queue 0.

You can optionally classify packets using their DSCP marking, instead of placing packets in Queue 0, if no match occurs. In the above example, if no match occurs against match-all qos-BE1, the classification logic continues:

4. Queue the packet according to the DSCP marking. The DSCP to Queue mapping will be as per the Table 41-5.

The behavior is similar for **trust dot1p fallback** in a Layer2 input policy map; the dot1p-to-queue mapping is according to Table 41-6.

To enable Fall Back to trust diffserve or dot1p:

Task	Command Syntax	Command Mode
Classify packets according to their DSCP value as a secondary option in case no match occurs against the configured class maps.	**trust** {**diffserve | dot1p**} **fallback**	POLICY-MAP-IN

Mapping dot1p values to service queues

Mapping dot1p values to service queues is available only on platforms: C S

On the C-Series and S-Series all traffic is by default mapped to the same queue, Queue 0. If you honor dot1p on ingress, then you can create service classes based the queueing strategy in Table 41-6 using the command **service-class dynamic dot1p** from INTERFACE mode. You may apply this queuing strategy globally by entering this command from CONFIGURATION mode.

- All dot1p traffic is mapped to Queue 0 unless **service-class dynamic dot1p** is enabled on an interface or globally.
- Layer 2 or Layer 3 service policies supersede dot1p service classes.

Guaranteeing bandwidth to dot1p-based service queues

Guarantee a minimum bandwidth to queues globally from CONFIGURATION mode with the command **service-class bandwidth-weight**. The command is applied in the same way as the bandwidth-weight command in an output QoS policy (see Allocate bandwidth to queue on page 863). The **bandwidth-weight** command in QOS-POLICY-OUT mode supersedes the **service-class bandwidth-weight command**.

Apply an input policy map to an interface

Apply an input policy map to an interface using the command **service-policy input** from INTERFACE mode. Specify the keyword **layer2** if the policy map you are applying a Layer 2 policy map; in this case, the INTERFACE must be in switchport mode. You can apply the same policy map to multiple interfaces, and you can modify a policy map after you apply it.

- You cannot apply a class-map and QoS policies to the same interface.
- You cannot apply an input Layer 2 QoS policy on an interface you also configure with **vlan-stack access**.
- If you apply a service policy that contains an ACL to more than one interface, FTOS uses ACL optimization to conserves CAM space. The ACL Optimization behavior detects when an ACL already exists in the CAM and rather than writing it to the CAM multiple times.

Create Output Policy Maps

Create Output Policy Maps is supported only on platform E

1. Create an output policy map using the command **policy-map-output** from CONFIGURATION mode.
2. Once you create an output policy map, do one or more of the following:
   - Apply an output QoS policy to a queue
   - Specify an aggregate QoS policy
   - Apply an output policy map to an interface
3. Apply the policy map to an interface. See page 61.

Apply an output QoS policy to a queue

Apply an output QoS policy to queues using the command **service-queue** from INTERFACE mode.

Specify an aggregate QoS policy

Specify an aggregate QoS policy using the command **policy-aggregate** from POLICY-MAP-OUT mode.

Apply an output policy map to an interface

Apply an input policy map to an interface using the command **service-policy output** from INTERFACE mode. You can apply the same policy map to multiple interfaces, and you can modify a policy map after you apply it.

QoS Rate Adjustment

The Ethernet packet format consists of:

- Preamble: 7 bytes Preamble
- Start Frame Delimiter (SFD): 1 byte
- Destination MAC Address: 6 bytes
- Source MAC Address: 6 bytes
- Ethernet Type/Length: 2 bytes
- Payload: (variable)
- Cyclic Redundancy Check (CRC): 4 bytes
- Inter-frame Gap (IFG): (variable)

By default, while rate limiting, policing, and shaping, FTOS does not include the Preamble, SFD, or the IFG fields. These fields are overhead; only the fields from MAC Destination Address to the CRC are used for forwarding and are included in these rate metering calculations. You can optionally include overhead fields in rate metering calculations by enabling QoS Rate Adjustment.

QoS Rate Adjustment is disabled by default, and no **qos-rate-adjust** is listed in the running-configuration.

Task	Command Syntax	Command Mode
Include a specified number of bytes of packet overhead to include in rate limiting, policing, and shaping calculations. For example, to include the Preamble and SFD, enter **qos-rate-adjust** 8. For variable length overhead fields you must know the number of bytes you want to include.	**qos-rate-adjust** *overhead-bytes* Default: Disabled C-Series and S-Series Range: 1-31 E-Series Range: 1-144	CONFIGURATION

Strict-priority Queueing

You can assign strict-priority to one unicast queue, 1-7, using the command **strict-priority** from CONFIGURATION mode. Strict-priority means that FTOS dequeues all packets from the assigned queue before servicing any other queues.

- The **strict-priority** supersedes **bandwidth-percentage** an **bandwidth-weight** percentage configurations.
- A queue with strict-priority can starve other queues in the same port-pipe.
- On the E-Series, this configuration is applied to the queue on both ingress and egress.

Weighted Random Early Detection

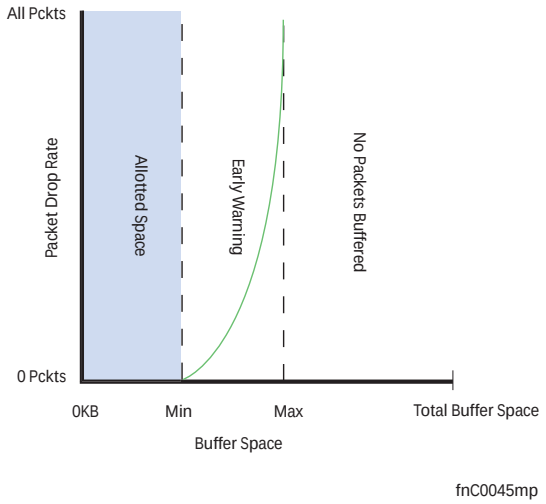
Weighted Random Early Detection is supported only on platform E

Weighted Random Early Detection (WRED) congestion avoidance mechanism that drops packets to prevent buffering resources from being consumed.

Traffic is a mixture of various kinds of packets. The rate at which some types of packets arrive might be greater than others. In this case, the space on the BTM (ingress or egress) can be consumed by only one or a few types of traffic, leaving no space for other types. A WRED profile can be applied to a policy-map so that specified traffic can be prevented from consuming too much of the BTM resources.

WRED uses a profile to specify minimum and maximum threshold values. The minimum threshold is the allotted buffer space for specified traffic, for example 1000KB on egress. If the 1000KB is consumed, packets will be dropped randomly at an exponential rate until the maximum threshold is reached (Figure 41-13); this is the “early detection” part of WRED. If the maximum threshold—2000KB, for example—is reached, then all incoming packets are dropped until less than 2000KB of buffer space is consumed by the specified traffic.

Figure 41-13. Packet Drop Rate for WREDI



You can create a custom WRED profile or use on of the five pre-defined profiles listed in Table 41-7.

Table 41-7. Pre-defined WRED Profiles

Default Profile Name	Minimum Threshold	Maximum Threshold
wred_drop	0	0
wred_ge_y	1024	2048
wred_ge_g	2048	4096
wred_teng_y	4096	8192
wred_teng_g	8192	16384

Create WRED Profiles

To create a WRED profile:

1. Create a WRED profile using the command **wred** from CONFIGURATION mode.
2. The command **wred** places you in WRED mode. From this mode, specify minimum and maximum threshold values using the command **threshold**.

Apply a WRED profile to traffic

Once you create a WRED profile you must specify to which traffic FTOS should apply the profile.

FTOS assigns a color (also called drop precedence)—red, yellow, or green—to each packet based on it DSCP value before queuing it. DSCP is a 6 bit field. Dell Force10 uses the first three bits of this field (DP) to determine the drop precedence. DP values of 110 and 100 map to yellow, and all other values map to green. If you do not configure FTOS to honor DSCP values on ingress (Honor DSCP values on ingress packets on page 865) see all traffic defaults to green drop precedence.

Assign a WRED profile to either yellow or green traffic from QOS-POLICY-OUT mode using the command **wred**.

Configure WRED for Storm Control

Configure WRED for Storm Control is supported only on platform E

Storm control limits the percentage of the total bandwidth that broadcast traffic can consume on an interface (if configured locally) or on all interfaces (if configured globally). For **storm-control broadcast 50 out**, the total bandwidth that broadcast traffic can consume on egress on a 1Gbs interface is 512Mbs. The method by which packets are selected to be dropped is the "tail-drop" method, where packets exceeding the specified rate are dropped.

WRED can be used in combination with storm control to regulate broadcast and unknown-unicast traffic. This feature is available through an additional option in command **storm-control** [**broadcast** | **unknown-unicast**] at CONFIGURATION. See the *FTOS Command Line Reference* for information on using this command.

Using the command **storm-control broadcast 50 out wred-profile**, for example, first the total bandwidth that broadcast traffic can consume is reduced to 50% of line rate. Even though broadcast traffic is restricted, the rate of outgoing broadcast traffic might be greater than other traffic, and if so, broadcast packets would consume too much buffer space. So, the **wred-profile** option is added to limit the amount of buffer space that broadcast traffic can consume.

Display Default and Configured WRED Profiles

Display default and configured WRED profiles and their threshold values using the command **show qos wred-profile** from EXEC mode, as shown in Figure 41-14.

Figure 41-14. Displaying WRED Profiles

```
FTOS#show qos wred-profile

Wred-profile-name       min-threshold   max-threshold
wred_drop               0               0
wred_ge_y               1000            2000
wred_ge_g               2000            4000
wred_teng_y             4000            8000
wred_teng_g             8000            16000
```

Display WRED Drop Statistics

Display the number of packets FTOS dropped by WRED Profile using the command **show qos statistics** from EXEC Privilege mode, as shown in Figure 41-15.

Figure 41-15. show qos statistics Command Example

```
 FTOS#show qos statistics wred-profile
Interface Gi 5/11
Queue#  Drop-statistic  WRED-name        Min      Max      Dropped Pkts

  0     Green           WRED1           10       100      51623
        Yellow          WRED2           20       100      51300
        Out of Profile                                    0
  1     Green           WRED1           10       100      52082
        Yellow          WRED2           20       100      51004
        Out of Profile                                    0
  2     Green           WRED1           10       100      50567
        Yellow          WRED2           20       100      49965
        Out of Profile                                    0
  3     Green           WRED1           10       100      50477
        Yellow          WRED2           20       100      49815
        Out of Profile                                    0
  4     Green           WRED1           10       100      50695
        Yellow          WRED2           20       100      49476
        Out of Profile                                    0
  5     Green           WRED1           10       100      50245
        Yellow          WRED2           20       100      49535
        Out of Profile                                    0
  6     Green           WRED1           10       100      50033
        Yellow          WRED2           20       100      49595
        Out of Profile                                    0
  7     Green           WRED1           10       100      50474
        Yellow          WRED2           20       100      49522
        Out of Profile                                    0
FTOS#
```

FTOS Behavior: The C-Series fetches the per-queue packet count via class-maps. The count is the number of packets matching the ACL entries in class-map. Every time the class-map or policy-map is modified, the ACL entries are re-written to the Forwarding Processor, and the queue statistics are cleared. This behavior is different from the E-Series. The E-Series fetches the packet count directly from counters at each queue, which allows queue statistics to persist until explicitly cleared via the CLI.

Allocating Bandwidth to Multicast Queues

Allocating Bandwidth to Multicast Queues is supported on platform: E

The E-Series has 128 multicast queues per port-pipe, which are transparent, and eight unicast queues per port. You can allocate a specific bandwidth percentage per port-pipe to multicast traffic using the command **queue egress multicast bandwidth-percentage** from CONFIGURATION mode.

- If you configure **bandwidth-percentage** for unicast only, 1/8 of the port bandwidth is reserved for multicast, and the remaining bandwidth is distributed based on your configuration.
- If you configure multicast bandwidth, after assigning the specified amount of bandwidth to multicast the remaining bandwidth is distributed according to the WFQ algorithm.
- If you configure **bandwidth-percentage** for both unicast and multicast, then bandwidth is assigned based on your configuration for multicast *then* unicast (based on the remaining available bandwidth), and the remaining bandwidth is distributed among the other queues.

For example, if you configure 70% bandwidth to multicast, 80% bandwidth to one queue in unicast and 0 % to all remaining unicast queues, then first, FTOS assigns 70% bandwidth to multicast, then FTOS derives the 80% bandwidth for unicast from the remaining 30% of total bandwidth.

Pre-calculating Available QoS CAM Space

Pre-calculating Available QoS CAM Space is supported on platforms: C E S

Before version 7.3.1 there was no way to measure the number of CAM entries a policy-map would consume (the number of CAM entries that a rule uses is not predictable; 1 to 16 entries might be used per rule depending upon its complexity). Therefore, it was possible to apply to an interface a policy-map that requires more entries than are available. In this case, the system writes as many entries as possible, and then generates an CAM-full error message (Message 1). The partial policy-map configuration might cause unintentional system behavior.

Message 1 QoS CAM Region Exceeded

```
%EX2YD:12 %DIFFSERV-2-DSA_QOS_CAM_INSTALL_FAILED: Not enough space in L3 Cam(PolicyQos) for
class 2 (Gi 12/20) entries on portpipe 1 for linecard 12
%EX2YD:12 %DIFFSERV-2-
DSA_QOS_CAM_INSTALL_FAILED: Not enough space in L3 Cam(PolicyQos) for class 5 (Gi 12/22)
entries on portpipe 1 for linecard 12
```

The command **test cam-usage** enables you to verify that there are enough available CAM entries *before* applying a policy-map to an interface so that you avoid exceeding the QoS CAM space and partial configurations. This command measures the size of the specified policy-map and compares it to the available CAM space in a partition for a specified port-pipe.

Test the policy-map size against the CAM space for a specific port-pipe or all port-pipes using these commands:

- **test cam-usage service-policy input** *policy-map* {**linecard | stack-unit** } *number* **port-set** *number*
- **test cam-usage service-policy input** *policy-map* {**linecard | stack-unit** } *all*

The output of this command, shown in Figure 41-16, displays:

- the estimated number of CAM entries the policy-map will consume
- whether or not the policy-map can be applied
- the number of interfaces in a port-pipe to which the policy-map can be applied

Specifically:

- **Available CAM** is the available number of CAM entries in the specified CAM partition for the specified line card or stack-unit port-pipe.
- **Estimated CAM** is the estimated number of CAM entries that the policy will consume when it is applied to an interface.

- **Status** indicates whether or not the specified policy-map can be completely applied to an interface in the port-pipe.
  - **Allowed** indicates that the policy-map can be applied because the estimated number of CAM entries is less or equal to the available number of CAM entries. The number of interfaces in the port-pipe to which the policy-map can be applied is given in parenthesis.
  - **Exception** indicates that the number of CAM entries required to write the policy-map to the CAM is greater than the number of available CAM entries, and therefore the policy-map cannot be applied to an interface in the specified port-pipe.

Note: The command **show cam-usage** provides much of the same information as **test cam-usage**, but whether or not a policy-map can be successfully applied to an interface cannot be determined without first measuring how many CAM entries the policy-map would consume; the command **test cam-usage** is useful because it provides this measurement.

Figure 41-16. test cam-usage Command Example

```
FTOS# test cam-usage service-policy input pmap_l2 linecard 0 port-set 0

Linecard | Port-pipe | CAM Partition | Available CAM | Estimated CAM | Status
==============================================================================
0           0           L2ACL           500            200        Allowed(2)
```

Viewing QoS CAM Entries

Viewing QoS CAM Entries is supported only on platform E

- View Layer 2 QoS CAM entries using the command **show cam layer3-qos** from EXEC Privilege mode.
- View Layer 3 QoS CAM entries using the command **show cam layer2-qos** from EXEC Privilege mode.

42

Routing Information Protocol

Routing Information Protocol is supported only on platforms: C E S

RIP is supported on the S-Series following the release of FTOS version 7.8.1.0, and on the C-Series with FTOS versions 7.6.1.0 and after.

RIP is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

Routing Information Protocol (RIP) is based on a distance-vector algorithm, it tracks distances or hop counts to nearby routers when establishing network connections.

RIP protocol standards are listed in the Appendix 63, Standards Compliance chapter.

Protocol Overview

RIP is the oldest interior gateway protocol. There are two versions of RIP: RIP version 1 (RIPv1) and RIP version 2 (RIPv2). These versions are documented in RFCs 1058 and 2453.

RIPv1

RIPv1 learns where nodes in a network are located by automatically constructing a routing data table. The routing table is established after RIP sends out one or more broadcast signals to all adjacent nodes in a network. Hop counts of these signals are tracked and entered into the routing table, which defines where nodes in the network are located.

The information that is used to update the routing table is sent as either a request or response message. In RIPv1, automatic updates to the routing table are performed as either one-time requests or periodic responses (every 30 seconds). RIP transports its responses or requests by means of UDP over port 520.

RIP must receive regular routing updates to maintain a correct routing table. Response messages containing a router's full routing table are transmitted every 30 seconds. If a router does not send an update within a certain amount of time, the hop count to that route is changed to unreachable (a route hop metric of 16 hops). Another timer sets the amount of time before the unreachable routes are removed from the routing table.

This first RIP version does not support VLSM or CIDR and is not widely used.

RIPv2

RIPv2 adds support for subnet fields in the RIP routing updates, thus qualifying it as a classless routing protocol. The RIPv2 message format includes entries for route tags, subnet masks, and next hop addresses. Another enhancement included in RIPv2 is multicasting for route updates on IP multicast address 224.0.0.9.

Implementation Information

FTOS supports both versions of RIP and allows you to configure one version globally and the other version or both versions on the interfaces. The C-Series and E-Series both support 1,000 RIP routes.

Table 42-1 displays the defaults for RIP in FTOS.

Table 42-1. RIP Defaults in FTOS

Feature	Default
Interfaces running RIP	Listen to RIPv1 and RIPv2 Transmit RIPv1
RIP timers	update timer = 30 seconds invalid timer = 180 seconds holddown timer = 180 seconds flush timer = 240 seconds
Auto summarization	Enabled
ECMP paths supported	16

Configuration Information

By default, RIP is disabled in FTOS. To configure RIP, you must use commands in two modes: ROUTER RIP and INTERFACE. Commands executed in the ROUTER RIP mode configure RIP globally, while commands executed in the INTERFACE mode configure RIP features on that interface only.

RIP is best suited for small, homogeneous networks. All devices within the RIP network must be configured to support RIP if they are to participate in the RIP.

Configuration Task List for RIP

- Enable RIP globally on page 879 (mandatory)
- Configure RIP on interfaces on page 880 (optional)
- Control RIP routing updates on page 881 (optional)
- Set send and receive version on page 882 (optional)
- Generate a default route on page 884 (optional)
- Control route metrics on page 885 (optional)
- Summarize routes on page 884 (optional)
- Control route metrics on page 885
- Debug RIP on page 885

For a complete listing of all commands related to RIP, refer to the *FTOS Command Reference.*

Enable RIP globally

By default, RIP is not enabled in FTOS. To enable RIP, use the following commands in sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**router rip**	CONFIGURATION	Enter ROUTER RIP mode and enable the RIP process on FTOS.
2	**network** *ip-address*	ROUTER RIP	Assign an IP network address as a RIP network to exchange routing information. You can use this command multiple times to exchange RIP information with as many RIP networks as you want.

After designating networks with which the system is to exchange RIP information, ensure that all devices on that network are configured to exchange RIP information.

The FTOS default is to send RIPv1, and to receive RIPv1 and RIPv2. To change the RIP version globally, use the **version** command in the ROUTER RIP mode.

When RIP is enabled, you can view the global RIP configuration by using the **show running-config** command in the EXEC mode or the **show config** command (Figure) in the ROUTER RIP mode.

Figure 42-1. show config Command Example in ROUTER RIP mode

```
FTOS(conf-router_rip)#show config
!
router rip
 network 10.0.0.0
FTOS(conf-router_rip)#
```

When the RIP process has learned the RIP routes, use the **show ip rip database** command in the EXEC mode to view those routes (Figure 385).

Figure 42-2. show ip rip database Command Example (Partial)

```
FTOS#show ip rip database
Total number of routes in RIP database: 978
160.160.0.0/16
        [120/1] via 29.10.10.12, 00:00:26, Fa 0/0
160.160.0.0/16          auto-summary
2.0.0.0/8
        [120/1] via 29.10.10.12, 00:01:22, Fa 0/0
2.0.0.0/8               auto-summary
4.0.0.0/8
        [120/1] via 29.10.10.12, 00:01:22, Fa 0/0
4.0.0.0/8               auto-summary
8.0.0.0/8
        [120/1] via 29.10.10.12, 00:00:26, Fa 0/0
8.0.0.0/8               auto-summary
12.0.0.0/8
        [120/1] via 29.10.10.12, 00:00:26, Fa 0/0
12.0.0.0/8              auto-summary
20.0.0.0/8
        [120/1] via 29.10.10.12, 00:00:26, Fa 0/0
20.0.0.0/8              auto-summary
29.10.10.0/24           directly connected,Fa 0/0
29.0.0.0/8              auto-summary
31.0.0.0/8
        [120/1] via 29.10.10.12, 00:00:26, Fa 0/0
31.0.0.0/8              auto-summary
192.162.2.0/24
        [120/1] via 29.10.10.12, 00:01:21, Fa 0/0
192.162.2.0/24          auto-summary
192.161.1.0/24
        [120/1] via 29.10.10.12, 00:00:27, Fa 0/0
192.161.1.0/24          auto-summary
192.162.3.0/24
        [120/1] via 29.10.10.12, 00:01:22, Fa 0/0
192.162.3.0/24          auto-summary
```

To disable RIP globally, use the **no router rip** command in the CONFIGURATION mode.

Configure RIP on interfaces

When you enable RIP globally on the system, interfaces meeting certain conditions start receiving RIP routes. By default, interfaces that are enabled and configured with an IP address in the same subnet as the RIP network address receive RIPv1 and RIPv2 routes and send RIPv1 routes.

Assign IP addresses to interfaces that are part of the same subnet as the RIP network identified in the **network** command syntax.

Control RIP routing updates

By default, RIP broadcasts routing information out all enabled interfaces, but you can configure RIP to send or to block RIP routing information, either from a specific IP address or a specific interface. To control which devices or interfaces receive routing updates, you must configure a direct update to one router and configure interfaces to block RIP updates from other sources.

To control the source of RIP route information, use the following commands, in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
neighbor *ip-address*	ROUTER RIP	Define a specific router to exchange RIP information between it and the Dell Force10 system. You can use this command multiple times to exchange RIP information with as many RIP networks as you want.
passive-interface *interface*	ROUTER RIP	Disable a specific interface from sending or receiving RIP routing information.

Another method of controlling RIP (or any routing protocol) routing information is to filter the information through a prefix list. A prefix lists is applied to incoming or outgoing routes. Those routes must meet the conditions of the prefix list; if not, FTOS drops the route. Prefix lists are globally applied on all interfaces running RIP. Configure the prefix list in the PREFIX LIST mode prior to assigning it to the RIP process.

For configuration information on prefix lists, see Chapter 17, IP Access Control Lists, Prefix Lists, and Route-maps, on page 47.

To apply prefix lists to incoming or outgoing RIP routes, use the following commands in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
distribute-list *prefix-list-name* **in**	ROUTER RIP	Assign a configured prefix list to all incoming RIP routes.
distribute-list *prefix-list-name* **out**	ROUTER RIP	Assign a configured prefix list to all outgoing RIP routes.

In addition to filtering routes, you can add routes from other routing instances or protocols to the RIP process. With the **redistribute** command syntax, you can include OSPF, static, or directly connected routes in the RIP process.

To add routes from other routing instances or protocols, use any of the following commands in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
redistribute {**connected** | **static**} [**metric** *metric-value*] [**route-map** *map-name*]	ROUTER RIP	Include directly connected or user-configured (static) routes in RIP. • *metric* range: 0 to 16 • *map-name*: name of a configured route map.
redistribute isis [**level-1** | **level-1-2** | **level-2**] [**metric** *metric-value*] [**route-map** *map-name*]	ROUTER RIP	Include IS-IS routes in RIP. • *metric* range: 0 to 16 • *map-name*: name of a configured route map. **Note:** IS-IS is not supported on the S-Series platform.
redistribute ospf *process-id* [**match external** {**1 | 2**} | **match internal**] [**metric** *value*] [**route-map** *map-name*]	ROUTER RIP	Include specific OSPF routes in RIP. Configure the following parameters: • *process-id* range: 1 to 65535 • *metric* range: 0 to 16 • *map-name*: name of a configured route map.

To view the current RIP configuration, use the **show running-config** command in the EXEC mode or the **show config** command in the ROUTER RIP mode.

Set send and receive version

To specify the RIP version, use the **version** command in the ROUTER RIP mode. To set an interface to receive only one or the other version, use the **ip rip send version** or the **ip rip receive version** commands in the INTERFACE mode.

To change the RIP version globally in FTOS, use the following command in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
version {**1** | **2**}	ROUTER RIP	Set the RIP version sent and received on the system.

You can set one RIP version globally on the system. This command sets the RIP version for RIP traffic on the interfaces participating in RIP unless the interface was specifically configured for a specific RIP version.

Use the **show config** command in the ROUTER RIP mode to see whether the **version** command is configured. You can also use the **show ip protocols** command in the EXEC mode to view the routing protocols configuration.

Figure 42-3 shows an example of the RIP configuration after the ROUTER RIP mode **version** command is set to RIPv2. When the ROUTER RIP mode **version** command is set, the interface (GigabitEthernet 0/0) participating in the RIP process is also set to send and receive RIPv2.

Figure 42-3. show ip protocols Command Example

```
FTOS#show ip protocols

 Routing Protocols is RIP
 Sending updates every 30 seconds, next due in 23
 Invalid after 180 seconds, hold down 180, flushed after 240
 Output delay 8 milliseconds between packets
 Automatic network summarization is in effect
 Outgoing filter for all interfaces is
 Incoming filter for all interfaces is
 Default redistribution metric is  1
 Default version control: receive version 2, send version 2
        Interface       Recv  Send
        GigabitEthernet 0/0   2     2
 Routing for Networks:
         10.0.0.0

 Routing Information Sources:
 Gateway         Distance      Last Update

 Distance: (default is 120)

FTOS#
```

RIPv2 configured globally and on the interface.

To configure the interfaces to send or receive different RIP versions from the RIP version configured globally, use either of the following commands in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ip rip receive version [**1**] [**2**]	INTERFACE	Set the RIP version(s) received on that interface.
ip rip send version [**1**] [**2**]	INTERFACE	Set the RIP version(s) sent out on that interface.

To configure an interface to receive or send both versions of RIP, include 1 and 2 in the command syntax. Figure 42-4 displays the command syntax for sending both RIPv1 and RIPv2 and receiving only RIPv2.

Figure 42-4. Configuring an interface to send both versions of RIP

```
FTOS(conf-if)#ip rip send version 1 2
FTOS(conf-if)#ip rip receive version 2
```

The **show ip protocols** command example Figure 42-5 confirms that both versions are sent out that interface. This interface no longer sends and receives the same RIP versions as FTOS does globally.

Figure 42-5. show ip protocols Command Example

```
FTOS#show ip protocols

 Routing Protocols is RIP
 Sending updates every 30 seconds, next due in 11
 Invalid after 180 seconds, hold down 180, flushed after 240
 Output delay 8 milliseconds between packets
 Automatic network summarization is in effect
 Outgoing filter for all interfaces is
 Incoming filter for all interfaces is
 Default redistribution metric is  1
 Default version control: receive version 2, send version 2
       Interface       Recv  Send
       FastEthernet 0/0   2     1 2
 Routing for Networks:
        10.0.0.0

 Routing Information Sources:
 Gateway         Distance      Last Update

 Distance: (default is 120)

FTOS#
```

RIPv2 configured globally

Different RIP versions configured for this interface

Generate a default route

Traffic is forwarded to the default route when the traffic's network is not explicitly listed in the routing table. Default routes are not enabled in RIP unless specified. Use the **default-information originate** command in the ROUTER RIP mode to generate a default route into RIP. In FTOS, default routes received in RIP updates from other routes are advertised if the **default-information originate** command is configured.

To configure FTOS to generate a default route, use the following command in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
default-information originate [**always**] [**metric** *value*] [**route-map** *route-map-name*]	ROUTER RIP	Specify the generation of a default route in RIP. Configure the following parameters: • **always**: enter this keyword to always generate a default route. • *value* range: 1 to 16. • *route-map-name*: name of a configured route map.

Use the **show config** command in the ROUTER RIP mode to confirm that the default route configuration is completed.

Summarize routes

Routes in the RIPv2 routing table are summarized by default, thus reducing the size of the routing table and improving routing efficiency in large networks. By default, the **autosummary** command in the ROUTER RIP mode is enabled and summarizes RIP routes up to the classful network boundary.

If you must perform routing between discontiguous subnets, disable automatic summarization. With automatic route summarization disabled, subnets are advertised.

The command **autosummary** requires no other configuration commands. To disable automatic route summarization, in the ROUTER RIP mode, enter **no autosummary**.

Note: If the **ip split-horizon** command is enabled on an interface, then the system does not advertise the summarized address.

Control route metrics

As a distance-vector protocol, RIP uses hop counts to determine the best route, but sometimes the shortest hop count is a route over the lowest-speed link. To manipulate RIP routes so that the routing protocol prefers a different route, you must manipulate the route by using the **offset** command.

Exercise caution when applying an **offset** command to routers on a broadcast network, as the router using the **offset** command is modifying RIP advertisements before sending out those advertisements.

The **distance** command also allows you to manipulate route metrics. Use the command to assign different weights to routes so that the ones with the lower weight or administrative distance assigned are preferred.

To set route metrics, use either of the following commands in the ROUTER RIP mode:

Command Syntax	Command Mode	Purpose
distance *weight* [*ip-address mask* [*access-list-name*]]	ROUTER RIP	Apply a weight to all routes or a specific route and ACL. Configure the following parameters: • *weight* range: 1 to 255 (default is 120) • *ip-address mask*: the IP address in dotted decimal format (A.B.C.D), and the mask in slash format (/x). • *access-list-name*: name of a configured IP ACL.
offset *access-list-name* {**in** | **out**} *offset* [*interface*]	ROUTER RIP	Apply an additional number to the incoming or outgoing route metrics. Configure the following parameters: • *access-list-name*: the name of a configured IP ACL • *offset* range: 0 to 16. • *interface*: the type, slot, and number of an interface.

Use the **show config** command in the ROUTER RIP mode to view configuration changes.

Debug RIP

The **debug ip rip** command enables RIP debugging. When debugging is enabled, you can view information on RIP protocol changes or RIP routes.

To enable RIP debugging, use the following command in the EXEC privilege mode:

Command Syntax	Command Mode	Purpose
debug ip rip [*interface* | **database** | **events** | **trigger**]	EXEC privilege	Enable debugging of RIP.

Figure 42-6 shows the confirmation when the debug function is enabled.

Figure 42-6. debug ip rip Command Example

```
FTOS#debug ip rip
RIP protocol debug is ON
FTOS#
```

To disable RIP, use the **no debug ip rip** command.

RIP Configuration Example

The example in this section shows the command sequence to configure RIPv2 on the two routers shown in Figure 42-7 — "Core 2" and "Core 3". The host prompts used in the example screenshots reflect those names. The screenshots are divided into the following groups of command sequences:

- Configuring RIPv2 on Core 2 on page 887
- Core 2 Output on page 887
- RIP Configuration on Core 3 on page 889
- Core 3 RIP Output on page 889
- RIP Configuration Summary on page 891

Figure 42-7. RIP Topology Example

Configuring RIPv2 on Core 2

Figure 42-8. Configuring RIPv2 on Core 2

```
Core2(conf-if-gi-2/31)#
Core2(conf-if-gi-2/31)#router rip
Core2(conf-router_rip)#ver 2
Core2(conf-router_rip)#network 10.200.10.0
Core2(conf-router_rip)#network 10.300.10.0
Core2(conf-router_rip)#network 10.11.10.0
Core2(conf-router_rip)#network 10.11.20.0
Core2(conf-router_rip)#show config
!
router rip
 network 10.0.0.0
 version 2
Core2(conf-router_rip)#
```

Core 2 Output

The screenshots in this section are:

- Figure 42-9: Using **show ip rip database** command to display Core 2 RIP database
- Figure 42-10: Using **show ip route** command to display Core 2 RIP setup
- Figure 42-11: Using **show ip protocols** command to display Core 2 RIP activity

Figure 42-9. Example of RIP Configuration Response from Core 2

```
Core2(conf-router_rip)#end
00:12:24: %RPM0-P:CP %SYS-5-CONFIG_I: Configured from console by  console
Core2#show ip rip database
Total number of routes in RIP database: 7
10.11.30.0/24
        [120/1] via 10.11.20.1, 00:00:03, GigabitEthernet 2/31
10.300.10.0/24          directly connected,GigabitEthernet 2/42
10.200.10.0/24          directly connected,GigabitEthernet 2/41
10.11.20.0/24           directly connected,GigabitEthernet 2/31
10.11.10.0/24           directly connected,GigabitEthernet 2/11
10.0.0.0/8              auto-summary
192.168.1.0/24
        [120/1] via 10.11.20.1, 00:00:03, GigabitEthernet 2/31
192.168.1.0/24          auto-summary
192.168.2.0/24
        [120/1] via 10.11.20.1, 00:00:03, GigabitEthernet 2/31
192.168.2.0/24          auto-summary

Core2#
```

Figure 42-10. Using show ip route Command to Show RIP Configuration on Core 2

```
Core2#show ip route

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set

Destination        Gateway                     Dist/Metric Last Change
       -----------        -------                       ----------- -----------
  C    10.11.10.0/24      Direct, Gi 2/11                 0/0    00:02:26
  C    10.11.20.0/24      Direct, Gi 2/31                 0/0    00:02:02
  R    10.11.30.0/24      via 10.11.20.1, Gi 2/31       120/1    00:01:20
  C    10.200.10.0/24     Direct, Gi 2/41                 0/0    00:03:03
  C    10.300.10.0/24     Direct, Gi 2/42                 0/0    00:02:42
  R    192.168.1.0/24     via 10.11.20.1, Gi 2/31       120/1    00:01:20
  R    192.168.2.0/24     via 10.11.20.1, Gi 2/31       120/1    00:01:20
Core2#
  R    192.168.1.0/24     via 10.11.20.1, Gi 2/31       120/1    00:05:22
  R    192.168.2.0/24     via 10.11.20.1, Gi 2/31       120/1    00:05:22

Core2#
```

Figure 42-11. Using show ip protocols Command to Show RIP Configuration Activity on Core 2

```
Core2#show ip protocols
Routing Protocol is "RIP"
 Sending updates every 30 seconds, next due in 17
 Invalid after 180 seconds, hold down 180, flushed after 240
 Output delay 8 milliseconds between packets
 Automatic network summarization is in effect
 Outgoing filter for all interfaces is
 Incoming filter for all interfaces is
 Default redistribution metric is  1
 Default version control: receive version 2, send version 2
        Interface       Recv  Send
        GigabitEthernet 2/42   2     2
        GigabitEthernet 2/41   2     2
        GigabitEthernet 2/31   2     2
        GigabitEthernet 2/11   2     2
 Routing for Networks:
         10.300.10.0
         10.200.10.0
         10.11.20.0
         10.11.10.0

 Routing Information Sources:
 Gateway           Distance      Last Update
 10.11.20.1          120              00:00:12

 Distance: (default is 120)

Core2#
```

RIP Configuration on Core 3

Figure 42-12. RIP Configuration on Core 3

```
Core3(conf-if-gi-3/21)#router rip
Core3(conf-router_rip)#version 2
Core3(conf-router_rip)#network 192.168.1.0
Core3(conf-router_rip)#network 192.168.2.0
Core3(conf-router_rip)#network 10.11.30.0
Core3(conf-router_rip)#network 10.11.20.0
Core3(conf-router_rip)#show config
!
router rip
 network 10.0.0.0
 network 192.168.1.0
 network 192.168.2.0
 version 2
Core3(conf-router_rip)#
```

Core 3 RIP Output

The screenshots in this section are:

- Figure 42-13: Using **show ip rip database** command to display Core 3 RIP database
- Figure 42-14: Using **show ip route** command to display Core 3 RIP setup
- Figure 42-15: Using **show ip protocols** command to display Core 3 RIP activity

Figure 42-13. Using show ip rip database Command for Core 3 RIP Setup

```
Core3#show ip rip database
Total number of routes in RIP database: 7
10.11.10.0/24
        [120/1] via 10.11.20.2, 00:00:13, GigabitEthernet 3/21
10.200.10.0/24
        [120/1] via 10.11.20.2, 00:00:13, GigabitEthernet 3/21
10.300.10.0/24
        [120/1] via 10.11.20.2, 00:00:13, GigabitEthernet 3/21
10.11.20.0/24           directly connected,GigabitEthernet 3/21
10.11.30.0/24           directly connected,GigabitEthernet 3/11
10.0.0.0/8              auto-summary
192.168.1.0/24          directly connected,GigabitEthernet 3/43
192.168.1.0/24          auto-summary
192.168.2.0/24          directly connected,GigabitEthernet 3/44
192.168.2.0/24          auto-summary
Core3#
```

Figure 42-14. Using show ip routes for Core 3 RIP Setup

```
Core3#show ip routes

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set

       Destination        Gateway                      Dist/Metric Last Change
       -----------        -------                      ----------- -----------
  R    10.11.10.0/24      via 10.11.20.2, Gi 3/21            120/1    00:01:14
  C    10.11.20.0/24      Direct, Gi 3/21                      0/0    00:01:53
  C    10.11.30.0/24      Direct, Gi 3/11                      0/0    00:06:00
  R    10.200.10.0/24     via 10.11.20.2, Gi 3/21            120/1    00:01:14
  R    10.300.10.0/24     via 10.11.20.2, Gi 3/21            120/1    00:01:14
  C    192.168.1.0/24     Direct, Gi 3/43                      0/0    00:06:53
  C    192.168.2.0/24     Direct, Gi 3/44                      0/0    00:06:26
Core3#
```

Figure 42-15. Using show ip protocols Command to Show RIP Configuration Activity on Core 3

```
Core3#show ip protocols

Routing Protocol is "RIP"
 Sending updates every 30 seconds, next due in 6
 Invalid after 180 seconds, hold down 180, flushed after 240
 Output delay 8 milliseconds between packets
 Automatic network summarization is in effect
 Outgoing filter for all interfaces is
 Incoming filter for all interfaces is
 Default redistribution metric is  1
 Default version control: receive version 2, send version 2
        Interface       Recv  Send
        GigabitEthernet 3/21   2     2
        GigabitEthernet 3/11   2     2
        GigabitEthernet 3/44   2     2
        GigabitEthernet 3/43   2     2
 Routing for Networks:
         10.11.20.0
         10.11.30.0
         192.168.2.0
         192.168.1.0

 Routing Information Sources:
 Gateway           Distance      Last Update
 10.11.20.2          120              00:00:22

 Distance: (default is 120)

Core3#
```

RIP Configuration Summary

Figure 42-16. Summary of Core 2 RIP Configuration Using Output of show run Command

```
!
interface GigabitEthernet 2/11
 ip address 10.11.10.1/24
 no shutdown
!
interface GigabitEthernet 2/31
 ip address 10.11.20.2/24
 no shutdown

!
interface GigabitEthernet 2/41
 ip address 10.200.10.1/24
 no shutdown

!
interface GigabitEthernet 2/42
 ip address 10.250.10.1/24
 no shutdown

router rip
version 2
10.200.10.0
10.300.10.0
10.11.10.0
10.11.20.0
```

Figure 42-17. Summary of Core 3 RIP Configuration Using Output of show run Command

```
!
interface GigabitEthernet 3/11
 ip address 10.11.30.1/24
 no shutdown

!
interface GigabitEthernet 3/21
 ip address 10.11.20.1/24
 no shutdown

!
interface GigabitEthernet 3/43
 ip address 192.168.1.1/24
 no shutdown

!
interface GigabitEthernet 3/44
 ip address 192.168.2.1/24
 no shutdown


!
router rip
version 2
network 10.11.20.0
network 10.11.30.0
network 192.168.1.0
network 192.168.2.0
```

43

Remote Monitoring

Remote Monitoring is supported on platform C E S

Remote Monitoring is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

This chapter describes the Remote Monitoring (RMON):

Remote Monitoring (RMON) is an industry-standard implementation that monitors network traffic by sharing network monitoring information. RMON provides both 32-bit and 64-bit monitoring facility and long-term statistics collection on Dell Force10 Ethernet Interfaces.

RMON operates with SNMP and monitors all nodes on a LAN segment. RMON monitors traffic passing through the router and segment traffic not destined for the router. The monitored interfaces may be chosen by using alarms and events with standard MIBs.

Implementation

You must configure SNMP prior to setting up RMON. For a complete SNMP implementation discussion, refer to Chapter 6, Simple Network Management Protocol (SNMP), on page 47.

Configuring RMON requires using the RMON CLI and includes the following tasks:

- Set rmon alarm
- Configure an RMON event
- Configure RMON collection statistics
- Configure RMON collection history
- Enable an RMON MIB collection history group

RMON implements the following standard RFCs (for details see Appendix 63, Standards Compliance):

- RFC-2819
- RFC-3273
- RFC-3434

Fault Recovery

RMON provides the following fault recovery functions:

Interface Down—When an RMON-enabled interface goes down, monitoring continues. However, all data values are registered as 0xFFFFFFFF (32 bits) or ixFFFFFFFFFFFFFFFF (64 bits). When the interface comes back up, RMON monitoring processes resumes.

Note: A Network Management System (NMS) should be ready to interpret a down interface and plot the interface performance graph accordingly.

Line Card Down—The same as Interface Down (see above).

RPM Down, RPM Failover—Master and standby RPMs run the RMON sampling process in the background. Therefore, when an RPM goes down, the other RPM maintains the sampled data—the new master RPM provides the same sampled data as did the old master—as long as the master RPM had been running long enough to sample all the data.

NMS backs up all the long-term data collection, and displays the failover downtime from the performance graph.

Chassis Down—When a chassis goes down, all sampled data is lost. But the RMON configurations are saved in the configuration file, and the sampling process continues after the chassis returns to operation.

Platform Adaptation—RMON supports all Dell Force10 chassis and all Dell Force10 Ethernet Interfaces.

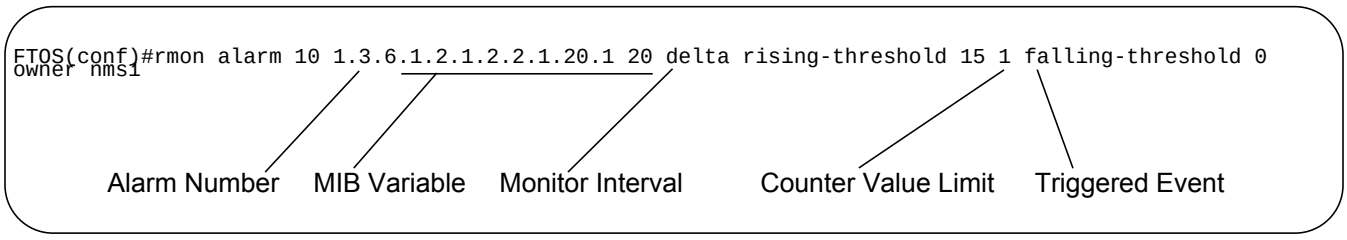
Set rmon alarm

To set an alarm on any MIB object, use the **rmon alarm** or **rmon hc-alarm** command in GLOBAL CONFIGURATION mode. To disable the alarm, use the **no** form of this command:

Command Syntax	Command Mode	Purpose
[no] rmon alarm number variable interval {delta | absolute} rising-threshold [value event-number] falling-threshold value event-number [owner string] or **[no] rmon hc-alarm number variable interval {delta | absolute} rising-threshold value event-number falling-threshold value event-number [owner string]**	CONFIGURATION	Set an alarm on any MIB object. Use the **no** form of this command to disable the alarm. Configure the alarm using the following optional parameters: • *number*: Alarm number, should be an integer from 1 to 65,535, the value must be unique in the RMON Alarm Table • *variable*: The MIB object to monitor—the variable must be in the SNMP OID format. For example, 1.3.6.1.2.1.1.3. The object type must be a 32-bit integer for the **rmon alarm** command and 64 bits for the **rmon hc-alarm** command. • *interval*: Time in seconds the alarm monitors the MIB variable, the value must be between 1 to 3,600. • **delta**: Tests the change between MIB variables, this is the *alarmSampleType* in the RMON Alarm table. • **absolute**: Tests each MIB variable directly, this is the *alarmSampleType* in the RMON Alarm table. • **rising-threshold** *value*: Value at which the rising-threshold alarm is triggered or reset. For the **rmon alarm** command this is a 32-bits value, for **rmon hc-alarm** command this is a 64-bits value. • *event-number*: Event number to trigger when the rising threshold exceeds its limit. This value is identical to the *alarmRisingEventIndex* in the alarmTable of the RMON MIB. If there is no corresponding rising-threshold event, the value should be zero. • **falling-threshold** *value*: Value at which the falling-threshold alarm is triggered or reset. For the **rmon alarm** command, this is a 32-bits value, for **rmon hc-alarm** command this is a 64bits value. • *event-number*: Event number to trigger when the falling threshold exceeds its limit. This value is identical to the *alarmFallingEventIndex* in the alarmTable of the RMON MIB. If there is no corresponding falling-threshold event, the value should be zero. • **owner** *string*: (Optional) Specifies an owner for the alarm, this is the alarmOwner object in the alarmTable of the RMON MIB. Default is a null-terminated string.

The following example configures an RMON alarm using the **rmon alarm** command.

Figure 43-1. rmon alarm Command Example



The above example configures RMON alarm number 10. The alarm monitors the MIB variable 1.3.6.1.2.1.2.2.1.20.1 (ifEntry.ifOutErrors) once every 20 seconds until the alarm is disabled, and checks the rise or fall of the variable. The alarm is triggered when the 1.3.6.1.2.1.2.2.1.20.1 value shows a MIB counter increase of 15 or more (such as from 100000 to 100015). The alarm then triggers event number 1, which is configured with the RMON event command. Possible events include a log entry or a SNMP trap. If the 1.3.6.1.2.1.2.2.1.20.1 value changes to 0 (falling-threshold 0), the alarm is reset and can be triggered again.

Configure an RMON event

To add an event in the RMON event table, use the **rmon event** command in GLOBAL CONFIGURATION mode. To disable RMON on the interface, use the **no** form of this command:

Command Syntax	Command Mode	Purpose
[no] rmon event number [log] [trap community] [description string] [owner string]	CONFIGURATION	*number*: Assigned event number, which is identical to the eventIndex in the eventTable in the RMON MIB. The value must be an integer from 1 to 65,535, the value must be unique in the RMON Event Table. *log*: (Optional) Generates an RMON log entry when the event is triggered and sets the eventType in the RMON MIB to log or log-and-trap. Default is no log. **trap** *community*: (Optional) SNMP community string used for this trap. Configures the setting of the eventType in the RMON MIB for this row as either snmp-trap or log-and-trap. This value is identical to the eventCommunityValue in the eventTable in the RMON MIB. Default is “public”. **description** *string*: (Optional) Specifies a description of the event, which is identical to the event description in the eventTable of the RMON MIB. Default is a null-terminated string. **owner** *string*: (Optional) Owner of this event, which is identical to the eventOwner in the eventTable of the RMON MIB. Default is a null-terminated string.

The following example shows the **rmon event** command.

Figure 43-2. rmon event Command Example

```
FTOS(conf)#rmon event 1 log trap eventtrap description “High ifOutErrors” owner nms1
```

The above configuration example creates RMON event number 1, with the description “High ifOutErrors”, and generates a log entry when the event is triggered by an alarm. The user *nms1* owns the row that is created in the event table by this command. This configuration also generates an SNMP trap when the event is triggered using the SNMP community string “eventtrap”.

Configure RMON collection statistics

To enable RMON MIB statistics collection on an interface, use the RMON collection statistics command in interface configuration mode. To remove a specified RMON statistics collection, use the **no** form of this command.

Command Syntax	Command Mode	Purpose
[**no**] **rmon collection statistics** {**controlEntry** *integer*} [**owner** *ownername*]	CONFIGURATION INTERFACE (config-if)	**controlEntry**: Specifies the RMON group of statistics using a value. *integer*: A value from 1 to 65,535 that identifies the RMON Statistics Table. The value must be unique in the RMON Statistic Table. **owner**: (Optional) Specifies the name of the owner of the RMON group of statistics. *ownername*: (Optional) Records the name of the owner of the RMON group of statistics. Default is a null-terminated string

The following command enables the RMON statistics collection on the interface, with an ID value of 20 and an owner of “john.”

Figure 43-3. rmon collection statistics Command Example

```
FTOS(conf-if-mgmt)#rmon collection statistics controlEntry 20 owner john
```

Configure RMON collection history

To enable the RMON MIB history group of statistics collection on an interface, use the **rmon collection history** command in interface configuration mode. To remove a specified RMON history group of statistics collection, use the **no** form of this command.

Command Syntax	Command Mode	Purpose
[no] rmon collection history {controlEntry integer} [owner ownername] [buckets bucket-number] [interval seconds]	CONFIGURATION INTERFACE (config-if)	**controlEntry**: Specifies the RMON group of statistics using a value. *integer*: A value from 1 to 65,535 that identifies the RMON group of statistics. The value must be a unique index in the RMON History Table. **owner**: (Optional) Specifies the name of the owner of the RMON group of statistics.Default is a null-terminated string. *ownername*: (Optional) Records the name of the owner of the RMON group of statistics. **buckets**: (Optional) Specifies the maximum number of buckets desired for the RMON collection history group of statistics. *bucket-number*: (Optional) A value associated with the number of buckets specified for the RMON collection history group of statistics. The value is limited to from 1 to 1000. Default is 50 (as defined in RFC-2819). **interval**: (Optional) Specifies the number of seconds in each polling cycle. *seconds*: (Optional) The number of seconds in each polling cycle. The value is ranged from 5 to 3,600 (Seconds). Default is 1,800 as defined in RFC-2819.

Enable an RMON MIB collection history group

The following command enables an RMON MIB collection history group of statistics with an ID number of 20 and an owner of “john”, both the sampling interval and the number of buckets use their respective defaults.

Figure 43-4. rmon collection history Command Example

```
FTOS(conf-if-mgmt)#rmon collection history controlEntry 20 owner john
```

44

Rapid Spanning Tree Protocol

Rapid Spanning Tree Protocol is supported on platforms: C E S

RSTP is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

Protocol Overview

Rapid Spanning Tree Protocol (RSTP) is a Layer 2 protocol—specified by IEEE 802.1w—that is essentially the same as Spanning-Tree Protocol (STP) but provides faster convergence and interoperability with switches configured with STP and MSTP.

FTOS supports three other variations of Spanning Tree, as shown in Table 44-1.

Table 44-1. FTOS Supported Spanning Tree Protocols

Force 10 Term	IEEE Specification
Spanning Tree Protocol	802.1d
Rapid Spanning Tree Protocol	802.1w
Multiple Spanning Tree Protocol	802.1s
Per-VLAN Spanning Tree Plus	Third Party

Configuring Rapid Spanning Tree

Configuring Rapid Spanning Tree is a two-step process:

1. Configure interfaces for Layer 2. See page 48.
2. Enable Rapid Spanning Tree Protocol. See page 49.

Related Configuration Tasks

- Add and Remove Interfaces on page 904
- Modify Global Parameters on page 904

- Modify Interface Parameters on page 906
- Configure an EdgePort on page 906
- Preventing Network Disruptions with BPDU Guard on page 1057
- Influence RSTP Root Selection on page 908
- Configuring Spanning Trees as Hitless on page 1064
- Fast Hellos for Link State Detection on page 909
- Flush MAC Addresses after a Topology Change on page 654

Important Points to Remember

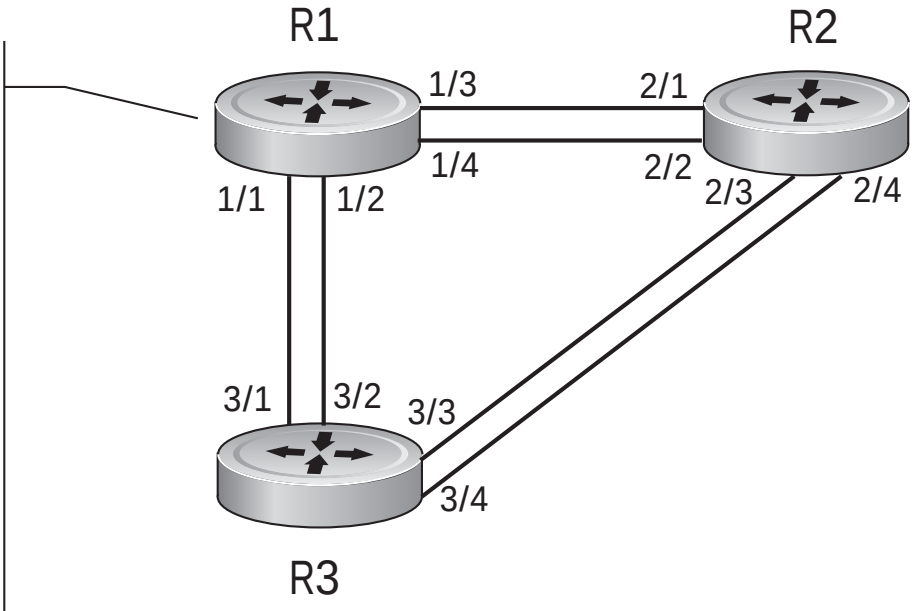
- RSTP is disabled by default.
- FTOS supports only one Rapid Spanning Tree (RST) instance.
- All interfaces in VLANs and all enabled interfaces in Layer 2 mode are automatically added to the RST topology.
- Avoid using the **range** command to add a large group of ports to a large group of VLANs; adding a group of ports to a range of VLANs sends multiple messages to the RSTP task. When using the **range** command, Dell Force10 recommends limiting the range to 5 ports and 40 VLANs.

Configure Interfaces for Layer 2 Mode

All interfaces on all bridges that will participate in Rapid Spanning Tree must be in Layer 2 and enabled.

Figure 44-1. Configuring Interfaces for Layer 2 Mode

```
R1(conf)# int range gi 1/1 - 4
R1(conf-if-gi-1/1-4)# switchport
R1(conf-if-gi-1/1-4)# no shutdown
R1(conf-if-gi-1/1-4)#show config
!
interface GigabitEthernet 1/1
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/2
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/3
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/4
 no ip address
 switchport
 no shutdown
```

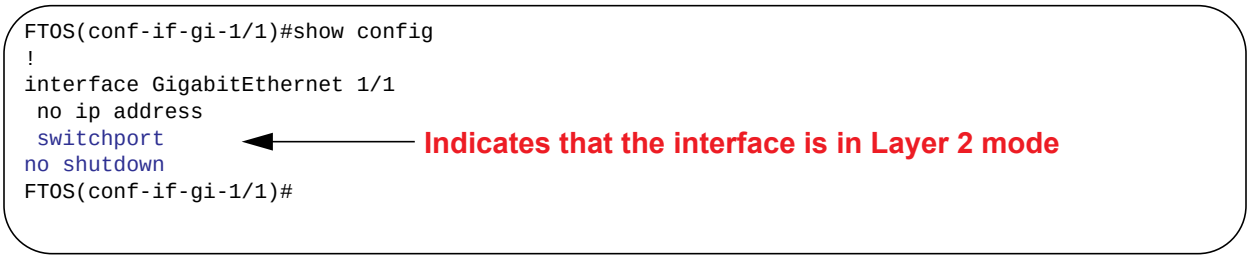


To configure the interfaces for Layer 2 and then enable them:

Step	Task	Command Syntax	Command Mode
1	If the interface has been assigned an IP address, remove it.	**no ip address**	INTERFACE
2	Place the interface in Layer 2 mode.	**switchport**	INTERFACE
3	Enable the interface.	**no shutdown**	INTERFACE

Verify that an interface is in Layer 2 mode and enabled using the **show config** command from INTERFACE mode.

Figure 44-2. Verifying Layer 2 Configuration



Enable Rapid Spanning Tree Protocol Globally

Rapid Spanning Tree Protocol must be enabled globally on all participating bridges; it is not enabled by default.

To enable Rapid Spanning Tree globally for all Layer 2 interfaces:

Step	Task	Command Syntax	Command Mode
1	Enter the PROTOCOL SPANNING TREE RSTP mode.	**protocol spanning-tree rstp**	CONFIGURATION
2	Enable Rapid Spanning Tree.	**no disable**	PROTOCOL SPANNING TREE RSTP

Note: To disable RSTP globally for all Layer 2 interfaces, enter the **disable** command from PROTOCOL SPANNING TREE RSTP mode.

Verify that Rapid Spanning Tree is enabled using the **show config** command from PROTOCOL SPANNING TREE RSTP mode.

Figure 44-3. Verifying RSTP is Enabled

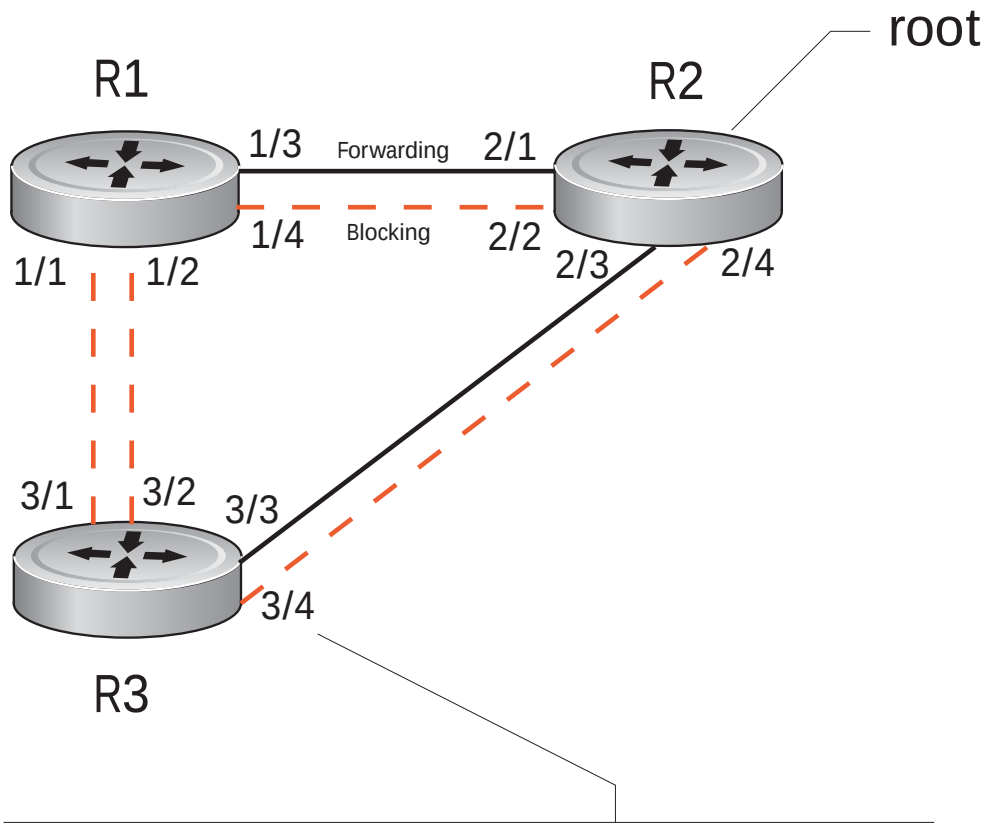
```
FTOS(conf-rstp)#show config
!
protocol spanning-tree rstp
 no disable
FTOS(conf-rstp)#
```

Indicates that Rapid Spanning Tree is enabled

When you enable Rapid Spanning Tree, all physical and port-channel interfaces that are enabled and in Layer 2 mode are automatically part of the RST topology.

- Only one path from any bridge to any other bridge is enabled.
- Bridges block a redundant path by disabling one of the link ports.

Figure 44-4. Rapid Spanning Tree Enabled Globally




```
Port 684 (GigabitEthernet 4/43) is alternate Discarding
Port path cost 20000, Port priority 128, Port Identifier 128.684
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.684, designated path cost 20000
Number of transitions to forwarding state 0
BPDU : sent 3, received 219
The port is not in the Edge port mode
```

View the interfaces participating in Rapid Spanning Tree using the **show spanning-tree rstp** command from EXEC privilege mode. If a physical interface is part of a port channel, only the port channel is listed in the command output.

Figure 44-5. show spanning-tree rstp Command Example

```
FTOS#show spanning-tree rstp
Root Identifier has priority 32768, Address 0001.e801.cbb4
Root Bridge hello time 2, max age 20, forward delay 15, max hops 0
Bridge Identifier has priority 32768, Address 0001.e801.cbb4
Configured hello time 2, max age 20, forward delay 15, max hops 0
We are the root
Current root has priority 32768, Address 0001.e801.cbb4
Number of topology changes 4, last change occurred 00:02:17 ago on Gi 1/26

Port 377 (GigabitEthernet 2/1) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.377
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.377, designated path cost 0
Number of transitions to forwarding state 1
BPDU : sent 121, received 9
The port is not in the Edge port mode

Port 378 (GigabitEthernet 2/2) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.378
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.378, designated path cost 0
Number of transitions to forwarding state 1
BPDU : sent 121, received 2
The port is not in the Edge port mode

Port 379 (GigabitEthernet 2/3) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.379
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.379, designated path cost 0
Number of transitions to forwarding state 1
BPDU : sent 121, received 5
The port is not in the Edge port mode

Port 380 (GigabitEthernet 2/4) is designated Forwarding
Port path cost 20000, Port priority 128, Port Identifier 128.380
Designated root has priority 32768, address 0001.e801.cbb4
Designated bridge has priority 32768, address 0001.e801.cbb4
Designated port id is 128.380, designated path cost 0


Number of transitions to forwarding state 1
BPDU : sent 147, received 3
The port is not in the Edge port mode
```

Confirm that a port is participating in Rapid Spanning Tree using the **show spanning-tree rstp brief** command from EXEC privilege mode.

Figure 44-6. show spanning-tree rstp brief Command Example

```
R3#show spanning-tree rstp brief
Executing IEEE compatible Spanning Tree Protocol
Root ID    Priority 32768, Address 0001.e801.cbb4
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID    Priority 32768, Address 0001.e80f.1dad
Configured hello time 2, max age 20, forward delay 15
Interface                                        Designated
 Name      PortID   Prio Cost    Sts Cost        Bridge ID        PortID
---------- -------- ---- ------- --- ------- -------------------- --------
Gi 3/1     128.681  128  20000   BLK 20000   32768 0001.e80b.88bd 128.469
Gi 3/2     128.682  128  20000   BLK 20000   32768 0001.e80b.88bd 128.470
Gi 3/3     128.683  128  20000   FWD 20000   32768 0001.e801.cbb4 128.379
Gi 3/4     128.684  128  20000   BLK 20000   32768 0001.e801.cbb4 128.380
Interface
 Name      Role   PortID   Prio Cost    Sts Cost    Link-type Edge
---------- ------ -------- ---- ------- --- ------- --------- ----
Gi 3/1     Altr   128.681  128  20000   BLK 20000   P2P       No
Gi 3/2     Altr   128.682  128  20000   BLK 20000   P2P       No
Gi 3/3     Root   128.683  128  20000   FWD 20000   P2P       No
Gi 3/4     Altr   128.684  128  20000   BLK 20000   P2P       No
R3#
```

Add and Remove Interfaces

- To add an interface to the Rapid Spanning Tree topology, configure it for Layer 2 and it is automatically added. If you previously disabled RSTP on the interface using the command **no spanning-tree 0**, re-enable it using the command **spanning-tree 0**.
- Remove an interface from the Rapid Spanning Tree topology using the command **no spanning-tree 0**. See also Removing an Interface from the Spanning Tree Group on page 1054 for BPDU Filtering behavior.

Modify Global Parameters

You can modify Rapid Spanning Tree parameters. The root bridge sets the values for forward-delay, hello-time, and max-age and overwrites the values set on other bridges participating in the Rapid Spanning Tree group.

- **Forward-delay** is the amount of time an interface waits in the Listening State and the Learning State before it transitions to the Forwarding State.
- **Hello-time** is the time interval in which the bridge sends RSTP Bridge Protocol Data Units (BPDUs).
- **Max-age** is the length of time the bridge maintains configuration information before it refreshes that information by recomputing the RST topology.

Note: Dell Force10 recommends that only experienced network administrators change the Rapid Spanning Tree group parameters. Poorly planned modification of the RSTG parameters can negatively impact network performance.

Table 44-2 displays the default values for RSTP.

Table 44-2. RSTP Default Values

RSTP Parameter		Default Value
Forward Delay		15 seconds
Hello Time		2 seconds
Max Age		20 seconds
Port Cost	100-Mb/s Ethernet interfaces	200000
	1-Gigabit Ethernet interfaces	20000
	10-Gigabit Ethernet interfaces	2000
	Port Channel with 100 Mb/s Ethernet interfaces	180000
	Port Channel with 1-Gigabit Ethernet interfaces	18000
	Port Channel with 10-Gigabit Ethernet interfaces	1800
Port Priority		128

To change these parameters, use the following commands, on the root bridge:

Task	Command Syntax	Command Mode
Change the forward-delay parameter. • Range: 4 to 30 • Default: 15 seconds	**forward-delay** *seconds*	PROTOCOL SPANNING TREE RSTP
Change the hello-time parameter. **Note:** With large configurations (especially those with more ports) Dell Force10 recommends that you increase the hello-time. Range: 1 to 10 Default: 2 seconds	**hello-time** *seconds*	PROTOCOL SPANNING TREE RSTP
Change the max-age parameter. Range: 6 to 40 Default: 20 seconds	**max-age** *seconds*	PROTOCOL SPANNING TREE RSTP

View the current values for global parameters using the **show spanning-tree rstp** command from EXEC privilege mode. See Figure 44-5.

Modify Interface Parameters

On interfaces in Layer 2 mode, you can set the port cost and port priority values.

- **Port cost** is a value that is based on the interface type. The default values are listed in Table 44-2. The greater the port cost, the less likely the port will be selected to be a forwarding port.
- **Port priority** influences the likelihood that a port will be selected to be a forwarding port in case that several ports have the same port cost.

To change the port cost or priority of an interface, use the following commands:

Task	Command Syntax	Command Mode
Change the port cost of an interface. Range: 0 to 65535 Default: see Table 44-2.	**spanning-tree rstp cost** *cost*	INTERFACE
Change the port priority of an interface. Range: 0 to 15 Default: 128	**spanning-tree rstp priority** *priority-value*	INTERFACE

View the current values for interface parameters using the **show spanning-tree rstp** command from EXEC privilege mode. See Figure 44-5.

Configure an EdgePort

The EdgePort feature enables interfaces to begin forwarding traffic approximately 30 seconds sooner. In this mode an interface forwards frames by default until it receives a BPDU that indicates that it should behave otherwise; it does not go through the Learning and Listening states. The **bpduguard shutdown-on-violation** option causes the interface hardware to be shutdown when it receives a BPDU. When only **bpduguard** is implemented, although the interface is placed in an Error Disabled state when receiving the BPDU, the physical interface remains up and spanning-tree will drop packets in the hardware after a BPDU violation. BPDUs are dropped in the software after receiving the BPDU violation. This feature is the same as PortFast mode in Spanning Tree.

△ **Caution:** Configure EdgePort only on links connecting to an end station. EdgePort can cause loops if it is enabled on an interface connected to a network.

To enable EdgePort on an interface, use the following command:

Task	Command Syntax	Command Mode
Enable EdgePort on an interface.	**spanning-tree rstp edge-port** [**bpduguard | shutdown-on-violation**]	INTERFACE

Verify that EdgePort is enabled on a port using the **show spanning-tree rstp** command from the EXEC privilege mode or the **show config** command from INTERFACE mode; Dell Force10 recommends using the **show config** command, as shown in Figure 44-7.

FTOS Behavior: Regarding **bpduguard shutdown-on-violation** behavior:

1. If the interface to be shutdown is a port channel then all the member ports are disabled in the hardware.
2. When a physical port is added to a port channel already in error disable state, the new member port will also be disabled in the hardware.
3. When a physical port is removed from a port channel in error disable state, the error disabled state is cleared on this physical port (the physical port will be enabled in the hardware).
4. The **reset linecard** command does not clear the error disabled state of the port or the hardware disabled state. The interface continues to be disables in the hardware.

   The error disabled state can be cleared with any of the following methods:

   - Perform an **shutdown** command on the interface.
   - Disable the **shutdown-on-violation** command on the interface (**no spanning-tree** *stp-id* **portfast** [**bpduguard |** [**shutdown-on-violation**]]).
   - Disable spanning tree on the interface (**no spanning-tree** in INTERFACE mode).
   - Disabling global spanning tree (**no spanning-tree** in CONFIGURATION mode).

Figure 44-7. EdgePort Enabled on Interface

```
FTOS(conf-if-gi-2/0)#show config
!
interface GigabitEthernet 2/0
 no ip address
 switchport
 spanning-tree rstp edge-port
 shutdown
FTOS(conf-if-gi-2/0)#
```

Indicates the interface is in EdgePort mode

Influence RSTP Root Selection

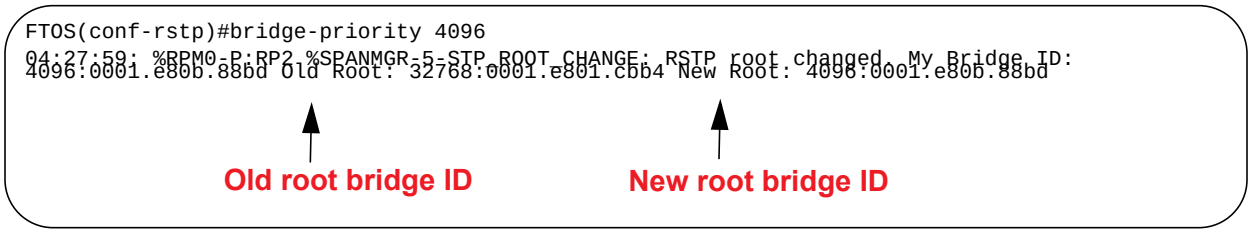
The Rapid Spanning Tree Protocol determines the root bridge, but you can assign one bridge a lower priority to increase the likelihood that it will be selected as the root bridge.

To change the bridge priority, use the following command:

Task	Command Syntax	Command Mode
Assign a number as the bridge priority or designate it as the primary or secondary root. *priority-value* range: 0 to 65535. The lower the number assigned, the more likely this bridge will become the root bridge. The default is 32768. Entries must be multiples of 4096.	**bridge-priority** *priority-value*	PROTOCOL SPANNING TREE RSTP

A console message appears when a new root bridge has been assigned. Figure 44-8 shows the console message after the **bridge-priority** command is used to make R2 the root bridge.

Figure 44-8. bridge-priority Command Example



SNMP Traps for Root Elections and Topology Changes

Enable SNMP traps for RSTP, MSTP, and PVST+ collectively using the command **snmp-server enable traps xstp**.

Fast Hellos for Link State Detection

Fast Hellos for Link State Detection is available only on platform: S

Use RSTP Fast Hellos to achieve sub-second link-down detection so that convergence is triggered faster. The standard RSTP link-state detection mechanism does not offer the same low link-state detection speed.

RSTP Fast Hellos decrease the hello interval to the order of milliseconds and all timers derived from the hello timer are adjusted accordingly. This feature does not inter-operate with other vendors, and is available only for RSTP.

Task	Command Syntax	Command Mode
Configure a hello time on the order of milliseconds.	**hello-time milli-second** *interval* Range: 50 - 950 milliseconds	PROTOCOL RSTP

```
FTOS(conf-rstp)#do show spanning-tree rstp brief
Executing IEEE compatible Spanning Tree Protocol
    Root ID    Priority 0, Address 0001.e811.2233
    Root Bridge hello time 50 ms, max age 20, forward delay 15
    Bridge ID    Priority 0, Address 0001.e811.2233
    We are the root
    Configured hello time 50 ms, max age 20, forward delay 15
```

Note: The hello time is encoded in BPDUs in increments of 1/256ths of a second. The standard minimum hello time in seconds is 1 second, which is encoded as 256. Millisecond hello times are encoded using values less than 256; the millisecond hello time equals (x/1000)*256.
Note: When millisecond hellos are configured, the default hello interval of 2 seconds is still used for edge ports; the millisecond hello interval is not used.

Configure a Root Guard

Use the Root Guard feature in a Layer 2 RSTP network to avoid bridging loops.

You enable root guard on a per-port or per-port-channel basis.

FTOS Behavior: The following conditions apply to a port enabled with root guard:

- Root guard is supported on any RSTP-enabled port or port-channel interface except when used as a stacking port.
- Root guard is supported on a port in any Spanning Tree mode:
  - Spanning Tree Protocol (STP)
  - Rapid Spanning Tree Protocol (RSTP)
  - Multiple Spanning Tree Protocol (MSTP)
  - Per-VLAN Spanning Tree Plus (PVST+)
- When enabled on a port, root guard applies to all VLANs configured on the port.
- Root guard and loop guard cannot be enabled at the same time on an RSTP port. For example, if you configure loop guard on a port on which root guard is already configured, the following error message is displayed:
  ```
  % Error: RootGuard is configured. Cannot configure LoopGuard.
  ```

To enable a root guard on an RSTP-enabled port or port-channel interface, enter the **spanning-tree rstp rootguard** command. Refer to STP Root Guard on page 1060 for more information on how to use the root guard feature.

Task	Command Syntax	Command Mode
Enable root guard on a port or port-channel interface.	**spanning-tree rstp rootguard**	INTERFACE INTERFACE PORT-CHANNEL

To disable RSTP root guard on a port or port-channel interface, enter the **no spanning-tree rstp rootguard** command in an interface configuration mode.

To verify the RSTP root guard configuration on a port or port-channel interface, enter the **show spanning-tree rstp guard [interface** *interface*] command in global configuration mode.

Configure a Loop Guard

The Loop Guard feature provides protection against Layer 2 forwarding loops (STP loops) caused by a hardware failure, such as a cable failure or an interface fault. When a cable or interface fails, a participating STP link may become unidirectional (STP requires links to be bidirectional) and an STP port does not receive BPDUs. When an STP blocking port does not receive BPDUs, it transitions to a forwarding state. This condition can create a loop in the network.

You enable loop guard on a per-port or per-port channel basis.

FTOS Behavior: The following conditions apply to a port enabled with loop guard:

- Loop guard is supported on any RSTP-enabled port or port-channel interface.
- Loop guard is supported on a port or port-channel in any Spanning Tree mode:
  - Spanning Tree Protocol (STP)
  - Rapid Spanning Tree Protocol (RSTP)
  - Multiple Spanning Tree Protocol (MSTP)
  - Per-VLAN Spanning Tree Plus (PVST+)
- Root guard and loop guard cannot be enabled at the same time on an RSTP port. For example, if you configure loop guard on a port on which root guard is already configured, the following error message is displayed:
  ```
  % Error: RootGuard is configured. Cannot configure LoopGuard.
  ```
- Enabling Portfast BPDU guard and loop guard at the same time on a port results in a port that remains in a blocking state and prevents traffic from flowing through it. For example, when Portfast BPDU guard and loop guard are both configured:

  - If a BPDU is received from a remote device, BPDU guard places the port in an err-disabled blocking state and no traffic is forwarded on the port.

  - If no BPDU is received from a remote device, loop guard places the port in a loop-inconsistent blocking state and no traffic is forwarded on the port.

To enable a loop guard on an RSTP-enabled port or port-channel interface, enter the **spanning-tree rstp loopguard** command. Refer to STP Loop Guard on page 1064 for more information on how to use the loop guard feature.

Task	Command Syntax	Command Mode
Enable loop guard on an RSTP-enabled port or port-channel interface.	**spanning-tree rstp loopguard**	INTERFACE INTERFACE PORT-CHANNEL

To disable RSTP loop guard on a port or port-channel interface, enter the **no spanning-tree rstp loopguard** command in an INTERFACE configuration mode.

To verify the RSTP loop guard configuration on a port or port-channel interface, enter the **show spanning-tree rstp guard** [**interface** *interface*] command in global configuration mode.

Displaying STP Guard Configuration

To verify the STP guard configured on RSTP port or port-channel interfaces, enter the **show spanning-tree rstp guard** command. Refer to Chapter 52, “Spanning Tree Protocol,” on page 1049 for information on how to configure and use the STP root guard, loop guard, and BPDU guard features.

Figure 44-9 shows an example for an RSTP network (instance 0) in which:

- Root guard is enabled on a port that is in a root-inconsistent state.
- Loop guard is enabled on a port that is in a forwarding state.
- BPDU guard is enabled on a port that is shut down.

Figure 44-9. Displaying STP Guard Configuration

```
FTOS#show spanning-tree rstp guard
Interface
Name      Instance    Sts          Guard type
---------  --------   ---------     ----------
Gi 0/1     0          INCON(Root)  Rootguard
Gi 0/2     0          FWD          Loopguard
Gi 0/3     0          EDS (Shut)   Bpduguard
```

45

Security

Security features are supported on platforms C E S

This chapter discusses several ways to provide access security to the Dell Force10 system. Platform-specific features are identified by the C, E or S icons (as shown below).

Security features are supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

For details on all commands discussed in this chapter, see the Security Commands chapter in the *FTOS Command Reference*.

AAA Accounting

AAA Accounting is part of the AAA security model (Accounting, Authentication, and Authorization), which includes services for authentication, authorization, and accounting. For details on commands related to AAA security, refer to the Security chapter in the *FTOS Command Reference*.

AAA Accounting enables tracking of services that users are accessing and the amount of network resources being consumed by those services. When AAA Accounting is enabled, the network server reports user activity to the security server in the form of accounting records. Each accounting record is comprised of accounting AV pairs and is stored on the access control server.

As with authentication and authorization, you must configure AAA Accounting by defining a named list of accounting methods, and then apply that list to various interfaces.

Configuration Task List for AAA Accounting

The following sections present the AAA Accounting configuration tasks:

- Enable AAA Accounting on page 914 (mandatory)
- Suppress AAA Accounting for null username sessions on page 915 (optional)
- Configure Accounting of EXEC and privilege-level command usage on page 915 (optional)
- Configure AAA Accounting for terminal lines on page 915 (optional)
- Monitor AAA Accounting on page 915 (optional)

Enable AAA Accounting

The **aaa accounting** command enables you to create a record for any or all of the accounting functions monitored. To enable AAA accounting, perform the following task in CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
aaa accounting {**system** | **exec** | **command** *level*} {*default* | *name*} {**start-stop** | **wait-start** | **stop-only**} {**tacacs+**}	CONFIGURATION	Enable AAA Accounting and create a record for monitoring the accounting function. The variables are: • **system**—sends accounting information of any other AAA configuration • **exec**—sends accounting information when a user has logged in to the EXEC mode • **command** *level*—sends accounting of commands executed at the specified privilege level • *default* | *name*—Enter the name of a list of accounting methods. • **start-stop—**Use for more accounting information, to send a start-accounting notice at the beginning of the requested event and a stop-accounting notice at the end. • **wait-start—**ensures that the TACACS+ security server acknowledges the start notice before granting the user's process request • **stop-only**—Use for minimal accounting; instructs the TACACS+ server to send a stop record accounting notice at the end of the requested user process. • **tacacs+** —Designate the security service. Currently, FTOS supports only TACACS+

Suppress AAA Accounting for null username sessions

When AAA Accounting is activated, the FTOS software issues accounting records for all users on the system, including users whose username string, because of protocol translation, is NULL. An example of this is a user who comes in on a line where the AAA Authentication **login method-list none** command is applied. To prevent accounting records from being generated for sessions that do not have usernames associated with them, perform the following task in CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
aaa accounting suppress null-username	CONFIGURATION	Prevent accounting records from being generated for users whose username string is NULL

Configure Accounting of EXEC and privilege-level command usage

The network access server monitors the accounting functions defined in the TACACS+ attribute/value (AV) pairs.

In the following sample configuration, AAA accounting is set to track all usage of EXEC commands and commands on privilege level 15.

```
FTOS(conf)#aaa accounting exec default start-stop tacacs+
FTOS(conf)#aaa accounting command 15 default start-stop tacacs+
```

System accounting can use only the default method list:

aaa accounting system default start-stop tacacs+

Configure AAA Accounting for terminal lines

Use the following commands to enable accounting with a named method list for a specific terminal line (where com15 and execAcct are the method list names):

```
FTOS(config-line-vty)# accounting commands 15 com15
FTOS(config-line-vty)# accounting exec execAcct
```

Monitor AAA Accounting

FTOS does not support periodic interim accounting, because the **periodic** command can cause heavy congestion when many users are logged in to the network.

No specific **show** command exists for TACACS+ accounting. To obtain accounting records displaying information about users currently logged in, perform the following task in Privileged EXEC mode:

Command Syntax	Command Mode	Purpose
show accounting	CONFIGURATION	Step through all active sessions and print all the accounting records for the actively accounted functions.

Figure 45-1. show accounting Command Example for AAA Accounting

```
FTOS#show accounting
Active accounted actions on tty2, User admin Priv 1
   Task ID 1, EXEC Accounting record, 00:00:39 Elapsed, service=shell
Active accounted actions on tty3, User admin Priv 1
   Task ID 2, EXEC Accounting record, 00:00:26 Elapsed, service=shell
```

AAA Authentication

FTOS supports a distributed client/server system implemented through Authentication, Authorization, and Accounting (AAA) to help secure networks against unauthorized access. In the Dell Force10 implementation, the Dell Force10 system acts as a RADIUS or TACACS+ client and sends authentication requests to a central RADIUS or TACACS+ server that contains all user authentication and network service access information.

Dell Force10 uses local usernames/passwords (stored on the Dell Force10 system) or AAA for login authentication. With AAA, you can specify the security protocol or mechanism for different login methods and different users. In FTOS, AAA uses a list of authentication methods, called method lists, to define the types of authentication and the sequence in which they are applied. You can define a method list or use the default method list. User-defined method lists take precedence over the default method list.

Configuration Task List for AAA Authentication

The following sections provide the configuration tasks:

- Configure login authentication for terminal lines
- Configure AAA Authentication login methods on page 918
- Enable AAA Authentication on page 919
- AAA Authentication—RADIUS on page 919

For a complete listing of all commands related to login authentication, refer to the Security chapter in the *FTOS Command Reference*.

Configure login authentication for terminal lines

You can assign up to five authentication methods to a method list. FTOS evaluates the methods in the order in which you enter them in each list. If the first method list does not respond or returns an error, FTOS applies the next method list until the user either passes or fails the authentication. If the user fails a method list, FTOS does not apply the next method list.

Configure AAA Authentication login methods

To configure an authentication method and method list, use these commands in the following sequence in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**aaa authentication login** {*method-list-name* | **default**} *method1* [... *method4*]	CONFIGURATION	Define an authentication method-list (*method-list-name*) or specify the **default**. The **default** method-list is applied to all terminal lines. Possible methods are: • **enable**—use the password defined by the **enable secret** or **enable password** command in the CONFIGURATION mode. • **line**—use the password defined by the password command in the LINE mode. • **local**—use the username/password database defined in the local configuration. • **none**—no authentication. • **radius**—use the RADIUS server(s) configured with the radius-server host command. • **tacacs+**—use the TACACS+ server(s) configured with the tacacs-server host command
2	**line** {**aux 0** | **console 0** | **vty** *number* [... *end-number*]}	CONFIGURATION	Enter the LINE mode.
3	**login authentication** {*method-list-name* | **default**}	LINE	Assign a *method-list-name* or the **default** list to the terminal line.

FTOS Behavior: If you use a method list on the console port in which RADIUS or TACACS is the last authentication method, and the server is not reachable, FTOS allows access even though the username and password credentials cannot be verified. Only the console port behaves this way, and does so to ensure that users are not locked out of the system in the event that network-wide issue prevents access to these servers.

To view the configuration, use the **show config** command in the LINE mode or the **show running-config** in the EXEC Privilege mode.

Note: Dell Force10 recommends that you use the **none** method only as a backup. This method does not authenticate users. The **none** and **enable** methods do not work with SSH.

You can create multiple method lists and assign them to different terminal lines.

Enable AAA Authentication

To enable AAA authentication, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
aaa authentication enable {*method-list-name* | **default**} *method1* [... *method4*]	CONFIGURATION	• **default**—Uses the listed authentication methods that follow this argument as the default list of methods when a user logs in. • *method-list-name*—Character string used to name the list of enable authentication methods activated when a user logs in. • *method1* [... *method4*]—Any of the following: RADIUS, TACACS, enable, line, none.

If the default list is not set, only the local enable is checked. This has the same effect as issuing: **aaa authentication enable default enable**

AAA Authentication—RADIUS

To enable authentication from the RADIUS server, and use TACACS as a backup, use the following commands:

Step	Command Syntax	Command Mode	Purpose
1	**aaa authentication enable default radius tacacs**	CONFIGURATION	To enable RADIUS and to set up TACACS as backup.
2	**radius-server host x.x.x.x key some-password**	CONFIGURATION	To establish host address and password.
3	**tacacs-server host x.x.x.x key some-password**	CONFIGURATION	To establish host address and password.

To get enable authentication from the RADIUS server, and use TACACS as a backup, issue the following commands:

```
FTOS(config)# aaa authentication enable default radius tacacs
Radius and TACACS server has to be properly setup for this.
FTOS(config)# radius-server host x.x.x.x key <some-password>
FTOS(config)# tacacs-server host x.x.x.x key <some-password>
```

To use local authentication for enable secret on console, while using remote authentication on VTY lines, perform the following steps:

```
FTOS(config)# aaa authentication enable mymethodlist radius tacacs
FTOS(config)# line vty 0 9
FTOS(config-line-vty)# enable authentication mymethodlist
```

Server-side configuration

TACACS+: When using TACACS+, Dell Force10 sends an initial packet with service type SVC_ENABLE, and then, a second packet with just the password. The TACACS server must have an entry for username $enable$.

RADIUS: When using RADIUS authentication, FTOS sends an authentication packet with the following:

```
Username: $enab15$
Password: <password-entered-by-user>
```

Therefore, the RADIUS server must have an entry for this username.

AAA Authorization

FTOS enables AAA new-model by default.You can set authorization to be either local or remote. Different combinations of authentication and authorization yield different results. By default, FTOS sets both to local.

Privilege Levels Overview

Limiting access to the system is one method of protecting the system and your network. However, at times, you might need to allow others access to the router and you can limit that access to a subset of commands. In FTOS, you can configure a privilege level for users who need limited access to the system.

Every command in FTOS is assigned a privilege level of 0, 1 or 15. You can configure up to 16 privilege levels in FTOS. FTOS is pre-configured with 3 privilege levels and you can configure 13 more. The three pre-configured levels are:

- **Privilege level 1**—is the default level for the EXEC mode. At this level, you can interact with the router, for example, view some show commands and Telnet and ping to test connectivity, but you cannot configure the router. This level is often called the “user” level. One of the commands available in Privilege level 1 is the **enable** command, which you can use to enter a specific privilege level.
- **Privilege level 0**—contains only the **end**, **enable** and **disable** commands.
- **Privilege level 15**—the default level for the **enable** command, is the highest level. In this level you can access any command in FTOS.

Privilege levels 2 through 14 are not configured and you can customize them for different users and access.

After you configure other privilege levels, enter those levels by adding the level parameter after the **enable** command or by configuring a user name or password that corresponds to the privilege level. Refer to Configure a username and password on page 921 for more information on configuring user names.

By default, commands in FTOS are assigned to different privilege levels. You can access those commands only if you have access to that privilege level. For example, to reach the **protocol spanning-tree** command, you must log in to the router, enter the **enable** command for privilege level 15 (this is the default level for the command) and then enter the CONFIGURATION mode.

You can configure passwords to control access to the box and assign different privilege levels to users. FTOS supports the use of passwords when you log in to the system and when you enter the **enable** command. If you move between privilege levels, you are prompted for a password if you move to a higher privilege level.

Configuration Task List for Privilege Levels

The following list has the configuration tasks for privilege levels and passwords.

- Configure a username and password on page 921 (mandatory)
- Configure the enable password command on page 922 (mandatory)
- Configure custom privilege levels on page 922 (mandatory)
- Specify LINE mode password and privilege on page 924 (optional)
- Enable and disabling privilege levels on page 925 (optional)

For a complete listing of all commands related to FTOS privilege levels and passwords, refer to the Security chapter in the *FTOS Command Reference*.

Configure a username and password

In FTOS, you can assign a specific username to limit user access to the system.

To configure a username and password, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
username *name* [**access-class** *access-list-name*] [**nopassword | password** [*encryption-type*] *password*] [**privilege** *level*]	CONFIGURATION	Assign a user name and password. Configure the optional and required parameters: • *name:* Enter a text string up to 63 characters long. • **access-class** *access-list-name:* Enter the name of a configured IP ACL. • **nopassword:** Do not require the user to enter a password. • *encryption-type:* Enter 0 for plain text or 7 for encrypted text. • *password:* Enter a string. • **privilege** *level* range: 0 to 15.

To view usernames, use the **show users** command in the EXEC Privilege mode.

Configure the enable password command

To configure FTOS, you must use the **enable** command to enter the EXEC Privilege level 15. After entering the command, FTOS requests that you enter a password. Privilege levels are not assigned to passwords, rather passwords are assigned to a privilege level. A password for any privilege level can always be changed. To change to a different privilege level, enter the **enable** command, followed by the privilege level. If you do not enter a privilege level, the default level 15 is assumed.

To configure a password for a specific privilege level, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
enable password [**level** *level*] [*encryption-mode*] *password*	CONFIGURATION	Configure a password for a privilege level. Configure the optional and required parameters: • **level** *level:* Specify a level 0 to 15. Level 15 includes all levels. • *encryption-type:* Enter 0 for plain text or 7 for encrypted text. • *password:* Enter a string. To change only the password for the enable command, configure only the *password* parameter.

To view the configuration for the **enable secret** command, use the **show running-config** command in the EXEC Privilege mode.

In custom-configured privilege levels, the **enable** command is always available. No matter what privilege level you entered FTOS, you can enter the **enable 15** command to access and configure all CLI.

Configure custom privilege levels

In addition to assigning privilege levels to the user, you can configure the privilege levels of commands so that they are visible in different privilege levels. Within FTOS, commands have certain privilege levels. With the privilege command, the default level can be changed or you can reset their privilege level back to the default.

- Assign the launch keyword (for example, **configure**) for the keyword's command mode.
- If you assign only the first keyword to the privilege level, all commands beginning with that keyword are also assigned to the privilege level. If you enter the entire command, the software assigns the privilege level to that command only.

To assign commands and passwords to a custom privilege level, you must be in privilege level 15 and use these commands in the following sequence in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**username** *name* [**access-class** *access-list-name*] [**privilege** *level*] [**nopassword | password** [*encryption-type*] *password*]	CONFIGURATION	Assign a user name and password. Configure the optional and required parameters: • *name:* Enter a text string (up to 63 characters). • **access-class** *access-list-name:* Enter the name of a configured IP ACL. • **privilege** *level* range: 0 to 15. • **nopassword:** Do not require the user to enter a password. • *encryption-type:* Enter 0 for plain text or 7 for encrypted text. • *password:* Enter a string.
2	**enable password** [**level** *level*] [*encryption-mode*] *password*	CONFIGURATION	Configure a password for privilege level. Configure the optional and required parameters: • **level** *level:* Specify a level 0 to 15. Level 15 includes all levels. • *encryption-type:* Enter 0 for plain text or 7 for encrypted text. • *password:* Enter a string up to 25 characters long. To change only the password for the enable command, configure only the *password* parameter.
3	**privilege** *mode* {**level** *level command* | **reset** *command*}	CONFIGURATION	Configure level and commands for a mode or reset a command's level. Configure the following required and optional parameters: • *mode:* Enter a keyword for the modes (exec, configure, interface, line, route-map, router) • **level** *level* range: 0 to 15. Levels 0, 1 and 15 are pre-configured. Levels 2 to 14 are available for custom configuration. • *command:* A FTOS CLI keyword (up to 5 keywords allowed). • **reset:** Return the command to its default privilege mode.

To view the configuration, use the **show running-config** command in the EXEC Privilege mode.

Figure 45-2 is an example of a configuration to allow a user "john" to view only the EXEC mode commands and all **snmp-serve**r commands. Since the **snmp-server** commands are "enable" level commands and, by default, found in the CONFIGURATION mode, you must also assign the launch command for the CONFIGURATION mode, **configure**, to the same privilege level as the **snmp-server** commands.

Figure 45-2. Configuring a Custom Privilege Level

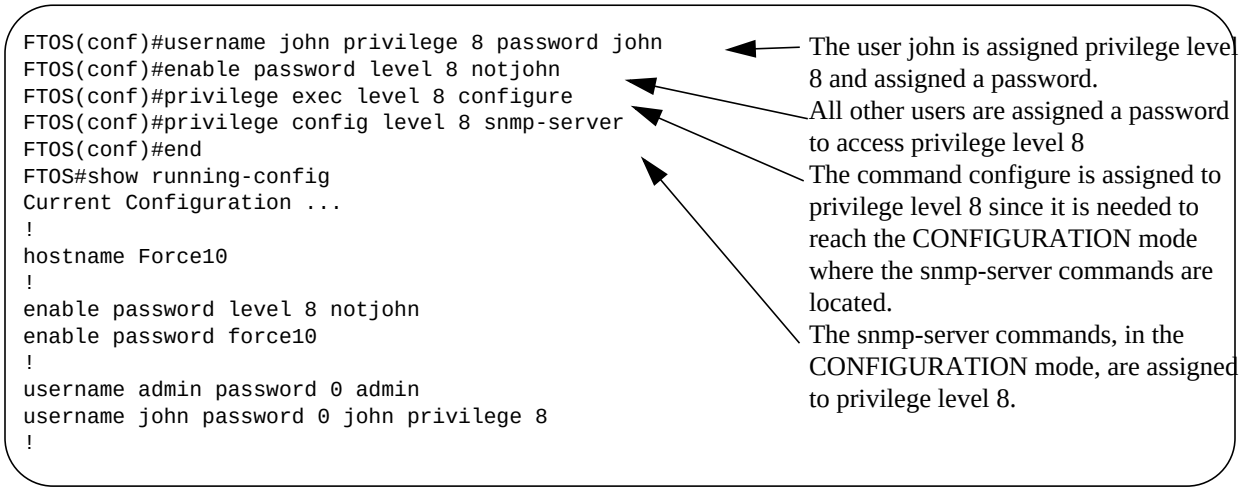


Figure 45-3 is a screen shot of the Telnet session for user “john”. The **show privilege** command output confirms that “john” is in privilege level 8. In the EXEC Privilege mode, “john” can access only the commands listed. In CONFIGURATION mode, “john” can access only the **snmp-server** commands.

Figure 45-3. User john’s Login and the List of Available Commands

```
apollo% telnet 172.31.1.53
Trying 172.31.1.53...
Connected to 172.31.1.53.
Escape character is '^]'.
Login: john
Password:
FTOS#show priv
Current privilege level is 8
FTOS#?
configure               Configuring from terminal
disable                 Turn off privileged commands
enable                  Turn on privileged commands
exit                    Exit from the EXEC
no                      Negate a command
show                    Show running system information
terminal                Set terminal line parameters
traceroute              Trace route to destination
FTOS#confi
FTOS(conf)#?
end                     Exit from Configuration mode
```

Specify LINE mode password and privilege

You can specify a password authentication of all users on different *terminal* lines. The user’s privilege level will be the same as the privilege level assigned to the terminal line, unless a more specific privilege level is is assigned to the user.

To specify a password for the terminal line, use the following commands, in any order, in the LINE mode:

Command Syntax	Command Mode	Purpose
privilege level *level*	LINE	Configure a custom privilege level for the terminal lines. • **level** *level* range: 0 to 15. Levels 0, 1 and 15 are pre-configured. Levels 2 to 14 are available for custom configuration.
password [*encryption-type*] *password*	LINE	Specify either a plain text or encrypted password. Configure the following optional and required parameters: • *encryption-type*: Enter 0 for plain text or 7 for encrypted text. • *password*: Enter a text string up to 25 characters long.

To view the password configured for a terminal, use the **show config** command in the LINE mode.

Enable and disabling privilege levels

Enter the **enable** or **enable privilege-level** command in the EXEC Privilege mode to set a user's security level. If you do not enter a privilege level, FTOS sets it to 15 by default.

To move to a lower privilege level, enter the command **disable** followed by the **level-number** you wish to set for the user in the EXEC Privilege mode. If you enter **disable** without a level-number, your security level is 1.

RADIUS

Remote Authentication Dial-In User Service (RADIUS) is a distributed client/server protocol. This protocol transmits authentication, authorization, and configuration information between a central RADIUS server and a RADIUS client (the Dell Force10 system). The system sends user information to the RADIUS server and requests authentication of the user and password. The RADIUS server returns one of the following responses:

- **Access-Accept**—the RADIUS server authenticates the user
- **Access-Reject**—the RADIUS server does not authenticate the user

If an error occurs in the transmission or reception of RADIUS packets, the error can be viewed by enabling the **debug radius** command.

Transactions between the RADIUS server and the client are encrypted (the users' passwords are not sent in plain text). RADIUS uses UDP as the transport protocol between the RADIUS server host and the client.

For more information on RADIUS, refer to RFC 2865, *Remote Authentication Dial-in User Service*.

RADIUS Authentication and Authorization

FTOS supports RADIUS for user authentication (text password) at login and can be specified as one of the login authentication methods in the **aaa authentication login** command.

When configuring AAA authorization, you can configure to limit the attributes of services available to a user. When authorization is enabled, the network access server uses configuration information from the user profile to issue the user's session. The user's access is limited based on the configuration attributes.

FTOS supports the following RADIUS attributes:

Code	Attribute
1	RADIUS_USER_NAME
2	RADIUS_USER_PASSWORD
4	RADIUS_NAS_IP_ADDRESS
5	RADIUS_NAS_PORT
11	RADIUS_FILTER_ID (for ACL)
26	RADIUS_VENDOR_SPECIFIC (privilege level/auto-command)
28	RADIUS_IDLE_TIMEOUT
61	RADIUS_NAS_PORT_TYPE
95	NAS_IPv6_ADDRESS

802.1x supported:

1	RADIUS_USER_NAME
4	RADIUS_NAS_IP_ADDRESS
5	RADIUS_NAS_PORT
24	RADIUS_STATE
30	RADIUS_CALLING_STATION_ID
61	RADIUS_NAS_PORT_TYPE
64	RADIUS_TUNNEL_TYPE
65	RADIUS_TUNNEL_MEDIUM_TYPE

79	RADIUS_EAP_MSG
80	RADIUS_MSG_AUTHENTICATOR
81	RADIUS_TUNNEL_PRIVATE_GROUP_ID
95	NAS_IPv6_ADDRESS

RADIUS exec-authorization stores a user-shell profile and that is applied during user login. You may name the relevant named-lists with either a unique name or the default name. When authorization is enabled by the RADIUS server, the server returns the following information to the client:

- Idle time
- ACL configuration information
- Auto-command
- Privilege level

After gaining authorization for the first time, you may configure these attributes.

Note: RADIUS authentication/authorization is done for every login. There is no difference between first-time login and subsequent logins.

Idle Time

Every session line has its own idle-time. If the idle-time value is not changed, the default value of 30 minutes is used. RADIUS specifies idle-time allow for a user during a session before timeout. When a user logs in, the lower of the two idle-time values (configured or default) is used. The idle-time value is updated if both of the following happens:

- The administrator changes the idle-time of the line on which the user has logged in
- The idle-time is lower than the RADIUS-returned idle-time

ACL

The RADIUS server can specify an ACL. If an ACL is configured on the RADIUS server, and if that ACL is present, user may be allowed access based on that ACL. If the ACL is absent, authorization fails, and a message is logged indicating the this.

RADIUS can specify an ACL for the user if both of the following are true:

- If an ACL is absent
- There is a very long delay for an entry, or a denied entry because of an ACL, and a message is logged

Note: The ACL name must be a string. Only standard ACLs in authorization (both RADIUS and TACACS) are supported. Authorization is denied in cases using Extended ACLs.

Auto-command

You can configure the system through the RADIUS server to automatically execute a command when you connect to a specific line. To do this, use the command **auto-command**. The auto-command is executed when the user is authenticated and before the prompt appears to the user.

Set access to privilege levels through RADIUS

Through the RADIUS server, you can use the command **privilege level** to configure a privilege level for the user to enter into when they connect to a session.This value is configured on the client system.

Configuration Task List for RADIUS

To authenticate users using RADIUS, at least one RADIUS server must be specified so that the system can communicate with and configure RADIUS as one of your authentication methods.

The following list includes the configuration tasks for RADIUS.

- Define a aaa method list to be used for RADIUS on page 928 (mandatory)
- Apply the method list to terminal lines on page 929 (mandatory except when using default lists)
- Specify a RADIUS server host on page 929 (mandatory)
- Set global communication parameters for all RADIUS server hosts on page 930 (optional)
- Monitor RADIUS on page 931 (optional)

For a complete listing of all FTOS commands related to RADIUS, refer to the Security chapter in the *FTOS Command Reference*.

Note: RADIUS authentication and authorization are done in a single step. Hence, authorization cannot be used independent of authentication. However, if RADIUS authorization is configured and authentication is not, then a message is logged stating this. During authorization, the next method in the list (if present) is used, or if another method is not present, an error is reported.

To view the configuration, use the **show config** in the LINE mode or the **show running-config** command in the EXEC Privilege mode.

Define a AAA method list to be used for RADIUS

To configure RADIUS to authenticate or authorize users on the system, you must create a AAA method list. Default method lists do not need to be explicitly applied to the line, so they are not mandatory. To create a method list, enter one of the following commands in CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
aaa authentication login *method-list-name* **radius**	CONFIGURATION	Enter a text string (up to 16 characters long) as the name of the method list you wish to use with the RADIUS authentication method.

Command Syntax	Command Mode	Purpose
aaa authorization exec {*method-list-name* | **default**} **radius tacacs+**	CONFIGURATION	Create methodlist with RADIUS and TACACS+ as authorization methods. Typical order of methods: RADIUS, TACACS+, Local, None. If authorization is denied by RADIUS, the session ends (**radius** should not be the last method specified).

Apply the method list to terminal lines

To enable RADIUS AAA login authentication for a method list, you must apply it to a terminal line. To configure a terminal line for RADIUS authentication and authorization, enter the following commands:

Command Syntax	Command Mode	Purpose
line {**aux 0** | **console 0** | **vty** *number* [*end-number*]}	CONFIGURATION	Enter the LINE mode.
login authentication {*method-list-name* | **default**}	LINE	Enable AAA login authentication for the specified RADIUS method list. This procedure is mandatory if you are not using default lists.
authorization exec *methodlist*	CONFIGURATION	To use the methodlist.

Specify a RADIUS server host

When configuring a RADIUS server host, you can set different communication parameters, such as the UDP port, the key password, the number of retries, and the timeout.

To specify a RADIUS server host and configure its communication parameters, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
radius-server host {*hostname* | *ipv4-address* | *ipv6-address*} [**auth-port** *port-number*] [**retransmit** *retries*] [**timeout** *seconds*] [**key** [*encryption-type*] *key*]	CONFIGURATION	Enter the host name or IP address of the RADIUS server host. Configure the optional communication parameters for the specific host: • **auth-port** *port-number* range: 0 to 65335. Enter a UDP port number. The default is 1812. • **retransmit** *retries* range: 0 to 100. Default is 3. • **timeout** *seconds* range: 0 to 1000. Default is 5 seconds. • **key** [*encryption-type*] *key:* Enter 0 for plain text or 7 for encrypted text, and a string for the key. The key can be up to 42 characters long. This key must match the key configured on the RADIUS server host. If these optional parameters are not configured, the global default values for all RADIUS host are applied.

To specify multiple RADIUS server hosts, configure the **radius-server host** command multiple times. If multiple RADIUS server hosts are configured, FTOS attempts to connect with them in the order in which they were configured. When FTOS attempts to authenticate a user, the software connects with the RADIUS server hosts one at a time, until a RADIUS server host responds with an accept or reject response.

If you want to change an optional parameter setting for a specific host, use the **radius-server host** command. To change the global communication settings to all RADIUS server hosts, refer to Set global communication parameters for all RADIUS server hosts on page 930.

To view the RADIUS configuration, use the **show running-config radius** command in the EXEC Privilege mode.

To delete a RADIUS server host, use the **no radius-server host** {*hostname* | *ip-address*} command.

Set global communication parameters for all RADIUS server hosts

You can configure global communication parameters (auth-port, key, retransmit, and timeout parameters) and specific host communication parameters on the same system. However, if both global and specific host parameters are configured, the specific host parameters override the global parameters for that RADIUS server host.

To set global communication parameters for all RADIUS server hosts, use any or all of the following commands in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
radius-server deadtime *seconds*	CONFIGURATION	Set a time interval after which a RADIUS host server is declared dead. • *seconds* range: 0 to 2147483647. Default: 0 seconds
radius-server key [*encryption-type*] *key*	CONFIGURATION	Configure a key for all RADIUS communications between the system and RADIUS server hosts. • *encryption-type:* Enter 7 to encrypt the password. Enter 0 to keep the password as plain text. • *key:* Enter a string. The key can be up to 42 characters long. You cannot use spaces in the key.
radius-server retransmit *retries*	CONFIGURATION	Configure the number of times FTOS retransmits RADIUS requests. • *retries* range: 0 to 100. Default is 3 retries.
radius-server timeout *seconds*	CONFIGURATION	Configure the time interval the system waits for a RADIUS server host response. • *seconds* range: 0 to 1000. Default is 5 seconds.

To view the configuration of RADIUS communication parameters, use the **show running-config** command in the EXEC Privilege mode.

Monitor RADIUS

To view information on RADIUS transactions, use the following command in the EXEC Privilege mode:

Command Syntax	Command Mode	Purpose
debug radius	EXEC Privilege	View RADIUS transactions to troubleshoot problems.

TACACS+

FTOS supports Terminal Access Controller Access Control System (TACACS+ client, including support for login authentication.

Configuration Task List for TACACS+

The following list includes the configuration task for TACACS+ functions:

- Choose TACACS+ as the Authentication Method
- Monitor TACACS+
- TACACS+ Remote Authentication and Authorization on page 933
- TACACS+ Remote Authentication and Authorization on page 933
- Specify a TACACS+ server host on page 934
- Choose TACACS+ as the Authentication Method on page 931

For a complete listing of all commands related to TACACS+, refer to the Security chapter in the *FTOS Command Reference*.

Choose TACACS+ as the Authentication Method

One of the login authentication methods available is TACACS+ and the user's name and password are sent for authentication to the TACACS hosts specified.To use TACACS+ to authenticate users, you must specify at least one TACACS+ server for the system to communicate with and configure TACACS+ as one of your authentication methods.

To select TACACS as the login authentication method, use these commands in the following sequence in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**tacacs-server host** {*ipv4-address* | *ipv6-address* | *host*}	CONFIGURATION	Configure a TACACS+ server host. Enter the IP address or host name of the TACACS+ server. Use this command multiple times to configure multiple TACACS+ server hosts.
2	**aaa authentication login** {*method-list-name* | **default**} **tacacs+** [*...method3*]	CONFIGURATION	Enter a text string (up to 16 characters long) as the name of the method list you wish to use with the TACAS+ authentication method The **tacacs+** method should not be the last method specified.
3	**line** {**aux 0** | **console 0** | **vty** *number* [*end-number*]}	CONFIGURATION	Enter the LINE mode.
4	**login authentication** {*method-list-name* | **default**}	LINE	Assign the *method-list* to the terminal line.

To view the configuration, use the **show config** in the LINE mode or the **show running-config tacacs+** command in the EXEC Privilege mode.

If authentication fails using the primary method, FTOS employs the second method (or third method, if necessary) automatically. For example, if the TACACS+ server is reachable, but the server key is invalid, FTOS proceeds to the next authentication method. In Figure 45-4, the TACACS+ is incorrect, but the user is still authenticated by the secondary method.

Figure 45-4. Failed Authentication

```
FTOS(conf)#
FTOS(conf)#do show run aaa
!
aaa authentication enable default tacacs+ enable
aaa authentication enable LOCAL enable tacacs+
aaa authentication login default tacacs+ local
aaa authentication login LOCAL local tacacs+
aaa authorization exec default tacacs+ none
aaa authorization commands 1 default tacacs+ none
aaa authorization commands 15 default tacacs+ none
aaa accounting exec default start-stop tacacs+
aaa accounting commands 1 default start-stop tacacs+
aaa accounting commands 15 default start-stop tacacs+
FTOS(conf)#
FTOS(conf)#do show run tacacs+
!
tacacs-server key 7 d05206c308f4d35b        Server key purposely changed to incorrect value
tacacs-server host 10.10.10.10 timeout 1
FTOS(conf)#tacacs-server key angeline
FTOS(conf)#%RPM0-P:CP %SEC-5-LOGIN_SUCCESS: Login successful for user admin on vty0
(10.11.9.209)
%RPM0-P:CP %SEC-3-AUTHENTICATION_ENABLE_SUCCESS: Enable password authentication
success on vty0 ( 10.11.9.209 )
%RPM0-P:CP %SEC-5-LOGOUT: Exec session is terminated for user admin on line vty0
(10.11.9.209)
FTOS(conf)#username angeline password angeline
FTOS(conf)#%RPM0-P:CP %SEC-5-LOGIN_SUCCESS: Login successful for user angeline on
vty0 (10.11.9.209)
%RPM0-P:CP %SEC-3-AUTHENTICATION_ENABLE_SUCCESS: Enable password authentication
success on vty0 ( 10.11.9.209 )   User authenticated using secondary method
```

Monitor TACACS+

To view information on TACACS+ transactions, use the following command in the EXEC Privilege mode:

Command Syntax	Command Mode	Purpose
debug tacacs+	EXEC Privilege	View TACACS+ transactions to troubleshoot problems.

TACACS+ Remote Authentication and Authorization

FTOS takes the access class from the TACACS+ server. Access class is the class of service that restricts Telnet access and packet sizes. If you have configured remote authorization, then FTOS ignores the access class you have configured for the VTY line. FTOS instead gets this access class information from the TACACS+ server. FTOS needs to know the username and password of the incoming user before it can fetch the access class from the server. A user, therefore, will at least see the login prompt. If the access class denies the connection, FTOS closes the Telnet session immediately.

Figure 45-5 demonstrates how to configure the **access-class** from a TACACS+ server. This causes the configured access-class on the VTY line to be ignored. If you have configured a **deny10** ACL on the TACACS+ server, FTOS downloads it and applies it. If the user is found to be coming from the 10.0.0.0 subnet, FTOS also immediately closes the Telnet connection. Note, that no matter where the user is coming from, they see the login prompt.

Figure 45-5. Specify a TACACS+ server host

```
FTOS#
FTOS(conf)#
FTOS(conf)#ip access-list standard deny10
FTOS(conf-ext-nacl)#permit 10.0.0.0/8
FTOS(conf-ext-nacl)#deny any
FTOS(conf)#
FTOS(conf)#aaa authentication login tacacsmethod tacacs+
FTOS(conf)#aaa authentication exec tacacsauthorization tacacs+
FTOS(conf)#tacacs-server host 25.1.1.2 key force10
FTOS(conf)#
FTOS(conf)#line vty 0 9
FTOS(config-line-vty)#login authentication tacacsmethod
FTOS(config-line-vty)#authorization exec tacauthor
FTOS(config-line-vty)#
FTOS(config-line-vty)#access-class deny10
FTOS(config-line-vty)#end
```

When configuring a TACACS+ server host, you can set different communication parameters, such as the the key password.

To specify a TACACS+ server host and configure its communication parameters, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
tacacs-server host {*hostname* | *ipv4-address* | *ipv6-address*} [**port** *port-number*] [**timeout** *seconds*] [**key** *key*]	CONFIGURATION	Enter the host name or IP address of the TACACS+ server host. Configure the optional communication parameters for the specific host: • **port** *port-number* range: 0 to 65335. Enter a TCP port number. The default is 49. • **timeout** *seconds* range: 0 to 1000. Default is 10 seconds. • **key** *key:* Enter a string for the key. The key can be up to 42 characters long. This key must match a key configured on the TACACS+ server host. This parameter should be the last parameter configured. If these optional parameters are not configured, the default global values are applied.

To specify multiple TACACS+ server hosts, configure the **tacacs-server host** command multiple times. If multiple TACACS+ server hosts are configured, FTOS attempts to connect with them in the order in which they were configured.

To view the TACACS+ configuration, use the **show running-config tacacs+** command in the EXEC Privilege mode.

To delete a TACACS+ server host, use the **no tacacs-server host** {*hostname* | *ip-address*} command.

```
freebsd2# telnet 2200:2200:2200:2200:2200::2202
Trying 2200:2200:2200:2200:2200::2202...
Connected to 2200:2200:2200:2200:2200::2202.
Escape character is '^]'.
Login: admin
Password:
FTOS#
FTOS#
```

Command Authorization

The AAA command authorization feature configures FTOS to send each configuration command to a TACACS server for authorization before it is added to the running configuration.

By default, the AAA authorization commands configure the system to check both EXEC mode and CONFIGURATION mode commands. Use the command **no aaa authorization config-commands** to enable only EXEC mode command checking.

If rejected by the AAA server, the command is not added to the running config, and messages similar to Message 1 are displayed.

Message 1 Configuration Command Rejection

```
04:07:48: %RPM0-P:CP %SEC-3-SEC_AUTHORIZATION_FAIL: Authorization failure  Command
authorization failed for user (denyall) on vty0 ( 10.11.9.209 )
```

Protection from TCP Tiny and Overlapping Fragment Attacks

Tiny and overlapping fragment attack is a class of attack where configured ACL entries—denying TCP port-specific traffic—can be bypassed, and traffic can be sent to its destination although denied by the ACL. RFC 1858 and 3128 proposes a countermeasure to the problem. This countermeasure is configured into the line cards and enabled by default.

SCP and SSH

Secure Shell (SSH) is a protocol for secure remote login and other secure network services over an insecure network. FTOS is compatible with SSH versions 1.5 and 2, both the client and server modes. SSH sessions are encrypted and use authentication.

FTOS supports both inbound and outbound SSH sessions using IPv4 or IPv6 addressing. Inbound SSH supports accessing the system through the management interface as well as through a physical Layer 3 interface.

For details on command syntax, see the Security chapter in the *FTOS Command Line Interface Reference*.

SCP is a remote file copy program that works with SSH and is supported by FTOS.

Note: The Windows-based WinSCP client software is not supported for secure copying between a PC and an FTOS-based system. Unix-based SCP client software is supported.

To use the SSH client, use the following command in the EXEC Privilege mode:

Command Syntax	Command Mode	Purpose
ssh {*hostname* | *hostip*} [**-l** *username* | **-p** *port-number* | **-v** {**1** | **2**}	EXEC Privilege	Open an SSH connection specifying the hostname or *hostip*, username, port number, and version of the SSH client. *hostip* is the IP address of the remote device, which can be an IPv4 address (A.B.C.D)or IPv6 address (X:X:X:X::X).

To enable the SSH server for version 1 and 2, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip ssh server {**enable | port** *port-number*}	CONFIGURATION	Configure the Dell Force10 system as an SCP/SSH server.

To enable the SSH server for version 1 or 2 only, use the following command:

Command Syntax	Command Mode	Purpose
ip ssh server version {**1|2**}	CONFIGURATION	Configure the Dell Force10 system as an SSH server that uses only version 1 or 2.

To view the SSH configuration, use the following command in EXEC Privilege mode:

Command Syntax	Command Mode	Purpose
show ip ssh	EXEC Privilege	Display SSH connection information.

Figure 45-6 on page 937 shows the use of the command **ip ssh server version 2** to enable SSH version 2, and the **show ip ssh** command to confirm the setting.

Figure 45-6. Specifying an SSH version

```
FTOS(conf)#ip ssh server version 2
FTOS(conf)#do show ip ssh
SSH server                : disabled.
SSH server version        : v2.
Password Authentication   : enabled.
Hostbased Authentication  : disabled.
RSA       Authentication  : disabled.
```

To disable SSH server functions, enter **no ip ssh server enable**.

Using SCP with SSH to copy a software image

To use Secure Copy (SCP) to copy a software image through an SSH connection from one switch to another, use the following procedure:

Step	Task	Command Syntax	Command Mode
1	On Chassis One, set the SSH port number (port 22 by default).	**ip ssh server port** *number*	CONFIGURATION
2	On Chassis One, enable SSH.	**ip ssh server enable**	CONFIGURATION
3	On Chassis Two, invoke SCP.	**copy scp: flash:**	CONFIGURATION
4	On Chassis Two, in response to prompts, enter the path to the desired file and enter the port number specified in Step 1.		EXEC Privilege

This example shows the use of SCP and SSH to copy a software image from one switch running SSH Server on UDP port 99 to the local switch:

Figure 45-7. Using SCP to copy from an SSH Server on another Switch

```
.FTOS#copy scp: flash:
Address or name of remote host []: 10.10.10.1
Port number of the server [22]: 99
Source file name []: test.cfg
User name to login remote host: admin
Password to login remote host:
```

Other SSH-related commands include:

- **crypto key generate:** Generate keys for the SSH server.
- **debug ip ssh:** Enables collecting SSH debug information.
- **ip scp topdir:** Identify a location for files used in secure copy transfer.
- **ip ssh authentication-retries:** Configure the maximum number of attempts that should be used to authenticate a user.

- **ip ssh connection-rate-limit:** Configure the maximum number of incoming SSH connections per minute.
- **ip ssh hostbased-authentication enable:** Enable hostbased-authentication for the SSHv2 server.
- **ip ssh key-size:** Configure the size of the server-generated RSA SSHv1 key.
- **ip ssh password-authentication enable:** Enable password authentication for the SSH server.
- **ip ssh pub-key-file:** Specify the file to be used for host-based authentication.
- **ip ssh rhostsfile:** Specify the rhost file to be used for host-based authorization.
- **ip ssh rsa-authentication enable:** Enable RSA authentication for the SSHv2 server.
- **ip ssh rsa-authentication:** Add keys for the RSA authentication.
- **show crypto:** Display the public part of the SSH host-keys.
- **show ip ssh client-pub-keys:** Display the client public keys used in host-based authentication.
- **show ip ssh rsa-authentication:** Display the authorized-keys for the RSA authentication.
- **ssh-peer-rpm**: Open an SSH connection to the peer RPM.

Secure Shell Authentication

Secure Shell (SSH) is disabled by default. Enable it using the command **ip ssh server enable**.

SSH supports three methods of authentication:

- SSH Authentication by Password on page 938
- RSA Authentication of SSH on page 939
- Host-based SSH Authentication on page 939

Important Points to Remember for SSH Authentication

- If more than one method is enabled, the order in which the methods are preferred is based on the *ssh_config* file on the Unix machine.
- When all the three authentication methods are enabled, password authentication is the backup method when the RSA method fails.
- The files *known_hosts* and *known_hosts2* are generated when a user tries to SSH using version 1 or version 2, respectively.

SSH Authentication by Password

Authenticate an SSH client by prompting for a password when attempting to connect to the Dell Force10 system. This is the simplest methods of authentication and uses SSH version 1.

Enable SSH password authentication using the command **ip ssh password-authentication enable** from CONFIGURATION mode. View your SSH configuration using the command **show ip ssh** from EXEC Privilege mode.

Figure 45-8. Enabling SSH Password Authentication

```
FTOS(conf)#ip ssh server enable
                % Please wait while SSH Daemon initializes ... done.
FTOS(conf)#ip ssh password-authentication enable
FTOS#sh ip ssh
SSH server                 : enabled.
Password  Authentication   : enabled.
Hostbased Authentication   : disabled.
RSA       Authentication   : disabled.
```

RSA Authentication of SSH

The following procedure authenticates an SSH client based on an RSA key using RSA authentication. This method uses SSH version 2:

Step	Task	Command Syntax	Command Mode
1	On the SSH client (Unix machine), generate an RSA key, as shown in Figure 45-9.		
2	Copy the public key *id_rsa.pub* to the Dell Force10 system.		
3	Disable password authentication if enabled.	**no ip ssh password-authentication enable**	CONFIGURATION
4	Enable RSA authentication.	**ip ssh rsa-authentication enable**	EXEC Privilege
5	Bind the public keys to RSA authentication.	**ip ssh rsa-authentication my-authorized-keys flash://** *public_key*	EXEC Privilege

Figure 45-9. Generating RSA Keys

```
admin@Unix_client#ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/home/admin/.ssh/id_rsa):
/home/admin/.ssh/id_rsa already exists.
Overwrite (y/n)? y
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/admin/.ssh/id_rsa.
Your public key has been saved in /home/admin/.ssh/id_rsa.pub.
```

Host-based SSH Authentication

Authenticate a particular host. This method uses SSH version 2.

To configure host-based authentication:

Step	Task	Command Syntax	Command Mode
1	Configure RSA Authentication. See RSA Authentication of SSH, above.		
2	Create shosts by copying the public RSA key to the to the file *shosts* in the diretory *.ssh*, and write the IP address of the host to the file.	**cp /etc/ssh/ssh_host_rsa_key.pub /.ssh/shosts**	
3	Create a list of IP addresses and usernames that are permitted to SSH in a file called *rhosts*, as shown in Figure 45-11.		
4	Copy the file shosts and rhosts to the Dell Force10 system.		
5	Disable password authentication and RSA authentication, if configured	• **no ip ssh password-authentication** • **no ip ssh rsa-authentication**	• CONFIGURATION • EXEC Privilege
6	Enable host-based authentication.	**ip ssh hostbased-authentication enable**	CONFIGURATION
7	Bind shosts and rhosts to host-based authentication.	**ip ssh pub-key-file flash://***filename* **ip ssh rhostsfile flash://***filename*	CONFIGURATION

Figure 45-10. Creating shosts

```
admin@Unix_client# cd /etc/ssh
admin@Unix_client# ls
moduli      sshd_config      ssh_host_dsa_key.pub  ssh_host_key.pub
ssh_host_rsa_key.pub  ssh_config  ssh_host_dsa_key  ssh_host_key
ssh_host_rsa_key
admin@Unix_client# cat ssh_host_rsa_key.pub
ssh-rsa      AAAAB3NzaC1yc2EAAAABIwAAAIEA8K7jLZRVfjqHJzUQmXxuIbZx/
AyWhVgJpOh39k8v3e8eQvLnHBIsqIL8jVv1QHhUeb7GaDlJvEDAMz30mvqQbJgXBBRTWgBpLWwL/
doyUXFufjiL9YmoVfkbKcFmxJEMkE3JyHanEi7hg34LChjk9hL1by8cYZP2kYS2lnSyQWk=
admin@Unix_client# ls
id_rsa  id_rsa.pub  shosts
admin@Unix_client# cat shosts
10.16.127.201  ssh-rsa AAAAB3NzaC1yc2EAAAABIwAAAIEA8K7jLZRVfjqHJzUQmXxuIbZx/AyW
hVgJpOh39k8v3e8eQvLnHBIsqIL8jVv1QHhUeb7GaDlJvEDAMz30mvqQbJgXBBRTWgBpLWwL/
doyUXFufjiL9YmoVfkbKcFmxJEMkE3JyHanEi7hg34LChjk9hL1by8cYZP2kYS2lnSyQWk=
```

Figure 45-11. Creating rhosts

```
admin@Unix_client# ls
id_rsa  id_rsa.pub  rhosts  shosts
admin@Unix_client# cat rhosts
10.16.127.201 admin
```

Client-based SSH Authentication

SSH from the chassis to the SSH client using using the command **ssh** *ip_address*. This method uses SSH version 1 or version 2. If the SSH port is a non-default value, use the command **ip ssh server port** *number*, to change the default port number. You may only change the port number when SSH is disabled. When must then still use the **-p** option with the command **ssh**.

Figure 45-12. Client-based SSH Authentication

```
FTOS#ssh 10.16.127.201 ?
-l                       User name option
-p                      SSH server port option (default 22)
-v                      SSH protocol version
```

Troubleshooting SSH

- You may not bind *id_rsa.pub* to RSA authentication while logged in via the console. In this case, Message 2 appears.

Message 2 RSA Authentication Error

```
%Error: No username set for this term.
```

- Host-based authentication must be enabled on the server (Dell Force10 system) and the client (Unix machine). Message 3 appears if you attempt to log in via SSH and host-based is disabled on the client. In this case, verify that host-based authentication is set to "Yes" in the file *ssh_config* (root permission is required to edit this file).

Message 3 Host-based Authentication Error

```
permission denied (host based)
```

- If the IP address in the RSA key does not match the IP address from which you attempt to log in, Message 4 appears. In this case, verify that the name and IP address of the client is contained in the file */etc/hosts*.

Message 4 RSA Authentication Error

```
getname info 8 failed
```

Telnet

To use Telnet with SSH, you must first enable SSH, as described above.

By default, the Telnet daemon is enabled. If you want to disable the Telnet daemon, use the following command, or disable Telnet in the startup config.

Use the [**no**] **ip telnet server enable** command to enable or disable the Telnet daemon.

```
FTOS(conf)#ip telnet server enable
FTOS(conf)#no ip telnet server enable
```

Trace Lists

The Trace Lists feature is supported only on the E-Series: E

You can log packet activity on a port to confirm the source of traffic attacking a system. Once the Trace list is enabled on the system, you view its traffic log to confirm the source address of the attacking traffic. In FTOS, Trace lists are similar to extended IP ACLs, except that Trace lists are not applied to an interface. Instead, Trace lists are enabled for all switched traffic entering the system.

The number of entries allowed per trace list is 1K.

In the E-Series, you can create a trace filter based on any of the following criteria:

- Source IP address
- Destination IP address
- Source TCP port number
- Destination TCP port number
- Source UDP port number
- Destination UDP port number

For trace lists, you can match criteria on specific or ranges of TCP or UDP ports or established TCP sessions.

Note: If there are unresolved next-hops and a trace-list is enabled, there is a possibility that the traffic hitting the CPU will not be rate-limited.

When creating a trace list, the sequence of the filters is important. You have a choice of assigning sequence numbers to the filters as you enter them, or FTOS assigns numbers in the order the filters were created. For more information on sequence numbering, refer to Chapter 21, IP Access Control Lists, Prefix Lists, and Route-maps, on page 419.

Configuration Tasks for Trace Lists

The following configuration steps include mandatory and optional steps.

- Creating a trace list on page 942 (mandatory)
- Apply trace lists on page 947 (mandatory)

For a complete listing of all commands related to trace lists, refer to the Security chapter in the *FTOS Command Reference*.

Creating a trace list

Trace lists filter and log traffic based on source and destination IP addresses, IP host addresses, TCP addresses, TCP host addresses, UDP addresses, and UDP host addresses. When configuring the Trace list filters, include the **count** and **bytes** parameters so that any hits to that filter are logged.

Since traffic passes through the filter in the order of the filter's sequence, you can configure the trace list by first entering the TRACE LIST mode and then assigning a sequence number to the filter.

To create a filter for packets with a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip trace-list** *trace-list-name*	CONFIGURATION	Enter the TRACE LIST mode by creating an trace list.
2	**seq** *sequence-number* {**deny** | **permit**} {**ip** | *ip-protocol-number*} {*source mask* | **any** | **host** *ip-address*} {*destination mask* | **any** | **host** *ip-address*} [**count** [**byte**] | **log**]	TRACE LIST	Configure a drop or forward filter. Configure the following required and optional parameters: • *sequence-number* range: 0 to, 4294967290. • **ip**: to specify IP as the protocol to filter for. • *ip-protocol-number* range: 0 to 255. • *source*: An IP address as the source IP address for the filter to match. • *mask:* a network mask • **any**: to match any IP source address • **host** *ip-address:* to match IP addresses in a host. • *destination*: An IP address as the source IP address for the filter to match. • **count:** count packets processed by the filter. • **byte:** count bytes processed by the filter. • **log:** is supported.

To create a filter for TCP packets with a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip trace-list** *trace-list-name*	CONFIGURATION	Create a trace list and assign it a unique name.

Step	Command Syntax	Command Mode	Purpose
2	**seq** *sequence-number* {**deny** | **permit**} **tcp** {*source mask* | **any** | **host** *ip-address*} [*operator port* [*port*]] {*destination mask* | **any** | **host** *ip-address*} [*operator port* [*port*]] [**established**] [**count** [**byte**] | **log**]	TRACE LIST	Configure a trace list filter for TCP packets. • *source*: An IP address as the source IP address for the filter to match. • *mask:* a network mask • **any**: to match any IP source address • **host** *ip-address:* to match IP addresses in a host. • *destination*: An IP address as the source IP address for the filter to match. • **count:** count packets processed by the filter. • **byte:** count bytes processed by the filter. • **log:** is supported.

To create a filter for UDP packets with a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ip trace-list** *access-list-name*	CONFIGURATION	Create a trace list and assign it a unique name.
2	**seq** *sequence-number* {**deny** | **permit**} **udp** {*source mask* | **any** | **host** *ip-address*} [*operator port* [*port*]] {*destination mask* | **any** | **host** *ip-address*} [*operator port* [*port*]] [**count** [**byte**] | **log**]	TRACE LIST	Configure a trace list filter for UDP packets. • *source*: An IP address as the source IP address for the filter to match. • *mask:* a network mask • **any**: to match any IP source address • **host** *ip-address:* to match IP addresses in a host. • *destination*: An IP address as the source IP address for the filter to match. • **count:** count packets processed by the filter. • **byte:** count bytes processed by the filter. • **log:** is supported.

When you create the filters with a specific sequence number, you can create the filters in any order and the filters are placed in the correct order.

Note: When assigning sequence numbers to filters, keep in mind that you might need to insert a new filter. To prevent reconfiguring multiple filters, assign sequence numbers in multiples of five or another number.

Figure 45-13 illustrates how the **seq** command orders the filters according to the sequence number assigned. In the example, filter 15 was configured before filter 5, but the **show config** command displays the filters in the correct order.

Figure 45-13. Trace list Using seq Command Example

```
FTOS(config-trace-acl)#seq 15 deny ip host 12.45.0.0 any log
FTOS(config-trace-acl)#seq 5 permit tcp 121.1.3.45 0.0.255.255 any
FTOS(config-trace-acl)#show conf
!
ip trace-list dilling
 seq 5 permit tcp 121.1.0.0 0.0.255.255 any
 seq 15 deny ip host 12.45.0.0 any log
```

If you are creating a Trace list with only one or two filters, you can let FTOS assign a sequence number based on the order in which the filters are configured. FTOS assigns filters in multiples of 5.

To configure a filter for a Trace list without a specified sequence number, use any or all of the following commands in the TRACE LIST mode:

Command Syntax	Command Mode	Purpose
{**deny** | **permit**} {**ip** | *ip-protocol-number*} {*source mask* | **any** | **host** *ip-address*} {*destination mask* | **any** | **host** *ip-address*} [**count** [**byte**] | **log**]	TRACE LIST	Configure a deny or permit filter to examine IP packets. Configure the following required and optional parameters: • **ip**: to specify IP as the protocol to filter for. • *ip-protocol-number* range: 0 to 255. • *source*: An IP address as the source IP address for the filter to match. • *mask:* a network mask • **any**: to match any IP source address • **host** *ip-address:* to match IP addresses in a host. • *destination*: An IP address as the source IP address for the filter to match. • **count:** count packets processed by the filter. • **byte:** count bytes processed by the filter. • **log:** is supported.

Command Syntax	Command Mode	Purpose
{**deny** | **permit**} **tcp** {*source mask* | **any** | **host** *ip-address*} [*operator port* [*port*]] {*destination mask* | **any** | **host** *ip-address*} [*operator port* [*port*]] [**established**] [**count** [**byte**] | **log**]	TRACE LIST	Configure a deny or permit filter to examine TCP packets. Configure the following required and optional parameters: • *source*: An IP address as the source IP address for the filter to match. • *mask:* a network mask • **any**: to match any IP source address • **host** *ip-address:* to match IP addresses in a host. • *destination*: An IP address as the source IP address for the filter to match. • **precedence** *precedence* range: 0 to 7. • **tos** *tos-value* range: 0 to 15 • **count:** count packets processed by the filter. • **byte:** count bytes processed by the filter. • **log:** is supported.
{**deny** | **permit**} **udp** {*source mask* | **any** | **host** *ip-address*} [*operator port* [*port*]] {*destination mask* | **any** | **host** *ip-address*} [*operator port* [*port*]] | **log**]	TRACE LIST	Configure a deny or permit filter to examine UDP packets. Configure the following required and optional parameters: • *source*: An IP address as the source IP address for the filter to match. • *mask:* a network mask • **any**: to match any IP source address • **host** *ip-address:* to match IP addresses in a host. • *destination*: An IP address as the source IP address for the filter to match. • **precedence** *precedence* range: 0 to 7. • **tos** *tos-value* range: 0 to 15 • **count:** count packets processed by the filter. • **byte:** count bytes processed by the filter. • **log:** is supported.

Figure 45-14 illustrates a Trace list in which the sequence numbers were assigned by the software. The filters were assigned sequence numbers based on the order in which they were configured (for example, the first filter was given the lowest sequence number). The **show config** command in the TRACE LIST mode displays the two filters with the sequence numbers 5 and 10.

Figure 45-14. Trace List Example

```
FTOS(config-trace-acl)#deny tcp host 123.55.34.0 any
FTOS(config-trace-acl)#permit udp 154.44.123.34 0.0.255.255 host 34.6.0.0
FTOS(config-trace-acl)#show config
!
ip trace-list nimule
 seq 5 deny tcp host 123.55.34.0 any
 seq 10 permit udp 154.44.0.0 0.0.255.255 host 34.6.0.0
```

To view all configured Trace lists and the number of packets processed through the Trace list, use the **show ip accounting trace-list** command (Figure 110) in the EXEC Privilege mode.

Apply trace lists

After you create a Trace list, you must enable it. Without enabling the Trace list, no traffic is filtered.

You can enable one Trace list.

To enable a Trace list, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ip trace-group *trace-list-name*	CONFIGURATION	Enable a configured Trace list to filter traffic.

To remove a Trace list, use the **no ip trace-group** *trace-list-name* command syntax.

Once the Trace list is enabled, you can view its log with the **show ip accounting trace-list** *trace-list-name* [**linecard** *number*] command.

Figure 45-15. show ip accounting trace-list Command Example

```
FTOS#show ip accounting trace-list dilling
Trace List dilling on linecard 0
 seq 2 permit ip host 10.1.0.0 any count (0 packets)
 seq 5 deny ip any any
FTOS#
```

VTY Line and Access-Class Configuration

Various methods are available to restrict VTY access in FTOS. These depend on which authentication scheme you use — line, local, or remote:

Table 45-1. VTY Access

Authentication Method	VTY access-class support?	Username access-class support?	Remote authorization support?
Line	YES	NO	NO
Local	NO	YES	NO
TACACS+	YES	NO	YES (with FTOS 5.2.1.0 and later)
RADIUS	YES	NO	YES (with FTOS 6.1.1.0 and later)

FTOS provides several ways to configure access classes for VTY lines, including:

- VTY Line Local Authentication and Authorization on page 948
- VTY Line Remote Authentication and Authorization on page 949

VTY Line Local Authentication and Authorization

FTOS retrieves the access class from the local database. To use this feature:

1. Create a username
2. Enter a password
3. Assign an access class
4. Enter a privilege level

Line authentication can be assigned on a per-VTY basis; it is a simple password authentication, using an access-class as authorization.

Local authentication is configured globally. You configure access classes on a per-user basis.

FTOS can assign different access classes to different users by username. Until users attempt to log in, FTOS does not know if they will be assigned a VTY line. This means that incoming users always see a login prompt even if you have excluded them from the VTY line with a **deny-all** access class. Once users identify themselves, FTOS retrieves the access class from the local database and applies it. (FTOS also subsequently can close the connection if a user is denied access).

Figure 45-16 shows how to allow or deny a Telnet connection to a user. Users will see a login prompt, even if they cannot login. No access class is configured for the VTY line. It defaults from the local database.

Figure 45-16. Example Access-Class Configuration Using Local Database

```
FTOS(conf)#user gooduser password abc privilege 10 access-class permitall
FTOS(conf)#user baduser password abc privilege 10 access-class denyall
FTOS(conf)#
FTOS(conf)#aaa authentication login localmethod local
FTOS(conf)#
FTOS(conf)#line vty 0 9
FTOS(config-line-vty)#login authentication localmethod
FTOS(config-line-vty)#end
```

Note: See also the section Chapter 8, IP Access Control Lists (ACL), Prefix Lists, and Route-maps.

VTY Line Remote Authentication and Authorization

FTOS retrieves the access class from the VTY line.

The Dell Force10 OS takes the access class from the VTY line and applies it to ALL users. FTOS does not need to know the identity of the incoming user and can immediately apply the access class. If the authentication method is radius, TACACS+, or line, and you have configured an access class for the VTY line, FTOS immediately applies it. If the access-class is **deny all** or **deny for the incoming subnet**, FTOS closes the connection without displaying the login prompt. Figure shows how to deny incoming connections from subnet 10.0.0.0 without displaying a login prompt. The example uses TACACS+ as the authentication mechanism.

Figure 45-17. Example Access Class Configuration Using TACACS+ Without Prompt

```
FTOS(conf)#ip access-list standard deny10
FTOS(conf-ext-nacl)#permit 10.0.0.0/8
FTOS(conf-ext-nacl)#deny any
FTOS(conf)#
FTOS(conf)#aaa authentication login tacacsmethod tacacs+
FTOS(conf)#tacacs-server host 256.1.1.2 key force10
FTOS(conf)#
FTOS(conf)#line vty 0 9
FTOS(config-line-vty)#login authentication tacacsmethod
FTOS(config-line-vty)#
FTOS(config-line-vty)#access-class deny10
FTOS(config-line-vty)#end
(same applies for radius and line authentication)
```

VTY MAC-SA Filter Support

FTOS supports MAC access lists which permit or deny users based on their source MAC address. With this approach, you can implement a security policy based on the source MAC address.

To apply a MAC ACL on a VTY line, use the same **access-class** command as IP ACLs (Figure 45-18). Figure 45-18 shows how to deny incoming connections from subnet 10.0.0.0 without displaying a login prompt..

Figure 45-18. Example Access Class Configuration Using TACACS+ Without Prompt

```
FTOS(conf)#mac access-list standard sourcemac
FTOS(config-std-mac)#permit 00:00:5e:00:01:01
FTOS(config-std-mac)#deny any
FTOS(conf)#
FTOS(conf)#line vty 0 9
FTOS(config-line-vty)#access-class sourcemac
FTOS(config-line-vty)#end
```

46

Service Provider Bridging

Service Provider Bridging is supported on platforms: C E S

This chapter contains the following major sections:

VLAN Stacking

VLAN Stacking is supported on platforms: C E S

VLAN Stacking is supported on E-Series ExaScale E X with FTOS 8.2.1.0. and later.

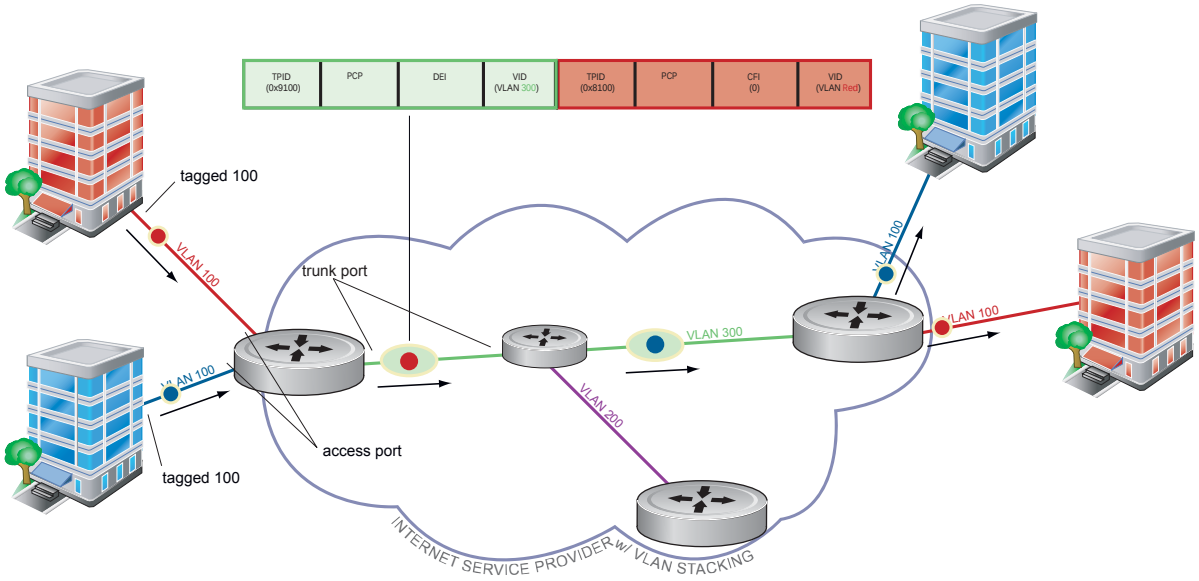
VLAN Stacking, also called Q-in-Q, is defined in IEEE 802.1ad—Provider Bridges, which is an amendment to IEEE 802.1Q—Virtual Bridged Local Area Networks. It enables service providers to use 802.1Q architecture to offer separate VLANs to customers with no coordination between customers, and minimal coordination between customers and the provider.

Using only 802.1Q VLAN tagging all customers would have to use unique VLAN IDs to ensure that traffic is segregated, and customers and the service provider would have to coordinate to ensure that traffic mapped correctly across the provider network. Even under ideal conditions, customers and the provider would still share the 4094 available VLANs.

Instead, 802.1ad allows service providers to add their own VLAN tag to frames traversing the provider network. The provider can then differentiate customers even if they use the same VLAN ID, and providers can map multiple customers to a single VLAN to overcome the 4094 VLAN limitation. Forwarding decisions in the provider network are based on the provider VLAN tag only, so the provider can map traffic through the core independently; the customer and provider need only coordinate at the provider edge.

In at the access point of a VLAN-stacking network, service providers add a VLAN tag, the S-Tag, to each frame before the 802.1Q tag. From this point, the frame is double-tagged. The service provider uses the S-Tag, to forward the frame traffic across its network. At the egress edge, the provider removes the S-Tag, so that the customer receives the frame in its original condition (Figure 46-1).

Figure 46-1. VLAN Stacking in a Service Provider Network



Important Points to Remember

- Interfaces that are members of the Default VLAN and are configured as VLAN-Stack access or trunk ports do not switch untagged traffic. To switch traffic, these interfaces must be added to a non-default VLAN-Stack-enabled VLAN.
- Dell Force10 cautions against using the same MAC address on different customer VLANs, on the same VLAN-Stack VLAN.
- You can ping across a trunk port only if both systems on the link are an E-Series. You cannot ping across the link if one or both of the systems is a C-Series or S-Series.
- This limitation becomes relevant if you enable the port as a multi-purpose port (carrying single-tagged and double-tagged traffic).

Configure VLAN Stacking

Configuring VLAN-Stacking is a three-step process:

1. Create access and trunk ports. See page 953.
2. Assign access and trunk ports to a VLAN. See page 954.
3. Make the VLAN VLAN-stacking capable.

Related Configuration Tasks

- Configure the Protocol Type Value for the Outer VLAN Tag on page 954
- FTOS Options for Trunk Ports on page 955
- VLAN Stacking in Multi-vendor Networks on page 956

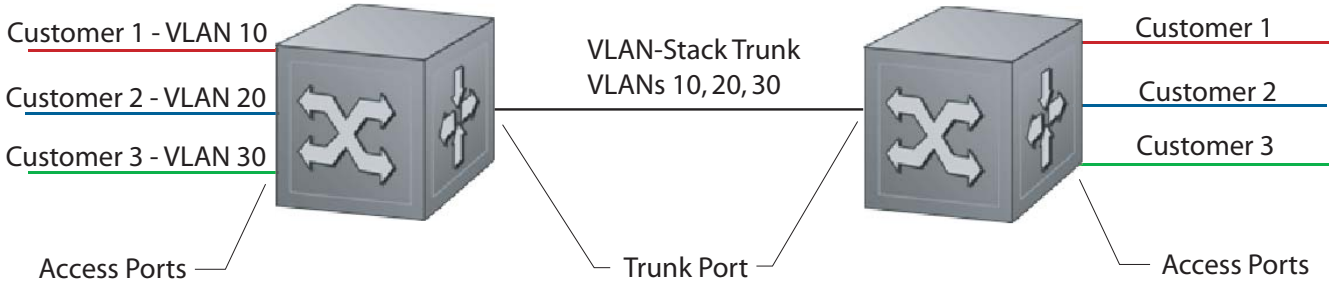
Create Access and Trunk Ports

An **access port** is a port on the service provider edge that directly connects to the customer. An access port may belong to only one service provider VLAN.

A **trunk port** is a port on a service provider bridge that connects to another service provider bridge and is a member of multiple service provider VLANs.

Physical ports and port-channels can be access or trunk ports.

Figure 46-2. Access and Trunk Ports



To create access and trunk ports:

Step	Task	Command Syntax	Command Mode
1	Assign the role of access port to a Layer 2 port on a provider bridge that is connected to a customer.	**vlan-stack access**	INTERFACE
2	Assign the role of trunk port to a Layer 2 port on a provider bridge that is connected to another provider bridge.	**vlan-stack trunk**	INTERFACE
3	Assign all access ports and trunk ports to service provider VLANs.	**member**	INTERFACE VLAN

Display the VLAN-Stacking configuration for a switchport using the command **show config** from INTERFACE mode, as shown in Figure 46-3.

Figure 46-3. Displaying the VLAN-Stack Configuration on a Layer 2 Port

```
FTOS#show run interface gi 7/0
!
interface GigabitEthernet 7/0
 no ip address
 switchport
 vlan-stack access
 no shutdown
FTOS#show run interface gi 7/12
!
interface GigabitEthernet 7/12
 no ip address
 switchport
 vlan-stack trunk
 no shutdown
```

Enable VLAN-Stacking for a VLAN

To enable VLAN-Stacking for a VLAN:

Task	Command Syntax	Command Mode
Enable VLAN-Stacking for the VLAN.	INTERFACE VLAN	**vlan-stack compatible**

Display the status and members of a VLAN using the **show vlan** command from EXEC Privilege mode. Members of a VLAN-Stacking-enabled VLAN are marked with an *M* in column *Q*.

Figure 46-4. Display the Members of a VLAN-Stacking-enabled VLAN

```
FTOS#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status    Q Ports
*   1      Active    U Gi 13/0-5,18
    2      Inactive
    3      Inactive
```

Configure the Protocol Type Value for the Outer VLAN Tag

The Tag Protocol Identifier (TPID) field of the S-Tag is user-configurable:

Task	Command Syntax	Command Mode
Select a value for the S-Tag TPID. Default: 9100	CONFIGURATION	**vlan-stack protocol-type**

Display the S-Tag TPID for a VLAN using the command **show running-config** from EXEC privilege mode. FTOS displays the S-Tag TPID only if it is a non-default value.

FTOS Options for Trunk Ports

802.1ad trunk ports may also be tagged members of a VLAN so that it can carry single and double-tagged traffic.

You can enable trunk ports to carry untagged, single-tagged, and double-tagged VLAN traffic by making the trunk port a hybrid port.

Step	Task	Command Syntax	Command Mode
1	Configure a trunk port to carry untagged, single-tagged, and double-tagged traffic by making it a hybrid port. **Note:** Note: On the C-Series and S-Series, a trunk port can be added to an 802.1Q VLAN as well as a Stacking VLAN only when the TPID 0x8100.	**portmode hybrid**	INTERFACE
2	Add the port to a 802.1Q VLAN as tagged or untagged.	[**tagged** | **untagged**]	INTERFACE VLAN

In Figure 46-5 GigabitEthernet 0/1 a trunk port that is configured as a hybrid port and then added to VLAN 100 as untagged VLAN 101 as tagged, and VLAN 103, which is a stacking VLAN.

Figure 46-5. Hybrid Port as VLAN-Stack Trunk Port and as Member of other VLANs

```
FTOS(conf)#int gi 0/1
FTOS(conf-if-gi-0/1)#portmode hybrid
FTOS(conf-if-gi-0/1)#switchport
FTOS(conf-if-gi-0/1)#vlan-stack trunk
FTOS(conf-if-gi-0/1)#show config
!
interface GigabitEthernet 0/1
 no ip address
 portmode hybrid
 switchport
 vlan-stack trunk
 shutdown
FTOS(conf-if-gi-0/1)#interface vlan 100
FTOS(conf-if-vl-100)#untagged gigabitethernet 0/1
FTOS(conf-if-vl-100)#interface vlan 101
FTOS(conf-if-vl-101)#tagged gigabitethernet 0/1
FTOS(conf-if-vl-101)#interface vlan 103
FTOS(conf-if-vl-103)#vlan-stack compatible
FTOS(conf-if-vl-103-stack)#member gigabitethernet 0/1
FTOS(conf-if-vl-103-stack)#do show vlan

Codes: * - Default VLAN, G - GVRP VLANs
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged, M - Vlan-stack

    NUM    Status    Description                  Q Ports
*   1      Inactive
    100    Inactive                               U Gi 0/1
    101    Inactive                               T Gi 0/1
    103    Inactive                               M Gi 0/1
```

VLAN Stacking in Multi-vendor Networks

The first field in the VLAN tag is the Tag Protocol Identifier (TPID), which is two bytes. In a VLAN-stacking network, once the frame is double tagged, the outer tag TPID must match the TPID of the next-hop system.

While 802.1Q requires that the inner tag TPID is 0x8100, it does not require a specific value for the outer tag TPID. Systems may use any two-byte value; FTOS uses 0x9100 (Figure 46-6) while non-Dell Force10 systems might use a different value.

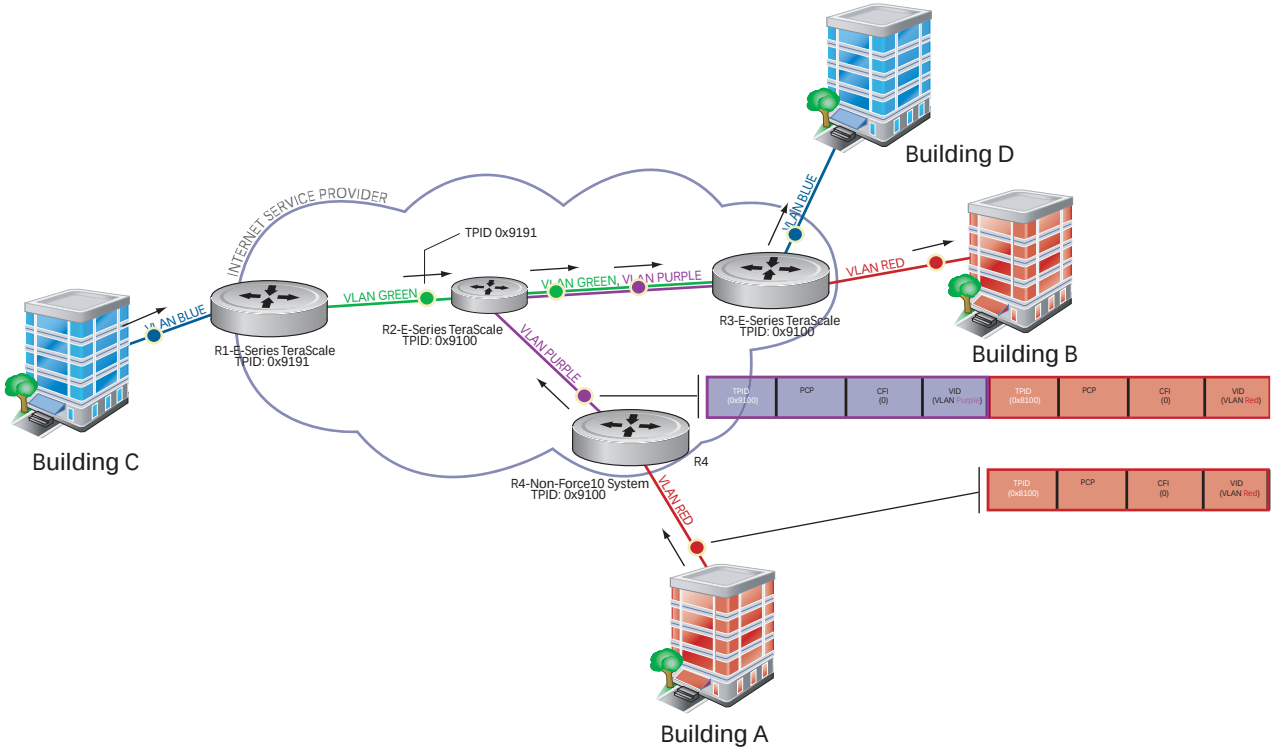
If the next-hop system's TPID does not match the outer-tag TPID of the incoming frame, the system drops the frame. For example, in Figure 46-6, the frame originating from Building A is tagged VLAN RED, and then double-tagged VLAN PURPLE on egress at R4. The TPID on the outer tag is 0x9100. R2's TPID must also be 0x9100, and it is, so R2 forwards the frame.

Given the matching-TPID requirement, there are limitations when you employ Dell Force10 systems at network edges, at which, frames are either double tagged on ingress (R4) or the outer tag is removed on egress (R3).

VLAN Stacking with E-Series TeraScale Systems

The default TPID for the outer VLAN tag is 0x9100. Although the TPID field is 16 bits, E-Series TeraScale only uses the first eight to make forwarding decisions, and as such makes no distinction between 0x8100 and any other TPID beginning with 0x81, for example, 0x8181. You can configure the first eight bits of the TPID using the command **vlan-stack protocol-type** command. In Figure 46-6, the frame originating from Building C is tagged 0x9191 on egress of R1. R2's TPID is 0x9100, but it its an E-Series TeraScale system and makes no distinction between 0x9191 and 0x9100, so it forwards the frame.

Figure 46-6. TPID Match and First-byte Match on the E-Series TeraScale

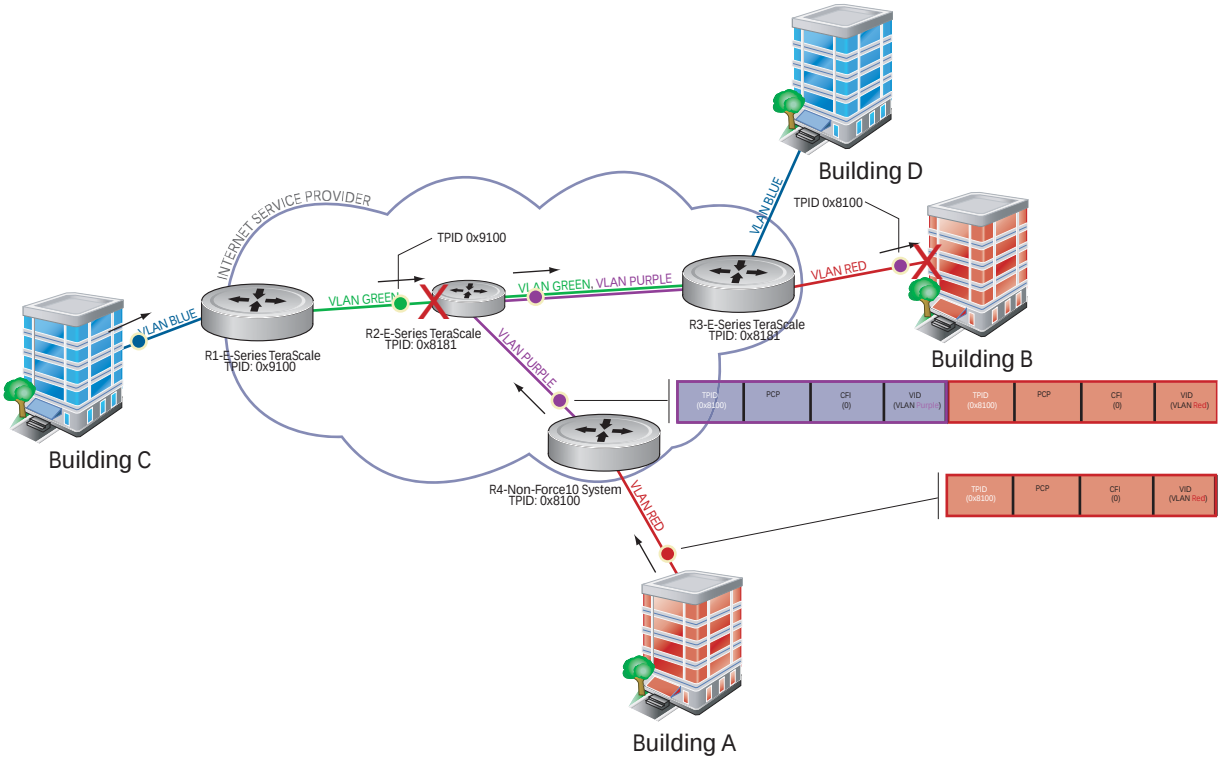


TPID 0x8100 on E-Series TeraScale Systems

E-Series TeraScale treats TPID 0x8100 as a normal VLAN even when on the outer tag. E-Series TeraScale makes forwarding decisions based strictly on the protocol type, without regard for whether the port is an access port. Therefore, when the outer tag has TPID 0x8100, the system does not remove it from frames egressing an access port. Still, although the frames cannot be decapsulated, the system is able to switch them. In Figure 46-7, the frame originating from Building A is double tagged on egress at R4 and is switched towards Building B, but is not decapsulated on egress at R2 because its TPID is 0x8181.

FTOS Behavior: The E-Series ExaScale and TeraScale forward frames with TPID 0x8100 even when its own TPID is not 0x8100. This behavior is required to service ARP and PVST packets, which use TPID 0x8100.

Figure 46-7. TPID Mismatch and 0x8100 Match on the E-Series TeraScale



VLAN Stacking with E-Series ExaScale Systems

E-Series ExaScale, beginning with FTOS version 8.2.1.0, allows you to configure both bytes of the 2-byte TPID. TeraScale systems allow you to configure the first byte only and thus, the system did not differentiate between TPIDs with a common first byte. For example 0x9100 and 0x91A8 were treated as the same TPID. In Figure 46-6, R2 forwards the frame with TPID 0x9191 which originated from Building C. In contrast, R2 drops the frame with TPID 0x9191 originating from Building C in Figure 46-8 because the frames TPID does not match both bytes of its own TPID.

FTOS Behavior: The E-Series ExaScale and TeraScale forwards frames with TPID 0x8100 even when its own TPID is not 0x8100. This behavior is required to service ARP and PVST packets, which use TPID 0x8100.

Figure 46-8. First-byte TPID Match on the E-Series ExaScale

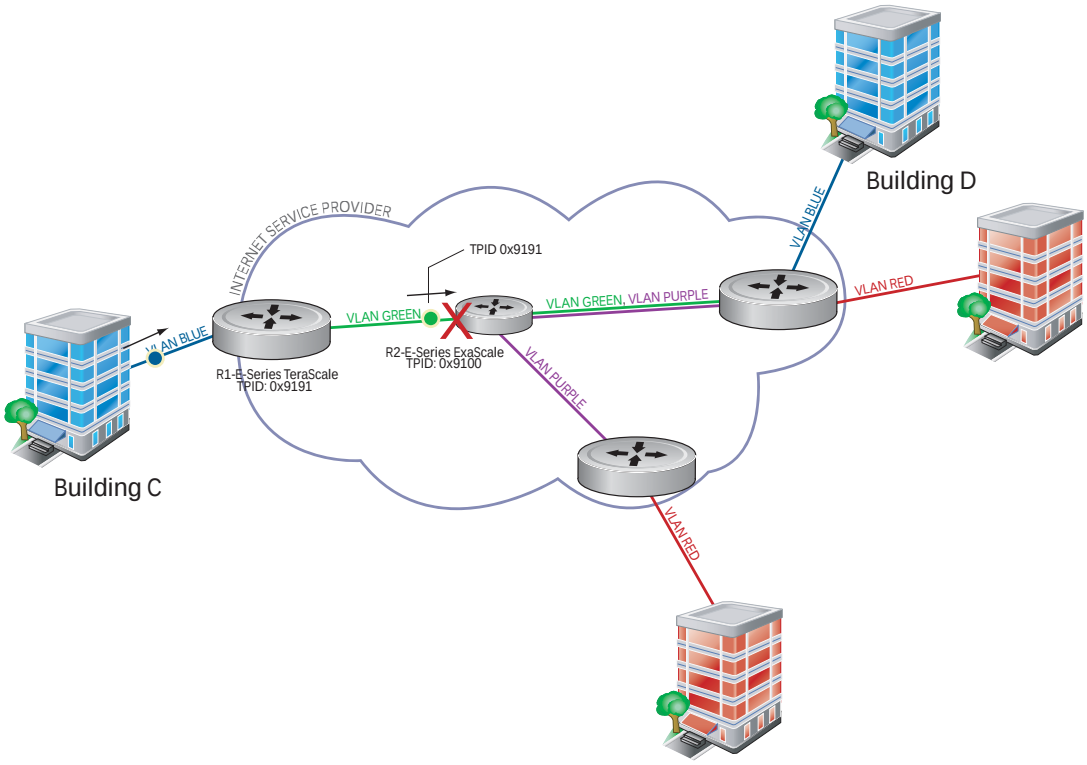


Table 46-1 details the outcome of matched and mis-matched TPIDs in a VLAN-stacking network with the E-Series.

Table 46-1. E-Series Behaviors for Mis-matched TPID

Network Position	Incoming Packet TPID	System TPID	Match Type	TeraScale Behavior	ExaScale Behavior
Core	0xUV**WX**	0xUV**YZ**	1st-byte match	switch as 0xUV**YZ**	drop
	0xUVWZ	0xQRST	mismatch	drop	drop
Egress Access Point	0xUV**WX**	0xUV**YZ**	1st-byte match	switch as 0xUV**YZ**	drop
	0x81**WX**	0x81**YZ**	1st-byte match	switch as is (no decapsulation)	drop
	0xUVWZ	0xQRST	mismatch	drop	drop

VLAN Stacking with C-Series and S-Series

The default TPID for the outer VLAN tag is 0x9100. Beginning with FTOS version 8.2.1.0, both the C-Series and S-Series allow you to configure both bytes of the 2-byte TPID. Previous versions allowed you to configure the first byte only, and thus, the systems did not differentiate between TPIDs with a common first byte. For example 0x8100 and any other TPID beginning with 0x81 were treated as the same TPID, as shown in Figure 46-9. Versions 8.2.1.0 and later differentiate between 0x9100 and 0x91XY, as shown in Figure 46-11.

You can configure the first eight bits of the TPID using the command **vlan-stack protocol-type**.

The TPID on the C-Series and S-Series systems is global. Ingress frames that do not match the system TPID are treated as untagged. This rule applies for both the outer tag TPID of a double-tagged frame and the TPID of a single-tagged frame.

For example, if you configure TPID 0x9100, then the system treats 0x8100 and untagged traffic the same and maps both types to the default VLAN, as shown by the frame originating from Building C in Figure 46-11. For the same traffic types, if you configure TPID 0x8100, then the system is able to differentiate between 0x8100 and untagged traffic and maps each to the appropriate VLAN, as shown by the packet originating from Building A in Figure 46-11.

Therefore, a mismatched TPID results in the port not differentiating between tagged and untagged traffic.

Figure 46-9. Single and Double-tag TPID Match on the C-Series and S-Series

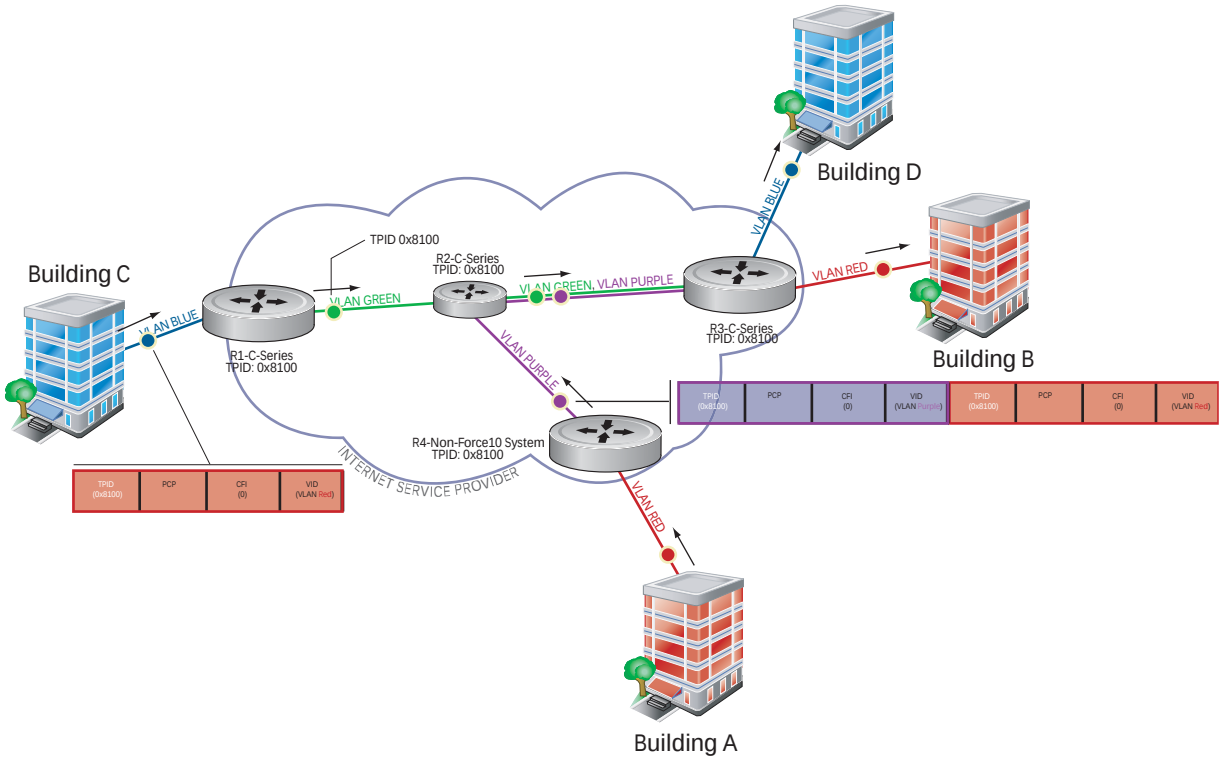


Figure 46-10. Single and Double-tag First-byte TPID Match on C-Series and S-Series

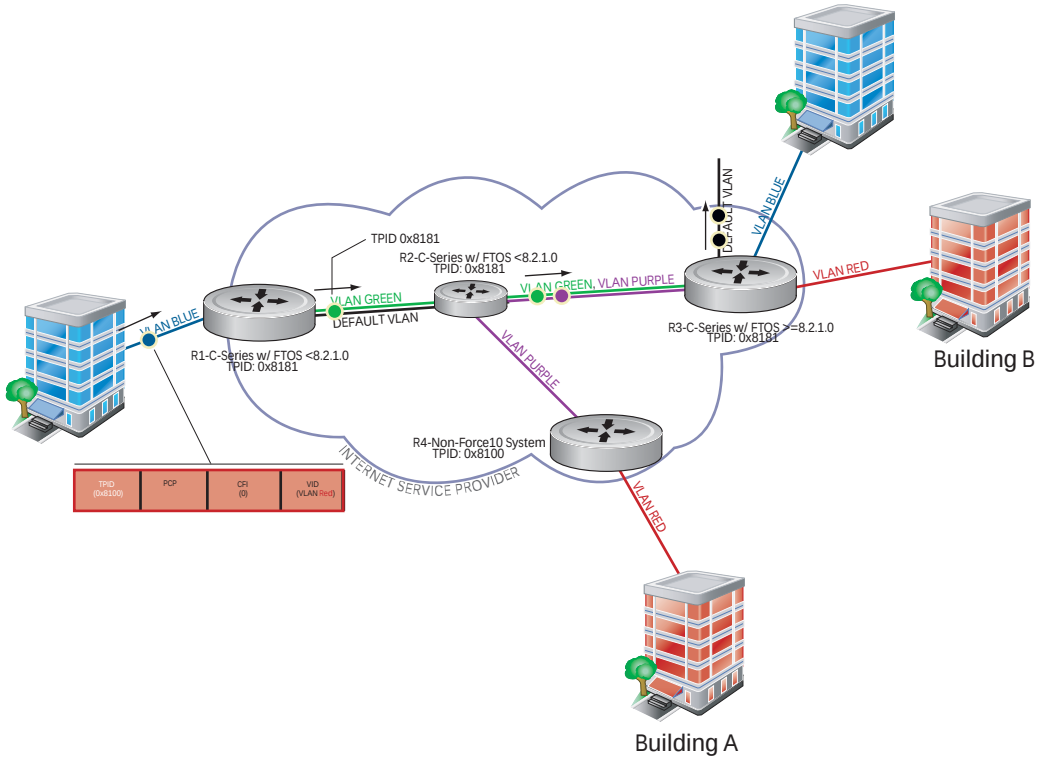


Figure 46-11. Single and Double-tag TPID Mismatch on the C-Series and S-Series

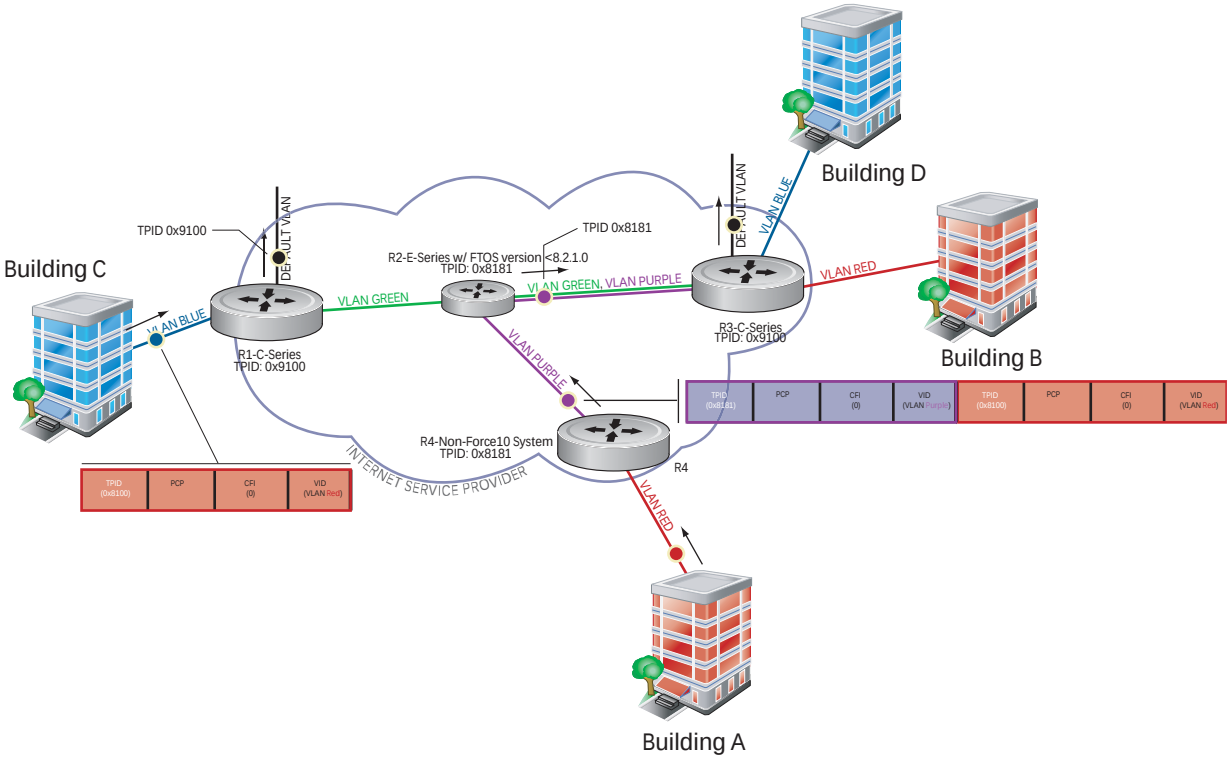


Table 46-2 details the outcome of matched and mismatched TPIDs in a VLAN-stacking network with the C-Series and S-Series.

Table 46-2. C-Series and S-Series Behaviors for Mis-matched TPID

Network Position	Incoming Packet TPID	System TPID	Match Type	Pre-8.2.1.0	8.2.1.0+
Ingress Access Point	untagged	0xUVWX	—	switch to default VLAN	switch to default VLAN
	single-tag (0x8100)	0xUVWX	single-tag mismatch	switch to default VLAN	switch to default VLAN
		0x8100	single-tag match	switch to VLAN	switch to VLAN
		0x81XY	single-tag first-byte match	switch to VLAN	switch to default VLAN
Core	untagged	0xUVWX	—	switch to default VLAN	switch to default VLAN
	double-tag 0xUVWX	0xUVWX	double-tag match	switch to VLAN	switch to VLAN
		0xUVYZ	double-tag first-byte match	switch to VLAN	switch to default VLAN
		0xQRST	double-tag mismatch	switch to default VLAN	switch to default VLAN
Egress Access Point	untagged	0xUVWX	—	switch to default VLAN	switch to default VLAN
	double-tag 0xUVWX	0xUVWX	double-tag match	switch to VLAN	switch to VLAN
		0xUVYZ	double-tag first-byte match	switch to VLAN	switch to default VLAN
		0xQRST	double-tag mismatch	switch to default VLAN	switch to default VLAN

VLAN Stacking Packet Drop Precedence

VLAN Stacking Packet Drop Precedence is available only on platform: C S

The Drop Eligible Indicator (DEI) bit in the S-Tag indicates to a service provider bridge which packets it should prefer to drop when congested.

Enable Drop Eligibility

You must enable Drop Eligibility globally before you can honor or mark the DEI value.

Task	Command Syntax	Command Mode
Make packets eligible for dropping based on their DEI value. By default, packets are colored green, and DEI is marked 0 on egress.	**dei enable**	CONFIGURATION

When Drop Eligibility is enabled, DEI mapping or marking takes place according to the defaults. In this case, the CFI is affected according to Table 46-3.

Table 46-3. Drop Eligibility Behavior

Ingress	Egress	DEI Disabled	DEI Enabled
Normal Port	Normal Port	Retain CFI	Set CFI to 0
Trunk Port	Trunk Port	Retain inner tag CFI	Retain inner tag CFI
		Retain outer tag CFI	Set outer tag CFI to 0
Access Port	Trunk Port	Retain inner tag CFI	Retain inner tag CFI
		Set outer tag CFI to 0	Set outer tag CFI to 0

Honor the Incoming DEI Value

To honor the incoming DEI value, you must explicitly map the DEI bit to an FTOS drop precedence; precedence can have one of three colors:

Precedence	Description
Green	High priority packets that are the least preferred to be dropped.
Yellow	Lower priority packets that are treated as best-effort.
Red	Lowest priority packets that are *always* dropped (regardless of congestion status).

Task	Command Syntax	Command Mode
Honor the incoming DEI value by mapping it to an FTOS drop precedence. You may enter the command once for 0 and once for 1. Packets with an unmapped DEI value are colored green.	**dei honor** {**0** | **1**} {**green** | **red** | **yellow**}	INTERFACE
Display the DEI-honoring configuration.	**show interface dei-honor** [**interface** *slot/port* | **linecard** *number* **port-set** *number*]	EXEC Privilege

Task	Command Syntax	Command Mode

```
FTOS#show interface dei-honor

Default Drop precedence: Green
Interface             CFI/DEI                 Drop precedence
-------------------------------------------------------------
Gi 0/1                0                       Green
Gi 0/1                1                       Yellow
Gi 8/9                1                       Red
Gi 8/40               0                       Yellow
```

Mark Egress Packets with a DEI Value

On egress, you can set the DEI value according to a different mapping than ingress (see Honor the Incoming DEI Value).

Task	Command Syntax	Command Mode
Set the DEI value on egress according to the color currently assigned to the packet.	**dei mark** {**green** | **yellow**} {**0** | **1**}	INTERFACE
Display the DEI-marking configuration.	**show interface dei-mark** [*interface slot/port* | **linecard** *number* **port-set** *number*]	EXEC Privilege

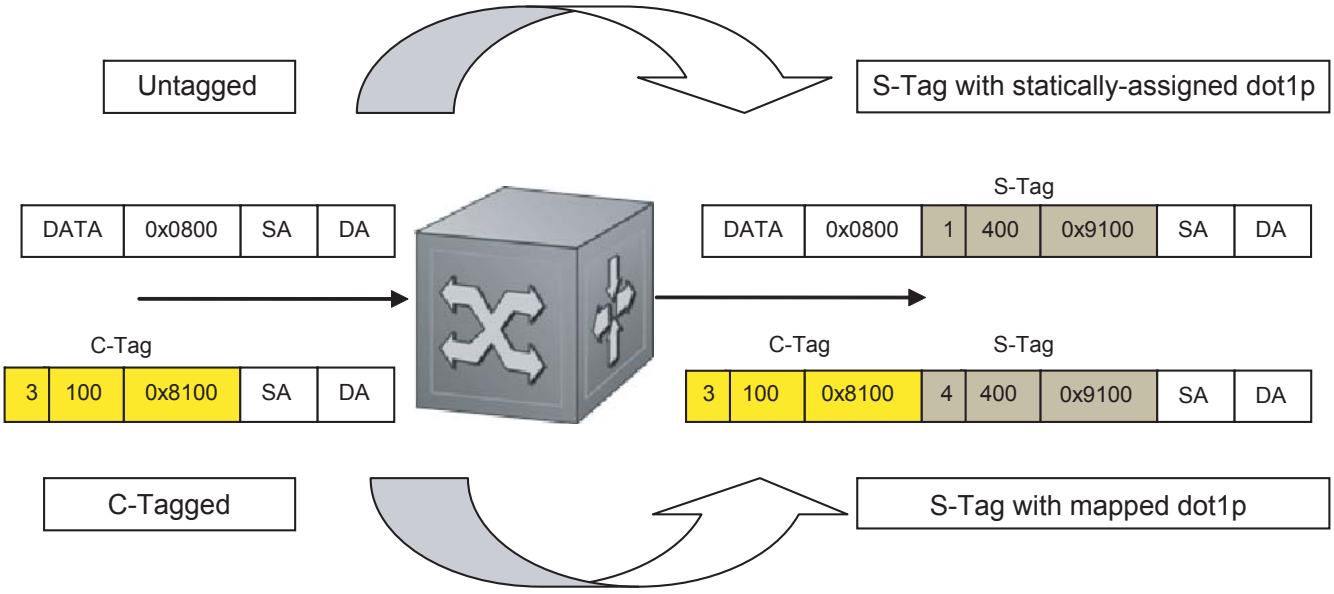
```
FTOS#show interface dei-mark

Default CFI/DEI Marking: 0
Interface             Drop precedence     CFI/DEI
-------------------------------------------------
Gi 0/1                Green               0
Gi 0/1                Yellow              1
Gi 8/9                Yellow              0
Gi 8/40               Yellow              0
```

Dynamic Mode CoS for VLAN Stacking

Dynamic Mode CoS for VLAN Stacking is available only on platforms: C S

One of the ways to ensure quality of service for customer VLAN-tagged frames is to use the 802.1p priority bits in the tag to indicate the level of QoS desired. When an S-Tag is added to incoming customer frames, the 802.1p bits on the S-Tag may be configured statically for each customer or derived from the C-Tag using *Dynamic Mode CoS*. Dynamic Mode CoS maps the C-Tag 802.1p value to a S-Tag 802.1p value.

Figure 46-12. Statically and Dynamically Assigned dot1p for VLAN Stacking



When configuring Dynamic Mode CoS, you have two options:

a mark the S-Tag dot1p and queue the frame according to the original C-Tag dot1p. In this case, you must have other dot1p QoS configurations; this option is classic dot1p marking.

b mark the S-Tag dot1p and queue the frame according to the S-Tag dot1p. For example, if frames with C-Tag dot1p values 0, 6 and 7 are mapped to an S-Tag dot1p value 0, then all such frames are sent to the queue associated with the S-Tag 802.1p value 0. This option requires two different CAM entries, each in a different Layer 2 ACL FP block.

Note: The ability to map incoming C-Tag dot1p to any S-Tag dot1p requires up to 8 entries to be installed in the Layer 2 QoS and Layer 2 ACL table for each configured customer VLAN. The scalability of this feature is limited by the impact of the 1:8 expansion in these CAM tables.

FTOS Behavior: For Option A above, when there is a conflict between the queue selected by Dynamic Mode CoS (**vlan-stack dot1p-mapping**) and a QoS configuration, the queue selected by Dynamic Mode CoS takes precedence. However, rate policing for the queue is determined by QoS configuration. For example, the following access-port configuration maps all traffic to Queue 0:

```
vlan-stack dot1p-mapping c-tag-dot1p 0-7 sp-tag-dot1p 1
```

However, if the following QoS configuration also exists on the interface, traffic is queued to Queue 0 but will be rate policed at 40Mbps (qos-policy-input for queue 3) since class-map "a" of Queue 3 also matches the traffic. This behavior is expected.

```
policy-map-input in layer2
service-queue 3 class-map a qos-policy 3
!
class-map match-any a layer2
match mac access-group a
!
mac access-list standard a
seq 5 permit any
!
qos-policy-input 3 layer2
rate-police 40
```

Likewise, in the configuration below, packets with dot1p priority 0 – 3 are marked as dot1p 7 in the outer tag and queued to Queue 3. Rate policing is according to **qos-policy-input 3**. All other packets will have outer dot1p 0 and hence are queued to Queue 1. They are therefore policed according to **qos-policy-input 1**.

A policy map output with rate shape for different queues can also be used.

```
policy-map-input in layer2
 service-queue 1 qos-policy 1
 service-queue 3 qos-policy 3
!
qos-policy-input 1 layer2
 rate-police 10
!
qos-policy-input 3 layer2
 rate-police 30
!
interface GigabitEthernet 0/21
 no ip address
 switchport
 vlan-stack access
 vlan-stack dot1p-mapping c-tag-dot1p 0-3 sp-tag-dot1p 7
 service-policy input in layer2
 no shutdown
```

To map C-Tag dot1p values to S-Tag dot1p values and mark the frames accordingly:

Step	Task	Command Syntax	Command Mode
1	Allocate CAM space to enable queuing frames according to the C-Tag or the S-Tag. **vman-qos**: mark the S-Tag dot1p and queue the frame according to the original C-Tag dot1p. This method requires half as many CAM entries as **vman-qos-dual-fp**. **vman-qos-dual-fp**: mark the S-Tag dot1p and queue the frame according to the S-Tag dot1p. This method requires twice as many CAM entries as **vman-qos** and FP blocks in multiples of 2.	**cam-acl l2acl** *number* **ipv4acl** *number* **ipv6acl** *number* **ipv4qos** *number* **l2qos** *number* **l2pt** *number* **ipmacacl** *number* **ecfmacl** *number* {**vman-qos** | **vman-qos-dual-fp**} *number* Default: 0 FP blocks for vman-qos and vman-qos-dual-fp	CONFIGURATION
2	The new CAM configuration is stored in NVRAM and takes effect only after a save and reload.	**copy running-config startup-config** **reload**	EXEC Privilege
3	Map C-Tag dot1p values to a S-Tag dot1p value. C-Tag values may be separated by commas, and dashed ranges are permitted. Dynamic Mode CoS overrides any Layer 2 QoS configuration in case of conflicts.	**vlan-stack dot1p-mapping c-tag-dot1p** *values* **sp-tag-dot1p** *value*	INTERFACE

Note: Since **dot1p-mapping** marks *and* queues packets, the only remaining applicable QoS configuration is rate metering. You may use Rate Shaping or Rate Policing.

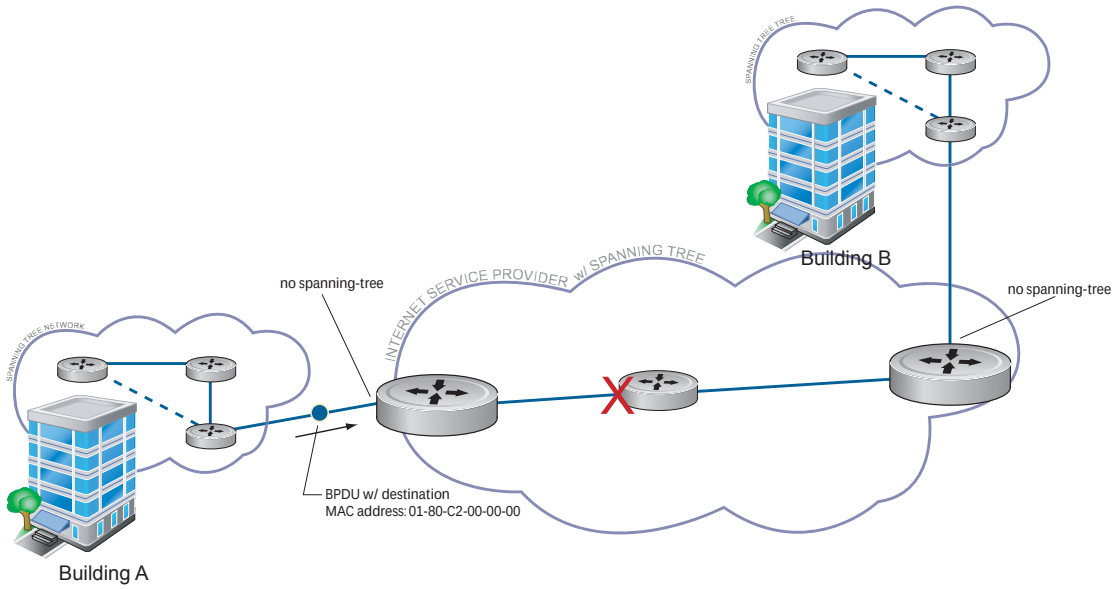
Layer 2 Protocol Tunneling

Layer 2 Protocol Tunneling (L2PT) is supported on platforms: C E S

L2PT is supported on E-Series ExaScale E X with FTOS 8.2.1.0. and later.

Spanning Tree BPDUs use a reserved destination MAC address called the Bridge Group Address, which is 01-80-C2-00-00-00. Only spanning-tree bridges on the LAN recognize this address and process the BPDU. When VLAN stacking is used to connect physically separate regions of a network, BPDUs attempting to traverse the intermediate network might be consumed and subsequently dropped because the intermediate network itself might be using Spanning Tree (Figure 46-13).

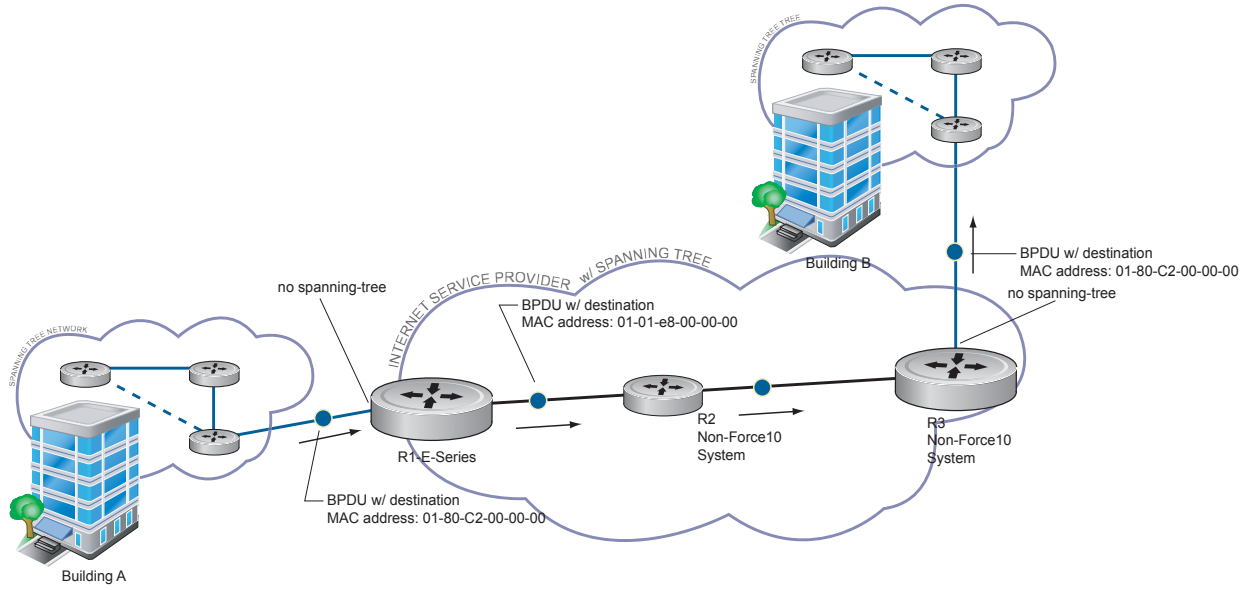
Figure 46-13. VLAN Stacking without L2PT



You might need to transport control traffic transparently through the intermediate network to the other region. Layer 2 Protocol Tunneling enables BPDUs to traverse the intermediate network by identifying frames with the Bridge Group Address, rewriting the destination MAC to a user-configured non-reserved address, and forwarding the frames. Since the frames now use a unique MAC address, BPDUs are treated as normal data frames by the switches in the intermediate network core. On egress edge of the intermediate network, the MAC address rewritten to the original MAC address and forwarded to the opposing network region (Figure 46-14).

FTOS Behavior: In FTOS versions prior to 8.2.1.0, the MAC address that Dell Force10 systems use to overwrite the Bridge Group Address on ingress was non-configurable. The value of the L2PT MAC address was the Force10-unique MAC address, 01-01-e8-00-00-00. As such, with these FTOS versions, Dell Force10 systems are required at the egress edge of the intermediate network because only FTOS could recognize the significance of the destination MAC address and rewrite it to the original Bridge Group Address. In FTOS version 8.2.1.0 and later, the L2PT MAC address is user-configurable, so you can specify an address that non-Dell Force10 systems can recognize and rewrite the address at egress edge.

Figure 46-14. VLAN Stacking with L2PT



Implementation Information

- L2PT is available for STP, RSTP, MSTP, and PVST+ BPDUs.
- No protocol packets are tunneled when VLAN Stacking is enabled.
- L2PT requires the default CAM profile.

Enable Layer 2 Protocol Tunneling

Step	Task	Command Syntax	Command Mode
1	Verify that the system is running the default CAM profile; you must use this CAM profile for L2PT.	**show cam-profile**	EXEC Privilege
2	Enable protocol tunneling globally on the system.	**protocol-tunnel enable**	CONFIGURATION
3	Tunnel BPDUs the VLAN.	**protocol-tunnel stp**	INTERFACE VLAN

Specify a Destination MAC Address for BPDUs

By default, FTOS uses a Force10-unique MAC address for tunneling BPDUs. You can configure another value.

Task	Command Syntax	Command Mode
Overwrite the BPDU with a user-specified destination MAC address when BPDUs are tunneled across the provider network. Default: 01:01:e8:00:00:00	**protocol-tunnel destination-mac**	CONFIGURATION

Rate-limit BPDUs on the E-Series

In order to rewrite the destination MAC address on BPDUs, they are forwarded to the RPM. You can rate-limit BPDUs to protect the RPM, in which case the system drops BPDUs when the threshold is reached.

Task	Command Syntax	Command Mode
Set a maximum rate at which the RPM will process BPDUs for L2PT. Default: 75 pps E-Series Range: 75 to 3000 pps	**protocol-tunnel rate-limit**	CONFIGURATION

Rate-limit BPDUs on the C-Series and S-Series

CAM space is allocated in sections called Field Processor (FP) blocks.

There are total 13 user-configurable FP blocks on the C-Series and S-Series. The default number of blocks for L2PT is 0; you must allocate at least one to enable BPDU rate-limiting.

Step	Task	Command Syntax	Command Mode
1	Create at least one FP group for L2PT. See CAM Allocation on page 289 for details on this command.	**cam-acl l2acl**	CONFIGURATION
2	Save the running-config to the startup-config.	**copy running-config startup-config**	EXEC Privilege
3	Reload the system.	**reload**	EXEC Privilege
4	Set a maximum rate at which the RPM will process BPDUs for L2PT. Default: no rate limiting C-Series Range: 64 to 640 kbps S-Series Range: 64 to 320 kbps	**protocol-tunnel rate-limit**	VLAN STACKING

Debug Layer 2 Protocol Tunneling

Task	Command Syntax	Command Mode
Display debugging information for L2PT.	**debug protocol-tunnel**	EXEC Privilege

Provider Backbone Bridging

Provider Backbone Bridging is supported only on platforms: C S

IEEE 802.1ad—Provider Bridges amends 802.1Q—Virtual Bridged Local Area Networks so that service providers can use 802.1Q architecture to offer separate VLANs to customers with no coordination between customers, and minimal coordination between customers and the provider.

802.1ad specifies that provider bridges operating Spanning Tree use a reserved destination MAC address called the Provider Bridge Group Address, 01-80-C2-00-00-08, to exchange BPDUs instead of the Bridge Group Address, 01-80-C2-00-00-00, originally specified in 802.1Q. Only bridges in the service provider network use this destination MAC address so these bridges treat BPDUs originating from the customer network as normal data frames, rather than consuming them.

The same is true for GVRP. 802.1ad specifies that provider bridges participating in GVRP use a reserved destination MAC address called the Provider Bridge GVRP Address, 01-80-C2-00-00-0D, to exchange GARP PDUs instead of the GVRP Address, 01-80-C2-00-00-21, specified in 802.1Q. Only bridges in the service provider network use this destination MAC address so these bridges treat GARP PDUs originating from the customer network as normal data frames, rather than consuming them.

Provider Backbone Bridging through IEEE 802.1ad eliminates the need for tunneling BPDUs with L2PT and increases the reliability of provider bridge networks as the network core need only learn the MAC addresses of core switches, as opposed to all MAC addresses received from attached customer devices.

Task	Command Syntax	Command Mode
Use the Provider Bridge Group address as the destination MAC address in BPDUs. The xstp keyword applies this functionality to STP, RSTP, and MSTP; this functionality is not available for PVST+.	**bpdu-destination-mac-address** [**stp** | **gvrp**] **provider-bridge-group**	CONFIGURATION

47

sFlow

sFlow is supported on platforms C E S

sFlow is supported on E-Series ExaScale E X with FTOS 8.1.1.0. and later.

Overview

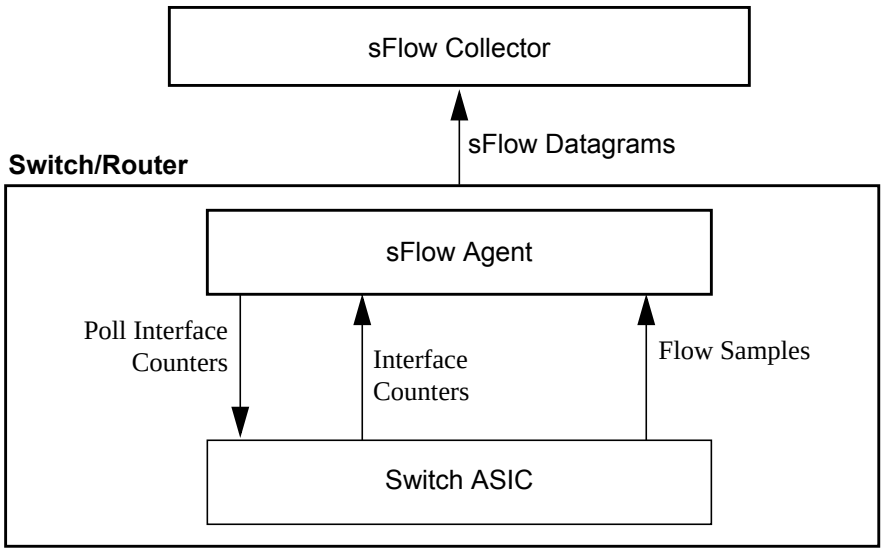
FTOS supports sFlow version 5. sFlow is a standard-based sampling technology embedded within switches and routers which is used to monitor network traffic. It is designed to provide traffic monitoring for high speed networks with many switches and routers. sFlow uses two types of sampling:

- Statistical packet-based sampling of switched or routed packet flows
- Time-based sampling of interface counters

The sFlow monitoring system consists of an sFlow Agent (embedded in the switch/router) and an sFlow collector. The sFlow Agent resides anywhere within the path of the packet, and combines the flow samples and interface counters into sFlow datagrams and forwards them to the sFlow Collector at regular intervals. The datagrams consists of information on, but not limited to, packet header, ingress and egress interfaces, sampling parameters, and interface counters.

Packet sampling is typically done by the ASIC. sFlow Collector analyses the sFlow datagrams received from different devices and produces a network-wide view of traffic flows.

Figure 47-1. sFlow Traffic Monitoring System



Implementation Information

Dell Force10 sFlow is designed so that the hardware sampling rate is per line card port-pipe and is decided based upon all the ports in that port-pipe. If sFlow is not enabled on any port specifically, then the global sampling rate is downloaded to that port and is to calculate the port-pipe's lowest sampling rate. This design supports, then, the possibility that sFlow might be configured on that port in the future. Back-off is triggered based on the port-pipe's hardware sampling rate.

For example, if port 1 in a the port-pipe has sFlow configured with a 16384 sampling rate while port 2 in the port-pipe has sFlow configured but no sampling rate set, FTOS applies a global sampling rate of 512 to port 2. The hardware sampling rate on the port-pipe is ten set at 512 because that is the lowest configured rate on the port-pipe. When a high traffic situation occurs, a back-off is triggered and the hardware sampling rate is backed-off from 512 to 1024. Note that port 1 maintains its sampling rate of 16384; port 1 is unaffected because it maintains its configured sampling rate of 16484.

To avoid the back-off, either increase the global sampling rate or configure all the line card ports with the desired sampling rate even if some ports have no sFlow configured.

Important Points to Remember

- The FTOS implementation of the sFlow MIB supports sFlow configuration via **snmpset**.
- Collection through management interface is supported on E-Series only
- Dell Force10 recommends that the sFlow Collector be connected to the Dell Force10 chassis through a line card port rather than the RPM Management Ethernet port.
- E-Series TeraScale sFlow sampling is done on a per-port-pipe basis.
- E-Series ExaScale, C-Series, and S-Series sFlow sampling is done on a per-port basis.

- FTOS exports all sFlow packets to the collector. A small sampling rate can equate to a large number of exported packets. A backoff mechanism will automatically be applied to reduce this amount. Some sampled packets may be dropped when the exported packet rate is high and the backoff mechanism is about to or is starting to take effect. The dropEvent counter, in the sFlow packet, will always be zero.
- Community list and local preference fields are not filled in extended gateway element in sFlow datagram.
- 802.1P source priority field is not filled in extended switch element in sFlow datagram.
- Only Destination and Destination Peer AS number are packed in the dst-as-path field in extended gateway element
- If packet being sampled is redirected using PBR (Policy-Based Routing), sFlow datagram may contain incorrect extended gateway/router information.
- Source VLAN field in the extended switch element will not be packed in case of routed packet.
- Destination VLAN field in the extended switch element will not be packed in case of Multicast packet.
- On the C-Series, Layer 3 and Layer 2 multicast traffic is not collected with sFlow sampling.
- On the S-Series, up to 700 packets can be sampled and processed per second.
- On the C-Series up to 1000 packets can be sampled and processed per second.
- On the E-Series, the maximum number of packets that can be sampled and processed per second is:
    - 7500 packets when no extended information packing is enabled.
    - 1000 packets when only extended-switch information packing is enabled.
    - 1600 packets when extended-router and/or extended-gateway information packing is enabled.

Enable and Disable sFlow

By default, sFlow is *disabled* globally on the system. To enable sFlow globally, use the **sflow enable** command in CONFIGURATION mode. Use the **no** version of this command to disable sFlow globally.

Command Syntax	Command Mode	Usage
[**no**] **sflow enable**	CONFIGURATION	Enable sFlow globally.

Enable and Disable on an Interface

By default, sFlow is *disabled* on all interfaces. To enable sFlow on a specific interface, use the **sflow enable** command in INTERFACE mode. Use the **no** version of this command to disable sFlow on an interface. This CLI is supported on physical ports and LAG ports.

Command Syntax	Command Mode	Usage
[**no**] **sflow enable**	INTERFACE	Enable sFlow on an interface.

sFlow Show Commands

FTOS includes the following sFlow display commands:

Show sFlow Globally

Use the following command to view sFlow statistics:

Command Syntax	Command Mode	Purpose
show sflow	EXEC	Display sFlow configuration information and statistics.

Figure 47-2 is a sample output from the **show sflow** command:

Figure 47-2. Command Example: show sflow

```
FTOS#show sflow
sFlow services are enabled                          <-- Indicates sFlow is globally enabled
Global default sampling rate: 32768
Global default counter polling interval: 20
1 collectors configured
Collector IP addr: 133.33.33.53, Agent IP addr: 133.33.33.116, UDP port: 6343
77 UDP packets exported
0 UDP packets dropped
165 sFlow samples collected
69 sFlow samples dropped due to sub-sampling

Linecard 1 Port set 0 H/W sampling rate 8192     <-- Indicates sFlow is enabled on linecards Gi 1/16 and Gi 1/17
  Gi 1/16: configured rate 8192, actual rate 8192, sub-sampling rate 1
  Gi 1/17: configured rate 16384, actual rate 16384, sub-sampling rate 2
```

Show sFlow on an Interface

Use the following command to view sFlow information on a specific interface:

Command Syntax	Command Mode	Purpose
show sflow interface *interface-name*	EXEC	Display sFlow configuration information and statistics on a specific interface.

Figure 47-3 is a sample output from the **show sflow interface** command.

Figure 47-3. Command Example: show sflow interface

```
FTOS#show sflow interface gigabitethernet 1/16
Gi 1/16
Configured sampling rate         :8192
Actual sampling rate             :8192
Sub-sampling rate                :2
Counter polling interval         :15
Samples rcvd from h/w            :33
Samples dropped for sub-sampling :6
```

The configuration, shown in Figure 47-2, is also displayed in the running configuration (Figure 47-4):

Figure 47-4. Command Example: show running-config interface

```
FTOS#show running-config interface gigabitethernet 1/16
!
interface GigabitEthernet 1/16
 no ip address
 mtu 9252
 ip mtu 9234
 switchport
 sflow enable
 sflow sample-rate 8192
 no shutdown
```

Show sFlow on a Line Card

Use the following command to view sFlow statistitics on a specified line card:

Command Syntax	Command Mode	Purpose
show sflow linecard *slot-number*	EXEC	Display sFlow configuration information and statistics on the specified interface.

Figure 47-5 is a sample output from the **show sflow linecard** command:

Figure 47-5. Command Example: show sflow linecard

```
FTOS#show sflow linecard 1
Linecard 1
  Samples rcvd from h/w            :165
  Samples dropped for sub-sampling :69
  Total UDP packets exported       :77
  UDP packets exported via RPM     :77
  UDP packets dropped              :
```

Configure Collectors

The **sflow collector** command allows you to configure sFlow collectors to which sFlow datagrams are forwarded. You can configure up to two sFlow collectors (IPv4 or IPv6). If you configure two collectors, traffic samples are sent to both devices.

sFlow collection through the Management interface is supported on platform: E.

IPv6 sFlow collectors and agents are supported on platforms: C E S.

Command Syntax	Command Mode	Usage
sflow collector {*ipv4-address* | *ipv6-address*} **agent-addr** {*ipv4-address* | *ipv6-address*} [*number* [**max-datagram-size** *number*]] | [**max-datagram-size** *number*]	CONFIGURATION	Configure an sFlow agent in the router and an sFlow collector to which sFlow datagrams are forwarded. Default UDP port: 6343 Default max-datagram-size: 1400

Polling Intervals

The **sflow polling-interval** command configures the polling interval for an interface in the maximum number of seconds between successive samples of counters to be sent to the collector. This command changes the global default counter polling (20 seconds) interval. You can configure an interface to use a different polling interval.

The polling interval can be configured globally (in CONFIGURATION mode) or by interface (in INTERFACE mode) by executing the interval command:

Command Syntax	Command Mode	Usage
sflow polling-interval *interval value*	CONFIGURATION or INTERFACE	Change the global default counter polling interval. *interval value*—in seconds. Range: 15 to 86400 seconds Default: 20 seconds

Sampling Rate

Sampling Rate is supported on platform E T.

The sFlow sampling rate is the number of packets that are skipped before the next sample is taken. sFlow does not have time-based packet sampling.

The **sflow sample-rate** command, when issued in CONFIGURATION mode, changes the default sampling rate. By default, the sampling rate of an interface is set to the same value as the current global default sampling rate.If the value entered is not a correct power of 2, the command generates an error message with the previous and next power-of-2 value. Select one of these two number and re-enter the command. (For more information on values in power-of-2, see Sub-sampling on page 979.)

The sample rate can be configured globally or by interface using the sample rate command:

Command Syntax	Command Mode	Usage
[**no**] **sflow sample-rate** *sample-rate*	CONFIGURATION or INTERFACE	Change the global or interface sampling rate. Rate must be entered in factors of 2 (eg, 4096, 8192). *sample-rate* range: 256-8388608 for C-Series and S-Series 2-8388608 for E-Series

Sub-sampling

Sub-sampling is available only on platform: E T

The sFlow sample rate is not the frequency of sampling, but the number of packets that are skipped before the next sample is taken. Although a sampling rate can be configured for each port, TeraScale line cards can support only a single sampling rate per port-pipe.

Therefore, sFlow Agent uses sub-sampling to create multiple sampling rates per port-pipe. To achieve different sampling rates for different ports in a port-pipe, sFlow Agent takes the lowest numerical value of the sampling rate of all the ports within the port-pipe, and configures all ports to this value. sFlow Agent is then able to skip samples on ports where you require a larger sampling rate value.

Sampling rates are configurable in powers of two. This allows the smallest sampling rate possible to be configured on the hardware, and also allows all other sampling rates to be available through sub-sampling.

For example, if Gig 1/0 and 1/1 are in a port-pipe, and they are configured with a sampling rate of 4096 on interface Gig 1/0, and 8192 on Gig 1/1, sFlow Agent does the following:

1. Configures the hardware to a sampling rate of 4096 for all ports with sFlow enabled on that port-pipe.
2. Configure interface Gig 1/0 to a sub-sampling rate of 1 to achieve an actual rate of 4096.
3. Configure interface Gig 1/1 to a sub-sampling rate of 2 to achieve an actual rate of 8192.

Note: Sampling rate backoff can change the sampling rate value that is set in the hardware. This equation shows the relationship between actual sampling rate, sub-sampling rate, and the hardware sampling rate for an interface:

*Actual sampling rate = sub-sampling rate * hardware sampling rate*

Note the absence of a configured rate in the equation. That is because when the hardware sampling rate value on the port-pipe exceeds the configured sampling rate value for an interface, the actual rate changes to the hardware rate. The sub-sampling rate never goes below a value of one.

Back-off Mechanism

If the sampling rate for an interface is set to a very low value, the CPU can get overloaded with flow samples under high-traffic conditions. In such a scenario, a binary back-off mechanism gets triggered, which doubles the sampling-rate (halves the number of samples per second) for all interfaces. The backoff mechanism continues to double the sampling-rate until CPU condition is cleared. This is as per sFlow version 5 draft. Once the back-off changes the sample-rate, users must manually change the sampling rate to the desired value.

As a result of back-off, the actual sampling-rate of an interface may differ from its configured sampling rate. The actual sampling-rate of the interface and the configured sample-rate can be viewed by using the **show sflow** command.

sFlow on LAG ports

When a physical port becomes a member of a LAG, it inherits the sFlow configuration from the LAG port.

Extended sFlow

Extended sFlow is supported fully on platform E

Platforms C and S support **extended-switch** information processing *only*.

Extended sFlow packs additional information in the sFlow datagram depending on the type of sampled packet. The following options can be enabled:

- **extended-switch** — 802.1Q VLAN ID and 802.1p priority information
- **extended-router** — Next-hop and source and destination mask length.
- **extended-gateway** — Source and destination AS number and the BGP next-hop.

Note: The entire AS path is not included. BGP community-list and local preference information are not included. These fields are assigned default values and are not interpreted by the collector.

Use the command **sflow** [**extended-switch**] [**extended-router**] [**extended-gateway**] **enable** command. By default packing of any of the extended information in the datagram is disabled.

Use the command **show sflow** to confirm that extended information packing is enabled, as shown in Figure 47-6.

Figure 47-6. Confirming that Extended sFlow is Enabled

```
FTOS#show sflow
sFlow services are enabled
Global default sampling rate: 4096
Global default counter polling interval: 15
Global extended information enabled: gateway, router, switch
1 collectors configured
Collector IP addr: 10.10.10.3, Agent IP addr: 10.10.0.0, UDP port: 6343
77 UDP packets exported
0 UDP packets dropped
165 sFlow samples collected
69 sFlow samples dropped due to sub-sampling

Linecard 1 Port set 0 H/W sampling rate 8192
  Gi 1/16: configured rate 8192, actual rate 8192, sub-sampling rate 1
  Gi 1/17: configured rate 16384, actual rate 16384, sub-sampling rate 2

Linecard 3 Port set 1 H/W sampling rate 16384
  Gi 3/40: configured rate 16384, actual rate 16384, sub-sampling rate 1
```

Extended sFlow settings show all 3 types are enabled

If none of the extended information is enabled, the **show** output is as shown in Figure 47-7.

Figure 47-7. Confirming that Extended sFlow is Disabled

```
FTOS#show sflow
sFlow services are disabled
Global default sampling rate: 32768
Global default counter polling interval: 20
Global extended information enabled: none
0 collectors configured
0 UDP packets exported
0 UDP packets dropped
0 sFlow samples collected
0 sFlow samples dropped due to sub-sampling
```

Indicates no Extended sFlow types enabled.

Important Points to Remember

- The IP destination address has to be learned via BGP in order to export extended-gateway data, prior to FTOS version 7.8.1.0.
- If the IP destination address is not learned via BGP the Dell Force10 system does not export extended-gateway data, prior to FTOS version 7.8.1.0.
- FTOS 7.8.1.0 and later enhances the sFlow implementation for real time traffic analysis on the E-Series to provide extended gateway information in cases where the destination IP addresses are learned by different routing protocols, and for cases where the destination is reachable over ECMP.
- If the IP source address is learned via IGP then *srcAS* and *srcPeerAS* are zero.
- The srcAS and srcPeerAS might be zero even though the IP source address is learned via BGP. The Dell Force10 system packs the srcAS and srcPeerAS information only if the route is learned via BGP and it is reachable via the ingress interface of the packet.

The previous points are summarized in following table.

Table 47-1. Extended Gateway Summary

IP SA	IP DA	srcAS and srcPeerAS	dstAS and dstPeerAS	Description
static/connected/IGP	static/connected/IGP	—	—	Extended gateway data is not exported because there is no AS information.
static/connected/IGP	BGP	0	Exported	src_as & src_peer_as are zero because there is no AS information for IGP.
BGP	static/connected/IGP	—	—	Prior to FTOS version 7.8.1.0, extended gateway data is not be exported because IP DA is not learned via BGP.
		Exported	Exported	7.8.1.0 allows extended gateway information in cases where the source and destination IP addresses are learned by different routing protocols, and for cases where is source is reachable over ECMP.
BGP	BGP	Exported	Exported	Extended gateway data is packed.

48

Simple Network Management Protocol

Simple Network Management Protocol is supported on platforms C E S

SNMP is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

Note: On Dell Force10 routers, standard and private SNMP MIBs are supported, including all Get and a limited number of Set operations (such as **set vlan** and **copy cmd**).

Protocol Overview

Network management stations use Simple Network Management Protocol (SNMP) to retrieve or alter management data from network elements. A datum of management information is called a *managed object*; the value of a managed object can be static or variable. Network elements store managed objects in a database called a *Management Information Base* (MIB).

MIBs are hierarchically structured and use *object identifiers* to address managed objects, but managed objects also have a textual name called an *object descriptor*.

Implementation Information

- FTOS supports SNMP version 1 as defined by RFC 1155, 1157, and 1212, SNMP version 2c as defined by RFC 1901, and SNMP version 3 as defined by RFC 2571.
- FTOS supports up to 15 trap receivers.
- The FTOS implementation of the sFlow MIB supports sFlow configuration via SNMP sets.
- SNMP traps for STP and MSTP state changes are based on BRIDGE MIB (RFC 1483) for STP and IEEE 802.1 *draft ruzin-mstp-mib-02* for MSTP.

Configure Simple Network Management Protocol

Note: The configurations in this chapter use a Unix environment with net-snmp version 5.4. This is only one of many RFC-compliant SNMP utilities you can use to manage your Dell Force10 system using SNMP. Also, these configurations use SNMP version 2c.

Configuring SNMP requires only a single step:

1. Create a community. See page 984.

Related Configuration Tasks

The following list contains configuration tasks for SNMP:

- Read Managed Object Values on page 985
- Write Managed Object Values on page 986
- Subscribe to Managed Object Value Updates using SNMP on page 988
- Copy Configuration Files on page 113
- Manage VLANs using SNMP on page 997
- Enable and Disable a Port using SNMP on page 1001
- Fetch Dynamic MAC Entries using SNMP on page 1001
- Deriving Interface Indices on page 1003
- Monitor Port-channels on page 1004

Important Points to Remember

- Typically, 5-second timeout and 3-second retry values on an SNMP server are sufficient for both LAN and WAN applications. If you experience a timeout with these values, increase the timeout value to greater than 3 seconds, and increase the retry value to greater than 2 on your SNMP server.
- Group ACLs override user ACLs in SNMPv3 configurations when both are configured and the user is part of the group.

Create a Community

The management station generates requests to either retrieve or alter the value of a management object and is called the *SNMP manage*r. A network element that processes SNMP requests is called an *SNMP agent*. An *SNMP community* is a group of SNMP agents and managers that are allowed to interact. Communities are necessary to secure communication between SNMP managers and agents; SNMP agents do not respond to requests from management stations that are not part of the community.

FTOS enables SNMP automatically when you create an SNMP community and displays Message 1. You must specify whether members of the community may only retrieve values (read), or retrieve and alter values (read-write).

To create an SNMP community:

Task	Command	Command Mode
Choose a name for the community.	**snmp-server community** *name* {**ro** | **rw**}	CONFIGURATION

Message 1 SNMP Enabled

```
22:31:23: %RPM1-P:CP %SNMP-6-SNMP_WARM_START: Agent Initialized - SNMP WARM_START.
```

View your SNMP configuration, using the command **show running-config snmp** from EXEC Privilege mode, as shown in Figure 48-1.

Figure 48-1. Creating an SNMP Community

```
FTOS#snmp-server community my-snmp-community ro
22:31:23: %RPM1-P:CP %SNMP-6-SNMP_WARM_START: Agent Initialized - SNMP WARM_START.
FTOS#do show running-config snmp
!
snmp-server community mycommunity ro
FTOS#
```

Read Managed Object Values

You may only retrieve (read) managed object values if your management station is a member of the same community as the SNMP agent.

There are several Unix SNMP commands that read data:

Task	Command
Read the value of a single managed object, as shown in Figure 48-2.	**snmpget -v** *version* **-c** *community agent-ip* {*identifier.instance* | *descriptor.instance*}

Figure 48-2. Reading the Value of a Managed Object

```
> snmpget -v 2c -c mycommunity 10.11.131.161 sysUpTime.0
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (32852616) 3 days, 19:15:26.16
> snmpget -v 2c -c mycommunity 10.11.131.161 .1.3.6.1.2.1.1.3.0
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (32856932) 3 days, 19:16:09.32
```

Read the value of the managed object directly below the specified object, as shown in Figure 48-3.	**snmpgetnext -v** *version* **-c** *community agent-ip* {*identifier.instance* | *descriptor.instance*}

Figure 48-3. Reading the Value of the Next Managed Object in the MIB

```
> snmpgetnext -v 2c -c mycommunity 10.11.131.161 .1.3.6.1.2.1.1.3.0
SNMPv2-MIB::sysContact.0 = STRING:
> snmpgetnext -v 2c -c mycommunity 10.11.131.161 sysContact.0
SNMPv2-MIB::sysName.0 = STRING: S50V_7.7
```

Read the value of many objects at once, as shown in Figure 48-4.	**snmpwalk -v** *version* **-c** *community agent-ip* {*identifier.instance* | *descriptor.instance*}

Task	Command

Figure 48-4. Reading the Value of Many Managed Objects at Once

```
> snmpwalk -v 2c -c mycommunity 10.11.131.161 .1.3.6.1.2.1.1
SNMPv2-MIB::sysDescr.0 = STRING: Force10 Networks Real Time Operating System Software
Force10 Operating System Version: 1.0
Force10 Application Software Version: E_MAIN4.7.6.350
Copyright (c) 1999-2007 by Force10 Networks, Inc.
Build Time: Mon May 12 14:02:22 PDT 2008
SNMPv2-MIB::sysObjectID.0 = OID: SNMPv2-SMI::enterprises.6027.1.3.1
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (32920954) 3 days, 19:26:49.54
SNMPv2-MIB::sysContact.0 = STRING:
SNMPv2-MIB::sysName.0 = STRING: S50V_7.7
SNMPv2-MIB::sysLocation.0 = STRING:
SNMPv2-MIB::sysServices.0 = INTEGER: 4
```

Write Managed Object Values

You may only alter (write) a managed object value if your management station is a member of the same community as the SNMP agent, and the object is writable.

To write or write-over the value of a managed object:

Task	Command
To write or write-over the value of a managed object, as shown in Figure 48-5.	**snmpset -v** *version* **-c** *community agent-ip* {*identifier.instance* | *descriptor.instance*}

Figure 48-5. Writing over the Current Value of a Managed Object

```
> snmpset -v 2c -c mycommunity 10.11.131.161 sysName.0 s "R5"
SNMPv2-MIB::sysName.0 = STRING: R5
```

Configure Contact and Location Information using SNMP

You may configure system contact and location information from the Dell Force10 system or from the management station using SNMP.

To configure system contact and location information from the Dell Force10 system:

Task	Command	Command Mode
Identify the system manager along with this person's contact information (e.g E-mail address or phone number). You may use up to 55 characters. **Default**: None	**snmp-server contact** *text*	CONFIGURATION
Identify the physical location of the system. For example, San Jose, 350 Holger Way, 1st floor lab, rack A1-1. You may use up to 55 characters. **Default**: None	**snmp-server location** *text*	CONFIGURATION

To configure the system from the manumitting station using SNMP:

Task	Command	Command Mode
Identify the system manager along with this person's contact information (e.g E-mail address or phone number). You may use up to 55 characters. **Default**: None	**snmpset -v** *version* **-c** *community agent-ip* ***sysContact.0*** **s "***contact-info***"**	CONFIGURATION
Identify the physical location of the system. For example, San Jose, 350 Holger Way, 1st floor lab, rack A1-1. You may use up to 55 characters. **Default**: None	**snmpset -v** *version* **-c** *community agent-ip* ***sysLocation.0*** **s "***location-info***"**	CONFIGURATION

Subscribe to Managed Object Value Updates using SNMP

By default, the Dell Force10 system displays some unsolicited SNMP messages (traps) upon certain events and conditions. You can also configure the system to send the traps to a management station. Traps cannot be saved on the system.

FTOS supports the following three sets of traps:

- **RFC 1157-defined traps**: coldStart, warmStart, linkDown, linkUp, authenticationFailure, egpNeighbborLoss
- **Dell Force10 enterpriseSpecific environment traps**: fan, supply, temperature
- **Dell Force10 enterpriseSpecific protocol traps**: bgp, ecfm, stp, vrrp, xstp,

To configure the system to send SNMP notifications:

Step	Task	Command	Command Mode
1	Configure the Dell Force10 system send notifications to an SNMP server.	**snmp-server host** *ip-address*	CONFIGURATION
2	Specify which traps the Dell Force10 system sends to the trap receiver. • Enable all Dell Force10 enterpriseSpecific and RFC-defined traps using the command **snmp-server enable traps** from CONFIGURATION mode. • Enable all of the RFC-defined traps using the command **snmp-server enable traps snmp** from CONFIGURATION mode.	**snmp-server enable traps**	CONFIGURATION
3	Specify the interfaces out of which FTOS sends SNMP traps.	**snmp-server trap-source**	CONFIGURATION

Table 48-1 lists the traps the RFC-defined SNMP traps and the command used to enable each. Note that the coldStart and warmStart traps are enabled using a single command.

Table 48-1. RFC 1157 Defined SNMP Traps on FTOS

Command Option	Trap
snmp authentication	`SNMP_AUTH_FAIL:SNMP Authentication failed.Request with invalid community string.`
snmp coldstart	`SNMP_COLD_START: Agent Initialized - SNMP COLD_START.` `SNMP_WARM_START: Agent Initialized - SNMP WARM_START.`
snmp linkdown	`PORT_LINKDN:changed interface state to down:%d`
snmp linkup	`PORT_LINKUP:changed interface state to up:%d`

To enable a subset of Dell Force10 enterprise-specific SNMP traps, enter any of the **snmp-server enable traps** command options in Table 48-2. Note that the **envmon** option enables all environment traps, including **fan**, **supply**, and **temperature**.

Table 48-2. Dell Force10 Enterprise-specific SNMP Traps

Command Option	Trap Examples
envmon	CARD_SHUTDOWN: %sLine card %d down - %s CARD_DOWN: %sLine card %d down - %s
	LINECARDUP: %sLine card %d is up
	CARD_MISMATCH: Mismatch: line card %d is type %s - type %s required.
	RPM_STATE: RPM1 is in Active State RPM_STATE: RPM0 is in Standby State
	RPM_DOWN: RPM 0 down - hard reset RPM_DOWN: RPM 0 down - card removed
	HOT_FAILOVER: RPM Failover Completed
	SFM_DISCOVERY: Found SFM 1
	SFM_REMOVE: Removed SFM 1
	MAJOR_SFM: Major alarm: Switch fabric down
	MAJOR_SFM_CLR: Major alarm cleared: Switch fabric up
	MINOR_SFM: MInor alarm: No working standby SFM
	MINOR_SFM_CLR: Minor alarm cleared: Working standby SFM present
	TASK SUSPENDED: SUSPENDED - svce:%d - inst:%d - task:%s
	RPM0-P:CP %CHMGR-2-CARD_PARITY_ERR
	ABNORMAL_TASK_TERMINATION: CRASH - task:%s %s
	CPU_THRESHOLD: Cpu %s usage above threshold. Cpu5SecUsage (%d)
	CPU_THRESHOLD_CLR: Cpu %s usage drops below threshold. Cpu5SecUsage (%d)
	MEM_THRESHOLD: Memory %s usage above threshold. MemUsage (%d)
	MEM_THRESHOLD_CLR: Memory %s usage drops below threshold. MemUsage (%d)
	DETECT_STN_MOVE: Station Move threshold exceeded for Mac %s in vlan %d
envmon supply	PEM_PRBLM: Major alarm: problem with power entry module %s
	PEM_OK: Major alarm cleared: power entry module %s is good
	MAJOR_PS: Major alarm: insufficient power %s
	MAJOR_PS_CLR: major alarm cleared: sufficient power
	MINOR_PS: Minor alarm: power supply non-redundant
	MINOR_PS_CLR: Minor alarm cleared: power supply redundant
envmon temperature	MINOR_TEMP: Minor alarm: chassis temperature
	MINOR_TEMP_CLR: Minor alarm cleared: chassis temperature normal (%s %d temperature is within threshold of %dC)
	MAJOR_TEMP: Major alarm: chassis temperature high (%s temperature reaches or exceeds threshold of %dC)
	MAJOR_TEMP_CLR: Major alarm cleared: chassis temperature lower (%s %d temperature is within threshold of %dC)
envmon fan	FAN_TRAY_BAD: Major alarm: fantray %d is missing or down
	FAN_TRAY_OK: Major alarm cleared: fan tray %d present
	FAN_BAD: Minor alarm: some fans in fan tray %d are down
	FAN_OK: Minor alarm cleared: all fans in fan tray %d are good

Table 48-2. Dell Force10 Enterprise-specific SNMP Traps

Command Option	Trap Examples
stp (includes only STP notifications)	%SPANMGR-5-STP_ROOT_CHANGE: STP root changed for vlan 1. My Bridge ID: 0:0001.e867.b1f8 Old Root: 0:0000.0000.0000 New Root: 0:0001.e867.b1f8. %SPANMGR-5-STP_NEW_ROOT: New Spanning Tree Root, Bridge ID Priority 32768, Address 0001.e801.fc35. %SPANMGR-5-STP_TOPOLOGY_CHANGE: Bridge port GigabitEthernet 11/38 transitioned from forwarding to discarding state. %SPANMGR-5-STP_TOPOLOGY_CHANGE: Topology change:BridgeAddr: 0001.e867.b1f8Bridge port Gi 3/19 transitioned from learning to forwarding state. %SPANMGR-5-BPDU_GUARD_RX_ERROR: Received Spanning Tree BPDU on BPDU guard port. Disable Port-channel 1.
xstp (includes MSTP, RSTP, and PVST+ notifications)	%SPANMGR-5-MSTP_NEW_ROOT_BRIDGE: Elected root bridge for instance 0. %SPANMGR-5-MSTP_NEW_ROOT_PORT: MSTP root changed to port Gi 11/38 for instance 0. My Bridge ID: 40960:0001.e801.fc35 Old Root: 40960:0001.e801.fc35 New Root: 32768:00d0:038a.2c01. %SPANMGR-5-MSTP_TOPOLOGY_CHANGE: Topology change BridgeAddr: 0001.e801.fc35 Mstp Instance Id 0 port Gi 11/38 transitioned from forwarding to discarding state. %SPANMGR-5-PVST_NEW_ROOT: Elected new root for instance 1. My Bridge ID: 0:0001.e867.b1f8 Old Root: 0:0000.0000.0000 New Root: 0:0001.e867.b1f8. %SPANMGR-5-PVST_TOPOLOGY_CHANGE: Topology change BridgeAddr: 0001.e867.b1f8 Pvst Instance Id: 1111 Bridge port Po 1 transitioned from forwarding to discarding state. %SPANMGR-5-PVST_TOPOLOGY_CHANGE: Topology change BridgeAddr: 0001.e867.b1f8 Pvst Instance Id: 1 Bridge port Po 1 transitioned from learning to forwarding state. %SPANMGR-5-RSTP_NEW_ROOT: New Rapid Spanning Tree Root. My Bridge Id: 0:0001.e867.b1f8 Old Root: 32768:0001.e867.b1f8 New Root: 0:0001.e867.b1f8. %SPANMGR-5-RSTP_TOPOLOGY_CHANGE: BridgeAddr: 0001.e867.b1f8 Bridge port Po 1 transitioned from forwarding to discarding state. %SPANMGR-5-RSTP_TOPOLOGY_CHANGE: BridgeAddr: 0001.e867.b1f8 Bridge port Gi 3/19 transitioned from learning to forwarding state.
vrrp	%VRRP-6-VRRP_MASTER: IPv4 vrid-1 on Vl 2 VRF default-vrf entering MASTER. %VRRP-6-VRRP_MASTER: IPv4 vrid-255 on Gi 3/2 VRF default-vrf leaving MASTER %VRRP-6-VRRP_MASTER: IPv6 vrid-150 on Po 1 VRF default-vrf entering MASTER. %VRRP-6-VRRP_MASTER: IPv6 vrid-100 on Po 1 VRF default-vrf leaving MASTER. %VRRP-6-VRRP_BACKUP: IPv4 vrid-1 on Vl 2 VRF default-vrf entering BACKUP. %VRRP-6-VRRP_BACKUP: IPv6 vrid-100 on Po 1 VRF default-vrf entering BACKUP.
ecfm	%ECFM-5-ECFM_XCON_ALARM: Cross connect fault detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000 %ECFM-5-ECFM_ERROR_ALARM: Error CCM Defect detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000 %ECFM-5-ECFM_MAC_STATUS_ALARM: MAC Status Defect detected by MEP 1 in Domain provider at Level 4 VLAN 3000 %ECFM-5-ECFM_REMOTE_ALARM: Remote CCM Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000 %ECFM-5-ECFM_RDI_ALARM: RDI Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000
<cr>	SNMP Copy Config Command Completed %RPM0-P:CP %SNMP-4-RMON_RISING_THRESHOLD: RMON rising threshold alarm from SNMP OID <oid> %RPM0-P:CP %SNMP-4-RMON_FALLING_THRESHOLD: RMON falling threshold alarm from SNMP OID <oid> %RPM0-P:CP %SNMP-4-RMON_HC_RISING_THRESHOLD: RMON high-capacity rising threshold alarm from SNMP OID <oid>

Copy Configuration Files Using SNMP

Use SNMP from a remote client to:

- copy the running-config file to the startup-config file, or
- copy configuration files from the Dell Force10 system to a server
- copy configuration files from a server to the Dell Force10 system

All of these tasks can be performed using IPv4 or IPv6 addresses. The examples in this section use IPv4 addresses; IPv6 addresses can be substituted for the IPv4 addresses in all of the examples.

The relevant MIBs for these functions are:

Table 48-3. MIB Objects for Copying Configuration Files via SNMP

MIB Object	OID	Object Values	Description
copySrcFileType	.1.3.6.1.4.1.6027.3.5.1.1.1.1.2	1 = FTOS file 2 = running-config 3 = startup-config	Specifies the type of file to copy from. Valid values are: • If the copySrcFileType is running-config or startup-config, the default copySrcFileLocation is flash. • If the copySrcFileType is a binary file, the copySrcFileLocation and copySrcFileName must also be specified.
copySrcFileLocation	.1.3.6.1.4.1.6027.3.5.1.1.1.1.3	1 = flash 2 = slot0 3 = tftp 4 = ftp 5 = scp	Specifies the location of source file. • If the copySrcFileLocation is FTP or SCP, copyServerAddress, copyUserName, and copyUserPassword must be specified.
copySrcFileName	.1.3.6.1.4.1.6027.3.5.1.1.1.1.4	Path (if file is not in current directory) and filename.	Specifies name of the file. • If copySourceFileType is set to running-config or startup-config, copySrcFileName is not required.
copyDestFileType	.1.3.6.1.4.1.6027.3.5.1.1.1.1.5	1 = FTOS file 2 = running-config 3 = startup-config	Specifies the type of file to copy to. • If the copySourceFileType is running-config or startup-config, the default copyDestFileLocation is flash. • If the copyDestFileType is a binary the copyDestFileLocation and copyDestFileName must be specified.
copyDestFileLocation	.1.3.6.1.4.1.6027.3.5.1.1.1.1.6	1 = flash 2 = slot0 3 = tftp 4 = ftp 5 = scp	Specifies the location of destination file. • If the copyDestFileLocation is FTP or SCP, copyServerAddress, copyUserName, and copyUserPassword must be specified.
copyDestFileName	.1.3.6.1.4.1.6027.3.5.1.1.1.1.7	Path (if file is not in default directory) and filename.	Specifies the name of destination file.
copyServerAddress	.1.3.6.1.4.1.6027.3.5.1.1.1.1.8	IP Address of the server	The IP address of the server. • If the copyServerAddress is specified so must copyUserName, and copyUserPassword.

Table 48-3. MIB Objects for Copying Configuration Files via SNMP

MIB Object	OID	Object Values	Description
copyUserName	.1.3.6.1.4.1.6027.3.5.1.1.1.1.9	Username for the server.	Username for for the FTP, TFTP, or SCP server. • If the copyUserName is specified so must copyUserPassword.
copyUserPassword	.1.3.6.1.4.1.6027.3.5.1.1.1.1.10	Password for the server.	Password for the FTP, TFTP, or SCP server.

To copy a configuration file:

Step	Task	Command Syntax	Command Mode
1	Create an SNMP community string with read/write privileges.	**snmp-server community** *community-name* **rw**	CONFIGURATION
2	Copy the *f10-copy-config.mib* MIB from the Dell Force10 iSupport webpage to the server to which you are copying the configuration file.		
3	On the server, use the command **snmpset** as shown: **snmpset -v** *snmp-version* **-c** *community-name* **-m** *mib_path*/**f10-copy-config.mib** *force10system-ip-address mib-object.index* {**i** | **a** | **s**} *object-value...* • Every specified object must have an object value, which must be preceded by the keyword **i**. See Table 6 for valid values. • *index* must be unique to all previously executed **snmpset** commands. If an index value has been used previously, a message like the one in Message 3 appears. In this case, increment the index value and enter the command again. • Use as many MIB Objects in the command as required by the MIB Object descriptions in Table 6 to complete the command. See Table 7 or examples. **Note:** You can use the entire OID rather than the object name. Use the form: *OID.index* **i** *object-value*, as shown in Figure 57.		

Message 2 snmpset Index Value Error

```
Error in packet.
Reason: notWritable (that object does not support modification)
Failed object: FORCE10-COPY-CONFIG-MIB::copySrcFileType.101
```

Table 7 shows examples of using the command **snmpset** to copy a configuration. These examples assume that:

- the server OS is Unix
- you are using SNMP version 2c
- the community name is public, and
- the file *f10-copy-config.mib* is in the current directory or in the snmpset tool path.

Note: In Unix, enter the command **snmpset** for help using this command. Place the file *f10-copy-config.mib* the directory from which you are executing the **snmpset** command or in the snmpset tool path.

Table 48-4. Copying Configuration Files via SNMP

Task
Copy the running-config to the startup-config using the following command from the Unix machine: **snmpset -v 2c -c public -m ./f10-copy-config.mib** *force10system-ip-address* **copySrcFileType.***index* **i 2 copyDestFileType.***index* **i 3**

Figure 48-6 show the command syntax using MIB object names, and Figure 48-7 shows the same command using the object OIDs. In both cases, the object is followed by a unique index number.

Figure 48-6. Copying Configuration Files via SNMP using Object-Name Syntax

```
> snmpset -v 2c -r 0 -t 60 -c public -m ./f10-copy-config.mib 10.10.10.10 copySrcFileType.101 i
2 copyDestFileType.101 i 3
FORCE10-COPY-CONFIG-MIB::copySrcFileType.101 = INTEGER: runningConfig(2)
FORCE10-COPY-CONFIG-MIB::copyDestFileType.101 = INTEGER: startupConfig(3)
```

Figure 48-7. Copying Configuration Files via SNMP using OID Syntax

```
> snmpset -v 2c -c public -m ./f10-copy-config.mib 10.10.10.10
.1.3.6.1.4.1.6027.3.5.1.1.1.2.100 i 2 .1.3.6.1.4.1.6027.3.5.1.1.1.5.100 i 3
FORCE10-COPY-CONFIG-MIB::copySrcFileType.100 = INTEGER: runningConfig(2)
FORCE10-COPY-CONFIG-MIB::copyDestFileType.100 = INTEGER: startupConfig(3)
```

Copy the startup-config to the running-config using the following command from a Unix machine:
snmpset -c private -v 2c *force10system-ip-address* **copySrcFileType.***index* **i 3 copyDestFileType.***index* **i 2**

Figure 48-8. Copying Configuration Files via SNMP using Object-Name Syntax

```
> snmpset -c public -v 2c -m ./f10-copy-config.mib 10.11.131.162 copySrcFileType.7 i 3
copyDestFileType.7 i 2
FORCE10-COPY-CONFIG-MIB::copySrcFileType.7 = INTEGER: runningConfig(3)
FORCE10-COPY-CONFIG-MIB::copyDestFileType.7 = INTEGER: startupConfig(2)
```

Figure 48-9. Copying Configuration Files via SNMP using OID Syntax

```
>snmpset -c public -v 2c 10.11.131.162 .1.3.6.1.4.1.6027.3.5.1.1.1.1.2.8 i 3
.1.3.6.1.4.1.6027.3.5.1.1.1.1.5.8 i 2
SNMPv2-SMI::enterprises.6027.3.5.1.1.1.1.2.8 = INTEGER: 3
SNMPv2-SMI::enterprises.6027.3.5.1.1.1.1.5.8 = INTEGER: 2
```

Copy the startup-config to the server via FTP using the following command from the Unix machine:

snmpset -v 2c -c public -m ./f10-copy-config.mib *force10system-ip-address* **copySrcFileType.***index* **i 2 copyDestFileName.***index* **s** *filepath/filename* **copyDestFileLocation.***index* **i 4 copyServerAddress.***index* **a** *server-ip-address* **copyUserName.***index* **s** *server-login-id* **copyUserPassword.***index* **s** *server-login-password*

Table 48-4. Copying Configuration Files via SNMP

Task
• *server-ip-address* must be preceded by the keyword **a**. • values for copyUsername and copyUserPassword must be preceded by the keyword **s**.

Figure 48-10. Copying Configuration Files via SNMP and FTP to a Remote Server

```
> snmpset -v 2c -c private -m ./f10-copy-config.mib 10.10.10.10 copySrcFileType.110 i 2
copyDestFileName.110 s /home/startup-config copyDestFileLocation.110 i 4 copyServerAddress.110
a 11.11.11.11 copyUserName.110 s mylogin copyUserPassword.110 s mypass
FORCE10-COPY-CONFIG-MIB::copySrcFileType.110 = INTEGER: runningConfig(2)
FORCE10-COPY-CONFIG-MIB::copyDestFileName.110 = STRING: /home/startup-config
FORCE10-COPY-CONFIG-MIB::copyDestFileLocation.110 = INTEGER: ftp(4)
FORCE10-COPY-CONFIG-MIB::copyServerAddress.110 = IpAddress: 11.11.11.11
FORCE10-COPY-CONFIG-MIB::copyUserName.110 = STRING: mylogin
FORCE10-COPY-CONFIG-MIB::copyUserPassword.110 = STRING: mypass
```

Copy the startup-config to the server via TFTP using the following command from the Unix machine:

Note: Verify that the file exists and its permissions are set to 777, and specify the relative path to the TFTP root directory.

snmpset -v 2c -c public -m ./f10-copy-config.mib *force10system-ip-address* **copySrcFileType.***index* **i 3 copyDestFileType.***index* **i 1 copyDestFileName.***index* **s** *filepathIfilename* **copyDestFileLocation.***index* **i 3 copyServerAddress.***index* **a** *server-ip-address*

Figure 48-11. Copying Configuration Files via SNMP and TFTP to a Remote Server

```
.snmpset -v 2c -c private -m ./f10-copy-config.mib 10.10.10.10
copySrcFileType.4 i 3
copyDestFileType.4 i 1
copyDestFileLocation.4 i 3
copyDestFileName.4 s /home/myfilename
copyServerAddress.4 a 11.11.11.11
```

Copy a binary file from the server to the startup-configuration on the Dell Force10 system via FTP using the following command from the Unix server:

snmpset -v 2c -c public -m ./f10-copy-config.mib *force10system-ip-address* **copySrcFileType.***index* **i 1 copySrcFileLocation.***index* **i 4 copySrcFileName.***index* **s** *filepathIfilename* **copyDestFileType.***index* **i 3 copyServerAddress.***index* **a** *server-ip-address* **copyUserName.***index* **s** *server-login-id* **copyUserPassword.***index* **s** *server-login-password*

Figure 48-12. Copying Configuration Files via SNMP and FTP from a Remote Server

```
> snmpset -v 2c -c private -m ./f10-copy-config.mib 10.10.10.10 copySrcFileType.10 i 1
copySrcFileLocation.10 i 4 copyDestFileType.10 i 3 copySrcFileName.10 s /home/myfilename
copyServerAddress.10 a 172.16.1.56 copyUserName.10 s mylogin copyUserPassword.10 s mypass
```

Dell Force10 provides additional MIB Objects to view copy statistics. These are provided in Table 48-5.

Table 48-5. MIB Objects for Copying Configuration Files via SNMP

MIB Object	OID	Values	Description
copyState	.1.3.6.1.4.1.6027.3.5.1.1.1.1.11	1= running 2 = successful 3 = failed	Specifies the state of the copy operation.
copyTimeStarted	.1.3.6.1.4.1.6027.3.5.1.1.1.1.12	Time value	Specifies the point in the up-time clock that the copy operation started.
copyTimeCompleted	.1.3.6.1.4.1.6027.3.5.1.1.1.1.13	Time value	Specifies the point in the up-time clock that the copy operation completed.
copyFailCause	.1.3.6.1.4.1.6027.3.5.1.1.1.1.14	1 = bad file name 2 = copy in progress 3 = disk full 4 = file exists 5 = file not found 6 = timeout 7 = unknown	Specifies the reason the copy request failed.
copyEntryRowStatus	.1.3.6.1.4.1.6027.3.5.1.1.1.1.15	Row status	Specifies the state of the copy operation. Uses CreateAndGo when you are performing the copy. The state is set to active when the copy is completed.

To obtain a value for any of the MIB Objects in Table 48-5:

Step	Task
1	Get a copy-config MIB object value. **snmpset -v 2c -c public -m ./f10-copy-config.mib** *force10system-ip-address* [*OID.index* | *mib-object.index* • *index* is the index value used in the **snmpset** command used to complete the copy operation. **Note:** You can use the entire OID rather than the object name. Use the form: *OID.index*, as shown in Figure 48-14.

Figure 48-13 and Figure 48-14 are examples of using the command **snmpget** to obtain a MIB object value. These examples assume that:

- the server OS is Unix
- you are using SNMP version 2c
- the community name is public, and
- the file *f10-copy-config.mib* is in the current directory.

Note: In Unix, enter the command **snmpset** for help using this command.

Figure 48-13 shows the command syntax using MIB object names, and Figure 48-14 shows the same command using the object OIDs. In both cases, the object is followed by same index number used in the **snmpset** command.

Figure 48-13. Obtaining MIB Object Values for a Copy Operation using Object-name Syntax

```
>snmpget -v2c -cprivate -m ./f10-copy-config.mib10.11.131.140copyTimeCompleted.110
FORCE10-COPY-CONFIG-MIB::copyTimeCompleted.110 = Timeticks: (1179831) 3:16:38.31
```

Figure 48-14. Obtaining MIB Object Values for a Copy Operation using OID Syntax

```
> snmpget -v 2c -c private 10.11.131.140 .1.3.6.1.4.1.6027.3.5.1.1.1.1.13.110
SNMPv2-SMI::enterprises.6027.3.5.1.1.1.1.13.110 = Timeticks: (1179831) 3:16:38.31
```

Manage VLANs using SNMP

The qBridgeMIB managed objects in the Q-BRIDGE-MIB, defined in RFC 2674, enable you to use SNMP manage VLANs.

Create a VLAN

Use the dot1qVlanStaticRowStatus object to create a VLAN. The snmpset operation in Figure 48-15 creates VLAN 10 by specifying a value of 4 for instance 10 of the dot1qVlanStaticRowStatus object.

Figure 48-15. Creating a VLAN using SNMP

```
> snmpset -v2c -c mycommunity 123.45.6.78 .1.3.6.1.2.1.17.7.1.4.3.1.5.10 i 4
SNMPv2-SMI::mib-2.17.7.1.4.3.1.5.10 = INTEGER: 4
```

Assign a VLAN Alias

Write a character string to the dot1qVlanStaticName object to assign a name to a VLAN, as shown in Figure 48-16.

Figure 48-16. Assign a VLAN Alias using SNMP

```
[Unix system output]

> snmpset -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.1.1107787786 s "My
VLAN"
SNMPv2-SMI::mib-2.17.7.1.4.3.1.1.1107787786 = STRING: "My VLAN"

[FTOS system output]

FTOS#show int vlan 10
Vlan 10 is down, line protocol is down
Vlan alias name is: My VLAN
Address is 00:01:e8:cc:cc:ce, Current address is 00:01:e8:cc:cc:ce
Interface index is 1107787786
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed auto
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 01:01:00
Queueing strategy: fifo
Time since last interface status change: 01:01:00
```

Display the Ports in a VLAN

FTOS identifies VLAN interfaces using an interface index number that is displayed in the output of the command **show interface vlan**, as shown in Figure 48-17.

Figure 48-17. Identifying the VLAN Interface Index Number

```
FTOS(conf)#do show interface vlan id 10
% Error: No such interface name.
R5(conf)#do show interface vlan 10
Vlan 10 is down, line protocol is down
Address is 00:01:e8:cc:cc:ce, Current address is 00:01:e8:cc:cc:ce
Interface index is 1107787786
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed auto
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:12:42
Queueing strategy: fifo
Time since last interface status change: 00:12:42
```

To display the ports in a VLAN, send an **snmpget** request for the object dot1qStaticEgressPorts using the interface index as the instance number, as shown for an S-Series in Figure 48-18.

Figure 48-18. Display the Ports in a VLAN in SNMP

```
> snmpget -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.2.1107787786
SNMPv2-SMI::mib-2.17.7.1.4.3.1.2.1107787786 = Hex-STRING:
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00
```

The table that the Dell Force10 system sends in response to the **snmpget** request is a table that contains hexadecimal (hex) pairs, each pair representing a group of eight ports.

- On the E-Series, 12 hex pairs represents a line card. Twelve pairs accommodates the greatest currently available line card port density, 96 ports.
- On the C-Series, 28 hex pairs represents a line card. Twenty-eight pairs accommodates the greatest currently available line card port density, 28 ports per port-pipe, and any remaining hex pairs are unused.
- On the S-Series, 7 hex pairs represents a stack unit. Seven pairs accommodates the greatest number of ports available on an S-Series, 56 ports. The last stack unit is assigned 8 pairs; the eighth pair is unused.

The first hex pair, 00 in Figure 48-18, represents ports 1-7 in Stack Unit 0. The next pair to the right represents ports 8-15. To resolve the hex pair into a representation of the individual ports, convert the hex pair to binary. Consider the first hex pair 00, which resolves to 0000 0000 in binary:

- On the E-Series and C-Series each position in the 8-character string is for one port, starting with Port 0 at the left end of the string, and ending with Port 7 at the right end. A 0 indicates that the port is not a member of the VLAN; a 1 indicates VLAN membership.
- On the S-Series, each position in the 8-character string is for one port, starting with Port 1 at the left end of the string, and ending with Port 8 at the right end. A 0 indicates that the port is not a member of the VLAN; a 1 indicates VLAN membership.

Figure 48-18 shows the output for an S-Series. All hex pairs are 00, indicating that no ports are assigned to VLAN 10. In Figure 48-19, Port 0/2 is added to VLAN 10 as untagged. And the first hex pair changes from 00 to 04.

Figure 48-19. Displaying Ports in a VLAN using SNMP

```
[FTOS system output]

R5(conf)#do show vlan id 10

Codes: * - Default VLAN, G - GVRP VLANs
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged, M - Vlan-stack

    NUM    Status    Description                    Q Ports
    10     Inactive                                 U Gi 0/2

[Unix system output]

> snmpget -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.2.1107787786
SNMPv2-SMI::mib-2.17.7.1.4.3.1.2.1107787786 = Hex-STRING: 40 00 00 00 00 00 00 00 00 00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00
```

The value 40 is in the first set of 7 hex pairs, indicating that these ports are in Stack Unit 0. The hex value 40 is 0100 0000 in binary. As described above, the left-most position in the string represents Port 1. The next position from the left represents Port 2 and has a value of 1, indicating that Port 0/2 is in VLAN 10. The remaining positions are 0, so those ports are not in the VLAN.

Note that the table contains none of the other information provided by the **show vlan** command, such as port speed or whether the ports are tagged or untagged.

Add Tagged and Untagged Ports to a VLAN

The value dot1qVlanStaticEgressPorts object is an array of all VLAN members.

The dot1qVlanStaticUntaggedPorts object is an array of only untagged VLAN members. All VLAN members that are not in dot1qVlanStaticUntaggedPorts are tagged.

- To add a tagged port to a VLAN, write the port to the dot1qVlanStaticEgressPorts object, as shown in Figure 48-20.
- To add an untagged port to a VLAN, write the port to the dot1qVlanStaticEgressPorts and dot1qVlanStaticUntaggedPorts objects, as shown in Figure 48-21.

Note: Whether adding a tagged or untagged port, you must specify values for both dot1qVlanStaticEgressPorts and dot1qVlanStaticUntaggedPorts.

In Figure 48-20, Port 0/2 is added as an untagged member of VLAN 10.

Figure 48-20. Adding Untagged Ports to a VLAN using SNMP

```
>snmpset -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.2.1107787786 x "40 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00"
.1.3.6.1.2.1.17.7.1.4.3.1.4.1107787786 x "40 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00"
SNMPv2-SMI::mib-2.17.7.1.4.3.1.2.1107787786 = Hex-STRING: 40 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00
SNMPv2-SMI::mib-2.17.7.1.4.3.1.4.1107787786 = Hex-STRING: 40 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00
```

In Figure 48-21, Port 0/2 is added as a tagged member of VLAN 10.

Figure 48-21. Adding Tagged Ports to a VLAN using SNMP

```
>snmpset -v2c -c mycommunity 10.11.131.185 .1.3.6.1.2.1.17.7.1.4.3.1.2.1107787786 x "40 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00"
.1.3.6.1.2.1.17.7.1.4.3.1.4.1107787786 x "00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00"
SNMPv2-SMI::mib-2.17.7.1.4.3.1.2.1107787786 = Hex-STRING: 40 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00
SNMPv2-SMI::mib-2.17.7.1.4.3.1.4.1107787786 = Hex-STRING: 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00
```

Enable and Disable a Port using SNMP

Step	Task	Command Syntax	Command Mode
1	Create an SNMP community on the Dell Force10 system.	**snmp-server community**	CONFIGURATION
2	From the Dell Force10 system, identify the interface index of the port for which you want to change the admin status. Or, from the management system, use the **snmpwwalk** command to identify the interface index.	**show interface**	EXEC Privilege
3	Enter the command **snmpset** to change the admin status using either the object descriptor or the OID. Choose integer 1 to change the admin status to Up, or 2 to change the admin status to Down. **snmpset with descriptor: snmpset -v** *version* **-c** *community* **agent-ip ifAdminStatus.***ifindex* **i** {**1** | **2**} **snmpset with OID: snmpset -v** *version* **-c** *community* **agent-ip .1.3.6.1.2.1.2.2.1.7.***ifindex* **i** {**1** | **2**}		

Fetch Dynamic MAC Entries using SNMP

Dell Force10 supports the RFC 1493 dot1d table for the default VLAN and the dot1q table for all other VLANs.

Note: The 802.1q Q-BRIDGE MIB defines VLANs with regard to 802.1d, as 802.1d itself does not define them. As a switchport must belong a VLAN (the default VLAN or a configured VLAN), all MAC address learned on a switchport are associated with a VLAN. For this reason, the Q-Bridge MIB is used for MAC address query. Moreover, specific to MAC address query, dot1dTpFdbTable is indexed by MAC address only for a single forwarding database, while dot1qTpFdbTable has two indices —VLAN ID and MAC address —to allow for multiple forwarding databases and considering that the same MAC address is learned on multiple VLANs. The VLAN ID is added as the first index so that MAC addresses can be read by VLAN, sorted lexicographically. The MAC address is are part of the OID instance, so in this case, lexicographic order is according to the most significant octet.

Table 48-6. MIB Objects for Fetching Dynamic MAC Entries in the Forwarding Database

MIB Object	OID	Description	MIB
dot1dTpFdbTable	.1.3.6.1.2.1.17.4.3	List the learned unicast MAC addresses on the default VLAN.	Q-BRIDGE MIB
dot1qTpFdbTable	.1.3.6.1.2.1.17.7.1.2.2	List the learned unicast MAC addresses on non-default VLANs.	
dot3aCurAggFdbTable	.1.3.6.1.4.1.6027.3.2.1.1.5	List the learned MAC addresses of aggregated links (LAG).	F10-LINK-AGGREGATION-MIB

In Figure 48-22, R1 has one dynamic MAC address, learned off of port GigabitEthernet 1/21, which a member of the default VLAN, VLAN 1. The SNMP walk returns the values for dot1dTpFdbAddress, dot1dTpFdbPort, and dot1dTpFdbStatus.

Each object is comprised an OID concatenated with an instance number. In the case of these objects, the instance number is the decimal equivalent of the MAC address; derive the instance number by converting each hex pair to its decimal equivalent. For example, the decimal equivalent of E8 is 232, and so the instance number for MAC address 00:01:e8:06:95:ac is .0.1.232.6.149.172.

The value of dot1dTpFdbPort is the port number of the port off which the system learns the MAC address. In this case, of GigabitEthernet 1/21, the manager returns the integer 118. The maximum line card port density on the E-Series is 96 ports, and line card numbering begins with 0; GigabitEthernet 1/21 is the 21st port on Line Card 1, and 96 + 21 yields 118.

Figure 48-22. Fetching Dynamic MAC Addresses on the Default VLAN

```
-----------------------------MAC Addresses on FTOS System-----------------------------
R1_E600#show mac-address-table
VlanId     Mac Address           Type   Interface        State
 1      00:01:e8:06:95:ac       Dynamic Gi 1/21         Active
-----------------------------Query from Management Station-----------------------------
>snmpwalk -v 2c -c techpubs 10.11.131.162 .1.3.6.1.2.1.17.4.3.1
SNMPv2-SMI::mib-2.17.4.3.1.1.0.1.232.6.149.172 = Hex-STRING: 00 01 E8 06 95 AC
SNMPv2-SMI::mib-2.17.4.3.1.2.0.1.232.6.149.172 = INTEGER: 118
SNMPv2-SMI::mib-2.17.4.3.1.3.0.1.232.6.149.172 = INTEGER: 3
```

In Figure 48-23, GigabitEthernet 1/21 is moved to VLAN 1000, a non-default VLAN. Use the objects dot1qTpFdbTable to fetch the MAC addresses learned on non-default VLANs. The instance number is the VLAN number concatenated with the decimal conversion of the MAC address.

Figure 48-23. Fetching Dynamic MAC Addresses on Non-default VLANs

```
-----------------------------MAC Addresses on FTOS System-----------------------------
R1_E600#show mac-address-table
VlanId     Mac Address           Type   Interface        State
 1000   00:01:e8:06:95:ac       Dynamic Gi 1/21         Active
-----------------------------Query from Management Station-----------------------------
>snmpwalk -v 2c -c techpubs 10.11.131.162 .1.3.6.1.2.1.17.7.1.2.2.1
SNMPv2-SMI::mib-2.17.7.1.2.2.1.2.1000.0.1.232.6.149.172 = INTEGER: 118
SNMPv2-SMI::mib-2.17.7.1.2.2.1.3.1000.0.1.232.6.149.172 = INTEGER: 3
```

Use dot3aCurAggFdbTable to fetch the learned MAC address of a port-channel. The instance number is the decimal conversion of the MAC address concatenated with the port-channel number.

Figure 48-24. Fetching Dynamic MAC Addresses on the Default VLAN

```
-----------------------------MAC Addresses on FTOS System-----------------------------
R1_E600(conf)#do show mac-address-table
VlanId     Mac Address           Type   Interface        State
 1000   00:01:e8:06:95:ac       Dynamic Po 1            Active
-----------------------------Query from Management Station-----------------------------
>snmpwalk -v 2c -c techpubs 10.11.131.162 .1.3.6.1.4.1.6027.3.2.1.1.5
SNMPv2-SMI::enterprises.6027.3.2.1.1.5.1.1.1000.0.1.232.6.149.172.1 = INTEGER: 1000
SNMPv2-SMI::enterprises.6027.3.2.1.1.5.1.2.1000.0.1.232.6.149.172.1 = Hex-STRING: 00 01 E8
06 95 AC
SNMPv2-SMI::enterprises.6027.3.2.1.1.5.1.3.1000.0.1.232.6.149.172.1 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.5.1.4.1000.0.1.232.6.149.172.1 = INTEGER: 1
```

Deriving Interface Indices

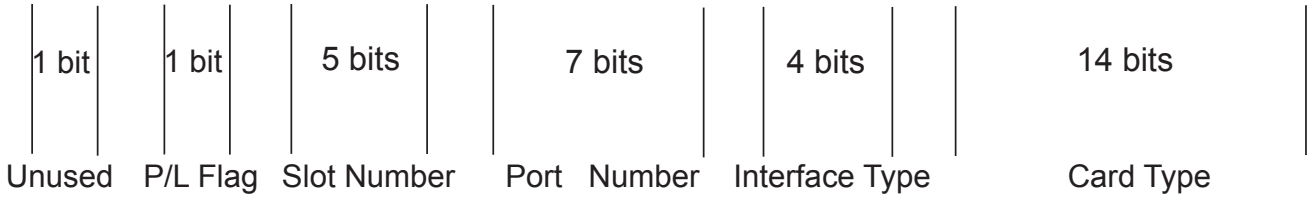
FTOS assigns an interface number to each (configured or unconfigured) physical and logical interface. Display the interface index number using the command **show interface** from EXEC Privilege mode, as shown in Figure 48-25.

Figure 48-25. Display the Interface Index Number

```
FTOS#show interface gig 1/21
GigabitEthernet 1/21 is up, line protocol is up
Hardware is Force10Eth, address is 00:01:e8:0d:b7:4e
    Current address is 00:01:e8:0d:b7:4e
Interface index is 72925242
[output omitted]
FTOS#show linecard all | grep 1
  1   online        online     E48TF    E48TF    7.7.1.1     48
```

The interface index is a binary number with bits that indicate the slot number, port number, interface type, and card type of the interface. FTOS converts this binary index number to decimal, and displays it in the output of the **show interface** command.

Figure 48-26. Interface Index Binary Calculations

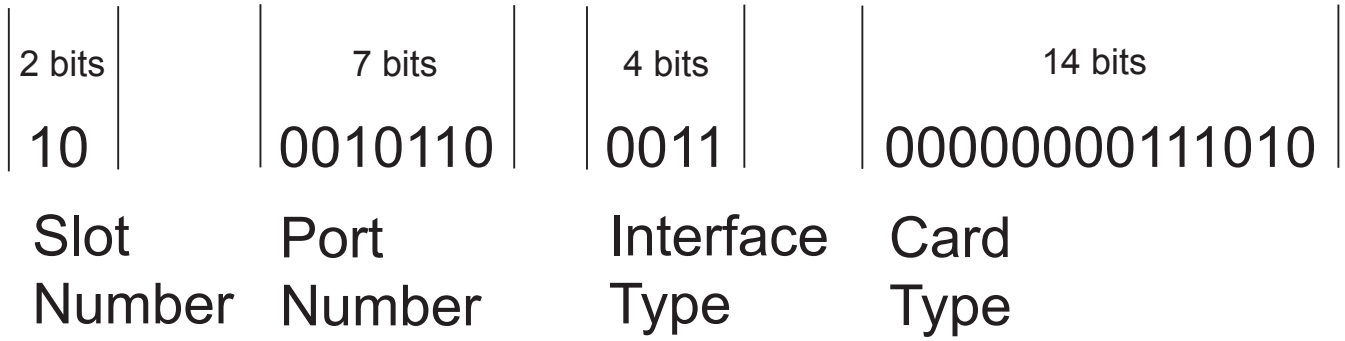


Starting from the least significant bit (LSB):

- the first 14 bits represent the card type
- the next 4 bits represent the interface type
- the next 7 bits represent the port number
- the next 5 bits represent the slot number
- the next 1 bit is 0 for a physical interface and 1 for a logical interface
- the next 1 bit is unused

For example, the index 72925242 is 10001011000110000000000111010 in binary. The binary interface index for GigabitEthernet 1/21 of a 48-port 10/100/1000Base-T line card with RJ-45 interface is shown in Figure 48-27. Notice that the physical/logical bit and the final, unused bit are not given. The interface is physical, so this must be represented by a 0 bit, and the unused bit is always 0. These two bits are not given because they are the most significant bits, and leading zeros are often omitted.

Figure 48-27. Binary Representation of Interface Index



For interface indexing, slot and port numbering begins with the binary one. If the Dell Force10 system begins slot and port numbering from 0, then the binary 1 represents slot and port 0. For example, the index number in Figure 48-27 gives the binary 2 for the slot number, though interface GigabitEthernet 1/21 belongs to Slot 1. This is because the port for this example is on an E-Series which begins numbering slots from 0. You must subtract 1 from the slot number 2, which yields 1, the correct slot number for interface 1/21.

Note that the interface index does not change if the interface reloads or fails over. On the S-Series, if the unit is renumbered (for any reason) the interface index will change during a reload.

Monitor Port-channels

To check the status of a Layer 2 port-channel, use f10LinkAggMib (.1.3.6.1.4.1.6027.3.2). Below, Po 1 is a switchport and Po 2 is in Layer 3 mode.

```
[senthilnathan@lithium ~]$ snmpwalk -v 2c -c public 10.11.1.1 .1.3.6.1.4.1.6027.3.2.1.1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.1.1 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.1.2 = INTEGER: 2
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.2.1 = Hex-STRING: 00 01 E8 13 A5 C7
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.2.2 = Hex-STRING: 00 01 E8 13 A5 C8
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.3.1 = INTEGER: 1107755009
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.3.2 = INTEGER: 1107755010
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.4.1 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.4.2 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.5.1 = Hex-STRING: 00 00
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.5.2 = Hex-STRING: 00 00
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.6.1 = STRING: "Gi 5/84 " << Channel member for Po1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.1.6.2 = STRING: "Gi 5/85 " << Channel member for Po2
dot3aCommonAggFdbIndex
SNMPv2-SMI::enterprises.6027.3.2.1.1.6.1.1.1107755009.1 = INTEGER: 1107755009
dot3aCommonAggFdbVlanId
SNMPv2-SMI::enterprises.6027.3.2.1.1.6.1.2.1107755009.1 = INTEGER: 1
dot3aCommonAggFdbTagConfig
SNMPv2-SMI::enterprises.6027.3.2.1.1.6.1.3.1107755009.1 = INTEGER: 2 (Tagged 1 or Untagged 2)
dot3aCommonAggFdbStatus
SNMPv2-SMI::enterprises.6027.3.2.1.1.6.1.4.1107755009.1 = INTEGER: 1 << Status active, 2 – status inactive
```

If we learn mac address for the LAG, status will be shown for those as well

```
dot3aCurAggVlanId
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.1.1.0.0.0.0.0.1.1 = INTEGER: 1
dot3aCurAggMacAddr
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.2.1.0.0.0.0.0.1.1 = Hex-STRING: 00 00 00 00 00 01
dot3aCurAggIndex
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.3.1.0.0.0.0.0.1.1 = INTEGER: 1
dot3aCurAggStatus
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.4.1.0.0.0.0.0.1.1 = INTEGER: 1 << Status active, 2 –
status inactive
```

For L3 lag we don’t have this support.

SNMP trap works fine for L2 / L3 / default mode LAG

```
SNMPv2-MIB::sysUpTime.0 = Timeticks: (8500842) 23:36:48.42
SNMPv2-MIB::snmpTrapOID.0 = OID: IF-MIB::linkDown
IF-MIB::ifIndex.33865785 = INTEGER: 33865785
SNMPv2-SMI::enterprises.6027.3.1.1.4.1.2 = STRING: "OSTATE_DN: Changed interface state to down: Gi
0/0"
2010-02-10 14:22:39 10.16.130.4 [10.16.130.4]:
SNMPv2-MIB::sysUpTime.0 = Timeticks: (8500842) 23:36:48.42
SNMPv2-MIB::snmpTrapOID.0 = OID: IF-MIB::linkDown
IF-MIB::ifIndex.1107755009 = INTEGER: 1107755009
SNMPv2-SMI::enterprises.6027.3.1.1.4.1.2 = STRING: "OSTATE_DN: Changed interface state to down: Po
1"
2010-02-10 14:22:40 10.16.130.4 [10.16.130.4]:
SNMPv2-MIB::sysUpTime.0 = Timeticks: (8500932) 23:36:49.32      SNMPv2-MIB::snmpTrapOID.0 = OID:
IF-MIB::linkUp IF-MIB::ifIndex.33865785 = INTEGER: 33865785
SNMPv2-SMI::enterprises.6027.3.1.1.4.1.2 = STRING: "OSTATE_UP: Changed interface state to up: Gi
0/0"
2010-02-10 14:22:40 10.16.130.4 [10.16.130.4]:
SNMPv2-MIB::sysUpTime.0 = Timeticks: (8500934) 23:36:49.34      SNMPv2-MIB::snmpTrapOID.0 = OID:
IF-MIB::linkUp IF-MIB::ifIndex.1107755009 = INTEGER: 1107755009
SNMPv2-SMI::enterprises.6027.3.1.1.4.1.2 = STRING: "OSTATE_UP: Changed interface state to up: Po
1"
```

Troubleshooting SNMP Operation

When you use SNMP to retrieve management data from an SNMP agent on a Dell Force10 router, take into account the following behavior:

- When you query an IPv4 icmpMsgStatsInPkts object in the ICMP table by using the **snmpwalk** command, the output for echo replies may be incorrectly displayed. Use the **show ip traffic** command to correctly display this information under ICMP statistics in the command output.
- When you query an icmpStatsInErrors object in the icmpStats table by using the **snmpget** or **snmpwalk** command, the output for IPv4 addresses may be incorrectly displayed. Use the **show ip traffic** command to correctly display this information under IP and ICMP statistics.
- When you query an IPv4 icmpMsgStatsInPkts object in the ICMP table by using the **snmpwalk** command, the echo response output may not be displayed. Use the **show ip traffic** command to correctly display ICMP statistics, such as echo response.

49

SONET/SDH

SONET/SDH is supported on platform E

SONET/SDH is supported on the E-Series ExaScale platform with FTOS 8.1.1.2 and later.

FTOS supports two line cards with SONET—Packet-Over-SONET (POS) and PPP-over-SONET/SDH.

This chapter covers the following topics:

Packet Over SONET (POS) Interfaces

POS Interfaces are supported on E-Series TeraScale E T platforms

Important Points to Remember

- PPP encapsulation must be configured before the interface is enabled for traffic.
- An IP address must be configured on an interface. POS line cards do not operate in Layer2 mode.
- SONET alarm reporting cannot be disabled.
- SONET uses synchronous transport signal (STS) framing. Configure framing only when the interface is shut down.
- The POS scramble-atm, C2 and J0 flags, and alarm reporting is supported.
- **The following features are not supported**:
  - The E-300 chassis does not support POS line cards.
  - VRRP is not supported on POS interfaces.
  - A POS interface cannot be configured as part of a LAG.

- Protection switching is not supported.
- POS interfaces cannot be mirrored ports.
- Configurable alarm thresholds (SF/SD BER, B1/B 2/B3 TC) are not supported.
- The CRC type and S1S0 flag cannot be changed.

Configuring POS Interfaces

POS interfaces require several configuration considerations, including

- Encapsulation
- MTU
- Clock Settings

Encapsulation

The E-Series' POS line card requires PPP encapsulation. A SONET interface without encapsulation is always administratively down.

Packet Over SONET interfaces require several configuration considerations.

When enabling encapsulation on an interface, PPP negotiation only begins after it has been turned on using the **no shutdown** command. You can enable authentication and other related commands after negotiation is completed. When removing encapsulation using the **no encap** command on a SONET interface, the interface is administratively shutdown and all configuration information (IP addresses for example) is deleted from the interface.

Equipment vendors use unique defaults for PPP encapsulation. When configuring PPP encapsulation between the E-Series and another vendor's equipment, verify the following settings:

- One side of the link is set using the command **clock source internal**.
- Default SONET settings are compatible. The E-Series defaults to ATM scrambling disabled; flag is C2 0xCF(207), J0 is 0xCF(207).
- A common PPP authentication method is configured. FTOS supports Challenge-Handshake Authentication Protocol (CHAP) and/or Password Authentication Protocol (PAP) authentication.
- The MTU and IP MTU settings on both ends of the link are the same. If you input the **ip mtu** command with a value which differs from the far-end interface, the interface on the E-Series will go down.
- Confirm that the MTU settings are the same on both end of the link. If you configure the **ip mtu** command with a different value on the far end of the link, the interface on the E-Series goes down.

Note: SONET uses synchronous transport signal (STS) framing. When framing is configured on an interface, it should only be done when the interface is shut down.

Configuring Maximum Transmission Unit (MTU)

Maximum Transmission Unit is an integer value that represents the greatest number of bytes that any given interface on the system can handle. MTU settings allow the router to determine if a large packet needs to be fragmented before transmission. PPP must be enabled on a SONET interface before MTU can become configurable. MTU size can be changed in INTERFACE mode by entering the command **mtu** *size.*

Figure 49-1. MTU configuration display

```
interface SONET 0/0
 no shutdown
FTOS(conf-if-so-0/0)#encap ?
ppp                    PPP encapsulation
FTOS(conf-if-so-0/0)#encap ppp
FTOS(conf-if-so-0/0)#mtu ?
<8-9252>               POS MTU size (default = 4506)
FTOS(conf-if-so-0/0)#ip mtu ?
<576-9230>             Interface IP MTU (default is 1500)
```

10GE WAN Physical Interface

10GE interfaces support LAN and WAN modes. When in WAN mode, the 10GE interface operates as a SONET interface. Use the **wanport** command in INTERFACE mode to transition a 10GE interface into WAN mode.

Note that the port must be in shutdown state before the **wanport** command can be executed successfully. (Figure 49-2).

Note: For E-Series ExaScale systems, you must configure all the ports in a port-pipe to either WANPHY or non-WANPHY. They cannot be mixed on the same port-pipe. If you configure port 3 for example to be a WANPHY port then ports 0-4 (same port pipe) all must be WANPHY as well.

Step	Task	Command Syntax	Command Mode
1	Place the port in shutdown state	**shutdown**	INTERFACE
2	Place the port in WAN mode	**wanport**	INTERFACE
3	Display the active/defective alarms	**show controllers tengigabitethernet** *slot/port*	EXEC

Figure 49-2. wanport command example

```
interface TenGigabitEthernet 13/0
 no ip address
 no shutdown
FTOS(conf-if-te-13/0)#
FTOS(conf-if-te-13/0)#wanport
% Error: Port should be in shutdown mode, config ignored Te 13/0.
FTOS(conf-if-te-13/0)#
FTOS(conf-if-te-13/0)#shutdown
FTOS(conf-if-te-13/0)#
```

error due to no shutdown state

Figure 49-3 displays the active alarms for the interface.

Figure 49-3. show controllers tengigabitEthernet command example

```
FTOS(conf-if-te-13/0)#exit
FTOS#show controllers te 13/0

Interface is TenGigabitEthernet 13/0

SECTION
LOF = 0    LOS = 0                              BIP(B1)  = 13

LINE
AIS = 0    RDI = 1              FEBE = 7633     BIP(B2) = 19264

PATH
AIS = 0    RDI = 0    LOP = 0   FEBE = 8554     BIP(B3) = 15685

Active Defects:  LRDI

Active Alarms:   LRDI

Alarm reporting enabled for:  SLOS SLOF B1-TCA LAIS LRDI B2-TCA PAIS PRDI PLOP B3-TCA SD SF

 Framing is SONET, AIS-shut is enabled
```

Enabled Alarms are listed here (default is none)

SONET Alarm Reporting

SONET equipment detects events and alarms at each of SONET's three layers—section, line, and path. Typically, a SONET device sends alarms both upstream and downstream to notify other devices of the problem condition. The GR-253-CORE Synchronous Optical Network (SONET) Transport Systems Common Generic Criteria specification defines several alarms:

- Section Loss of Signal (SLOS)
- Section Loss of Frame (SLOF)
- Alarm Indication Signal - Line (AIS-L)
- Signal Degrade Bit Error Rate (SD-BER)
- Signal Failure Bit Error Rate (SF-BER)
- Remote Defect Indication - Line (RDI-L)

While performance monitoring provides advanced alert of link degradation, alarms indicate a failure. Fault management involves alarm monitoring and generation, reporting, logging, correlation, and clearing.

E-Series POS and 10GE WAN interfaces support the SONET alarms shown in Table 49-1:

- Section alarms—SLOS, SLOF
- Line alarms—AIS, RDI, FEBE(REI), SD, SF
- Path Alarms—AIS, RDI, FEBE(REI), LOP

Since E-Series is Terminal Equipment (TE), it must support the alarms in Table 49-1.

Table 49-1. Supported SONET Alarms

SONET/SDH Layer	Alarm	Description
Section/Regenerator	LOF	Loss of Frame condition—when a severely errored frame (SEF) defect on the incoming SONET signal and persists for 3 milliseconds
	LOS	Loss of Sync condition—when an all-zero pattern on the incoming SONET signal last 19 (+/-3) microseconds or longer. This defect might also be reported if the received signal level drops below the specified threshold.
Line/Multiplexing	AIS	Line Alarm Indication Signal is sent by the section terminating equipment (STE) to alert the downstream line terminating equipment (LTE) that a LOS or LOF defect has been detected on the incoming SONET section.
	RDI	Line Remote Defect Indication is reported by the downstream LTE when it detects LOF, LOS, or AIS.
	FEBE	Line Far End Block Errors (accumulated from the M0 or M1 byte) is reported when the downstream LTE detects BIP (B2) errors.
	SD	Signal Degrade is sourced from B2 BIP (BER). The threshold is fixed at 10^6.
	SF	Signal Failure is sourced from B2 BIP (BER). The threshold is fixed at 10^3.
Path/Section	AIS	Path Alarm Indication Signal is sent by the LTE to alert the downstream path terminating equipment (PTE) that it has detected a defect on its incoming line signal.
	RDI	Path Remote Defect Indication is reported by the downstream PTE when it detects a defect on the incoming signal.
	FEBE	Path Far End Block Errors (accumulated from G1 byte) is reported when the downstream PTE detects BIP (B3) errors.
	LOP	Loss of pointer is a result of an invalid pointer (H1,H2) or an excess number of new data flag (NDF) enable indications.

Use the **alarm-report** command to configure the SONET alarms that a POS or 10 GE WAN interface can activate.Table 49-2 defines the alarms that you can enable.

Task	Command Syntax	Command Mode
Specify which POS/SDH alarms to report to the remote SNMP server.	**alarm-report** {**lais** | **lrdi** | **pais** | **plop** | **prdi** | **sd-ber** | **sf-ber** | **slof** | **slos**}	INTERFACE

To view active alarms and defects, use the **show controllers sonet** command in EXEC Privilege mode.

Note: Historical data is not saved. The command input will show current information only.

Table 49-2. Alarm Definitions

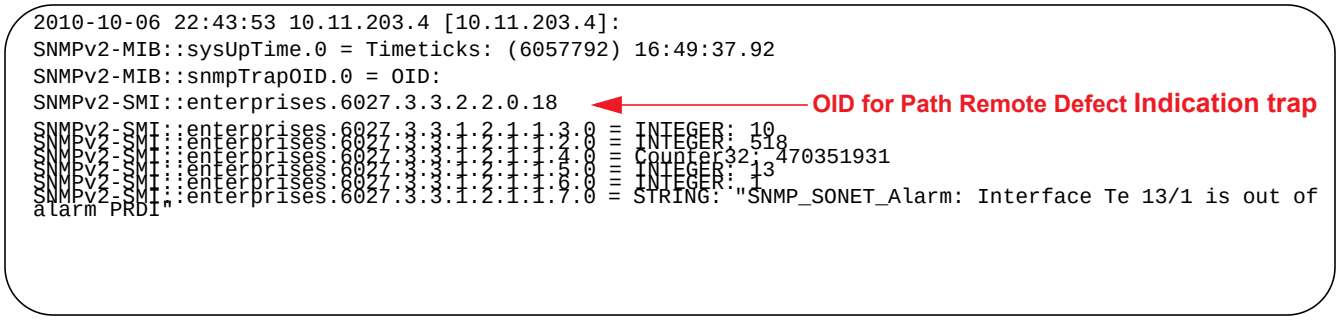
Alarm	Description
lais	Line Alarm Indication Signal
lrdi	Line Remote Defect Indication
pais	Path Alarm Indication Signal
plop	Path loss of Pointer
prdi	Path Remote Defect Indication
sd-ber	LBIP BER in excess of Signal Degradation threshold
sf-ber	LBIP BER in excess of Signal Failure threshold
slof	Section Loss of Frame
slos	Section Loss of Signal

When an E-Series POS or 10GE WAN interface detects a SONET alarm, a Syslog and SNMP trap message are generated containing information about the alarm condition.

SONET TRAP Example

SONET Traps on page 1015 describes the traps and OIDs for SONET alarms that are reported on an SNMP trap receiver. Figure 49-4 shows an example of a SONET trap.

Figure 49-4. SONET Trap example



SONET Syslog Example

Syslog messages are generated for Critical, Major, and Minor alarm conditions detected on a SONET interface according to the alarm hierarchy. For example, if a critical alarm condition is detected, a Syslog message is reported for the critical condition, but not for any major and minor alarms that may also be found. If a minor alarm condition is detected, a major or critical condition may also be reported.

Figure 49-5 shows an example of the Syslog messages generated for SONET alarms:

- Line Alarm Indication Signal (lais)
- Path Remote Defect Indication (prdi)
- Section Loss of Frame (slof)

Figure 49-5. SONET Syslog Message examples

```
Oct 6 04:46:51: %EXW4PF:13 %SONETAGT-5-ALARM: Interface Te 13/1 is out of alarm LAIS
Oct 6 04:46:42: %EXW4PF:13 %SONETAGT-6-ALARM: Interface Te 13/1 is out of alarm PRDI
Oct 6 04:46:44: %EXW4PF:13 %SONETAGT-4-ALARM: Interface Te 13/1 is in alarm SLOF
```

Events that Bring Down a SONET Interface

Down State configuration is supported on E-Series TeraScale (E)(T) platforms.

You can configure the SONET interface to change to a “down state” when certain SONET events are reported. When the event (or trigger) occurs, FTOS brings down the SONET interface. You can use the **delay triggers** command to indicate a 100ms delay in bringing down the SONET interface once the event or trigger is detected.

Task	Command Syntax	Command Mode
Delay triggering the line or path alarms with a 100ms delay.	**delay triggers** {**line [lrdi | sd-ber | sf-ber] | path [pais | prdi]** }	INTERFACE

By default, certain alarms (LOS, LOF, LAIS, PLOP) bring the line protocol down immediately. Use this command, with the **line** option, to delay that trigger event by 100ms.

By default, path alarms (AIS, RDI, LOP) do not cause (or trigger) the interface line protocol to go down. The **delay triggers** command, used with the **path** option, can be used to trigger this action.

Note: FTOS does not support configurable thresholds; **delay triggers** uses a default value of 100 ms.

SONET Port Recovery Mechanism

This feature automatically clears a condition that could cause a SONET port to hang, and stop sending and receiving data. When enabled, FTOS continuously polls status registers on SONET line cards. A port hang is declared when backpressure is detected on the port, and the port is brought down and then back up to clear the condition.The default detection interval is 60 seconds.

Task	Command Syntax	Command Mode
Implement a detection interval to find and recover hung SONET ports	**sonet-port-recover detection-interval** *interval (16-500 seconds; default is 60 seconds)*	INTERFACE

To keep a port in shutdown use the use the **hardware monitor mac action-on-error port-shutdown** command.

SONET MIB

Table 49-3 lists the managed objects supported in the SONET MIB, as defined in RFC 2558.

Table 49-3. SONET MIB: Managed Objects

SONET Managed Object	Description
sonetMediumType	Sonet or SDH depending on the configuration
sonetMediumTimeElapsed	Time in seconds (up to 900 seconds) since the line card is up. Resets after 900 seconds has elapsed
sonetMediumValidIntervals	The number of previous intervals for which valid data has been stored.
sonetMediumLineCoding	This variable describes the line coding for this interface—Non-Return to Zero (NRZ).
sonetMediumCircuitIdentifier	This variable contains the transmission vendor's circuit identifier to facilitate troubleshooting. Note that the circuit identifier, if available, is also represented by ifPhysAddress.
sonetMediumInvalidIntervals	Displays seconds in current 15 minute intervals when data could not be collected.
sonetMediumLoopbackConfig	Displays if loopback is line or internal
sonetSESthresholdSet	Displays which recognized set of SES thresholds is supported.

SONET Traps

Table 49-4 describes SONET traps supported in the Force10-specific MIB.

Table 49-4. SONET Traps and OIDs

Trap	OID	Trap Object
SONET_S_LOS Section Loss of Signal	1.3.6.1.4.1.6027.3.3.2.2.0.1	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_S_LOF Section Loss of Frame	1.3.6.1.4.1.6027.3.3.2.2.0.2	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_L_AIS Line Alarm Indication Signal	1.3.6.1.4.1.6027.3.3.2.2.0.9	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6

Table 49-4. SONET Traps and OIDs (continued)

Trap	OID	Trap Object
SONET_L_RDI Line Remote Defect Indication	1.3.6.1.4.1.6027.3.3.2.2.0.10	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_L_FEBE Line Far-end Background Block Errors	1.3.6.1.4.1.6027.3.3.2.2.0.11	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_L_B2TCA Line B2 Threshold Crossing Alert	1.3.6.1.4.1.6027.3.3.2.2.0.12	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_P_AIS Path Alarm Indication Signal	1.3.6.1.4.1.6027.3.3.2.2.0.17	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_P_RDI Path Remote Defect Indication	1.3.6.1.4.1.6027.3.3.2.2.0.18	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_P_FEBE Path Far-end Background Block Errors	1.3.6.1.4.1.6027.3.3.2.2.0.19	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_P_LOP Path Loss of Pointer	1.3.6.1.4.1.6027.3.3.2.2.0.20	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_P_NEWPTR Path New Pointer	1.3.6.1.4.1.6027.3.3.2.2.0.21	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)

Table 49-4. SONET Traps and OIDs (continued)

Trap	OID	Trap Object
SONET_P_PSE	1.3.6.1.4.1.6027.3.3.2.2.0.22	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_P_NSE	1.3.6.1.4.1.6027.3.3.2.2.0.23	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_P_B3TCA Line B3 Threshold Crossing Alert	1.3.6.1.4.1.6027.3.3.2.2.0.24	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_SD_BER Signal Degrade Bit Error Rate	1.3.6.1.4.1.6027.3.3.2.2.0.27	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_SF_BER Signal Failure Bit Error Rate	1.3.6.1.4.1.6027.3.3.2.2.0.28	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)
SONET_LOC Loss of Cell Delineation	1.3.6.1.4.1.6027.3.3.2.2.0.29	alarm state (1.3.6.1.4.1.6027.3.3.1.2.1.1.3), alarm type(1.3.6.1.4.1.6027.3.3.1.2.1.1.2), ifindex(1.3.6.1.4.1.6027.3.3.1.2.1.1.4), slot(1.3.6.1.4.1.6027.3.3.1.2.1.1.5), port(1.3.6.1.4.1.6027.3.3.1.2.1.1.6)

50

Stacking S-Series Switches

Stacking S-Series Switches is supported on platform S

Note: S-Series Stackin is not supported on the S60 system.

This chapter contains the following sections:

S-Series Stacking Overview

Up to eight S-Series systems can be interconnected so that all of the units function as a single unit.

A stack is analogous to an E-Series or C-Series system with redundant RPMs and multiple line cards. FTOS elects a primary and secondary management unit, and all other units are member units. The forwarding database resides on the primary, and all other units maintain a sychnronized local copy. Each unit in the stack makes forwarding decisions based on their local copy.

FTOS presents all of the units like line cards; for example, to access GigabitEthernet Port 1 on Stack Unit 0, enter **interface gigabitethernet 0/1** from CONFIGURATION mode.

High Availability on S-Series Stacks

S-Series stacks have primary and secondary management units analogous to Dell Force10 Route Processor Modules (Message 50-1). The management units synchronize the running configuration and protocol states so that the system fails over in the event of a hardware or software fault on the primary. In such an event, or when the primary is removed, the secondary unit becomes the stack manager and FTOS elects a new secondary. FTOS resets the failed management unit, and once online, it becomes a member unit; the remaining members remain online.

Figure 50-1. S-Series Stack Manager Redundancy

```
Stack#show redundancy

--  Stack-unit Status  --
-----------------------------------------------
 Mgmt ID:                          0
 Stack-unit ID:                    1
 Stack-unit Redundancy Role:       Primary
 Stack-unit State:                 Active
 Stack-unit SW Version:            7.8.1.0
 Link to Peer:                     Up

--  PEER Stack-unit Status  --
-----------------------------------------------
 Stack-unit State:                 Standby
 Peer stack-unit ID:               2
 Stack-unit SW Version:            7.8.1.0

--  Stack-unit Redundancy Configuration  --
-----------------------------------------------
 Primary Stack-unit:               mgmt-id    0
 Auto Data Sync:                   Full
 Failover Type:                    Hot Failover
 Auto reboot Stack-unit:           Enabled
 Auto failover limit:              3 times in 60 minutes

--  Stack-unit Failover Record  --
-----------------------------------------------
 Failover Count:                   0
 Last failover timestamp:          None
 Last failover Reason:             None
 Last failover type:               None

--  Last Data Block Sync Record:  --
-----------------------------------------------
 Stack Unit Config:         succeeded  Mar 24 2009 20:35:14
   Start-up Config:         failed  Mar 24 2009 20:35:14
 Runtime Event Log:         succeeded  Mar 24 2009 20:35:14
    Running Config:         succeeded  Mar 24 2009 20:35:14
           ACL Mgr:         succeeded  Mar 24 2009 20:35:14
```

Management Unit Selection on S-Series Stacks

FTOS has a selection algorithm to decide which stack units will be the primary and secondary management units. During the bootup of a single unit or the stack, FTOS compares the priority values of all of the units in the stack and elects the unit with the numerically highest priority the primary, and the next highest priority the secondary.

For example, if you add a powered standalone unit to a stack, either the standalone unit or the stack reloads (excluding the new unit), depending on which has the higher priority, the new unit or the existing stack manager. If the new unit has the higher priority, it becomes the new stack manager after the stack reloads.

All switches have a default priority of 0; if a priority tie occurs, the system with the highest MAC address supersedes, as shown in Figure 50-2.

Figure 50-2. Electing the Stack Manager

```
Stack>show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------
  0   Standby      online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Member       online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
Stack#show system stack-unit 0 | grep priority
Master priority : 0
Stack#show system stack-unit 1 | grep priority
Master priority : 0
Stack#show system stack-unit 2 | grep priority
Master priority : 0
Stack#show system stack-unit 0| grep "Burned In MAC"
Burned In MAC   : 00:01:e8:d5:ef:81
Stack#show system stack-unit 0 | grep "Burned In MAC"
Burned In MAC   : 00:01:e8:d5:ef:81
Stack#show system stack-unit 1 | grep "Burned In MAC"
Burned In MAC   : 00:01:e8:d5:f9:6f
Stack#show system stack-unit 2 | grep "Burned In MAC"
Burned In MAC   : 00:01:e8:cc:cc:cc
```

MAC Addressing on S-Series Stacks

The S-Series has three MAC addressees: the chassis MAC, interface MAC, and null interface MAC. All interfaces in the stack use the interface MAC address of the management unit (stack manager), and the chassis MAC for the stack is the master's chassis MAC. The stack continues to use the master's chassis MAC address even after a failover. The MAC address is not refreshed until the stack is reloaded and a different unit becomes the stack manager.

Note: If the removed management unit is brought up as a standalone unit or as part of a different stack, there is a possibility of MAC address collisions.

In Figure 50-3 and Figure 50-4, a standalone is added to a stack. The standalone and the stack master have the same priority, but the standalone has a lower MAC address, so the standalone reboots. In Figure 50-4 and Figure 50-5, a standalone is added to a stack. The standalone has a higher priority than the stack, so the stack (excluding the new unit) reloads.

Figure 50-3. Adding a Standalone with a Lower MAC Address to a Stack— Before

```
------------------------------STANDALONE BEFORE CONNECTION---------------------------------
Standalone#show system brief

Stack MAC : 00:01:e8:d5:ef:81

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------
  0   Management   online         S50V        S50V        7.8.1.0     52
  1   Member       not present
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
Standalone#show system | grep priority
Master priority : 0
-----------------------------------STACK BEFORE CONNECTION---------------------------------
Stack#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------
  0   Standby      online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
Stack#show system stack-unit 0 | grep priority
Master priority : 0
Stack#show system stack-unit 1 | grep priority
Master priority : 0
```

Figure 50-4. Adding a Standalone with a Lower MAC Address and Equal Priority to a Stack—After

```
------------------------------STANDALONE AFTER CONNECTION---------------------------------
Standalone#%STKUNIT0-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is present
00:20:20: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
00:20:22: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present

Going for reboot. Reason is Stack merge
[bootup messages omitted]
-----------------------------------STACK AFTER CONNECTION---------------------------------
Stack# 3w1d14h: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
3w1d14h: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 2 down - card removed
3w1d14h: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
3w1d14h: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 2 (type S50V, 52 ports)
3w1d14h: %S50V:2 %CHMGR-0-PS_UP: Power supply 0 in unit 2 is up
3w1d14h: %STKUNIT1-M:CP %CHMGR-5-STACKUNITUP: Stack unit 2 is up

Stack#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------
  0   Standby      online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Member       online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
```

Figure 50-5. Adding a Standalone with a Lower MAC Address but Higher Priority to a Stack—Before

```
------------------------------STANDALONE BEFORE CONNECTION---------------------------------
Standalone#show system brief

Stack MAC : 00:01:e8:d5:ef:81

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-----------------------------------------------------------------------------
  0   Member       not present    S50V
  1   Member       not present    S50N
  2   Management   online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
Stack#show system | grep priority
Master priority : 1
-----------------------------------STACK BEFORE CONNECTION---------------------------------
Stack##show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-----------------------------------------------------------------------------
  0   Standby      online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
Stack#show system stack-unit 0 | grep priority
Master priority : 0
Stack#show system stack-unit 1 | grep priority
Master priority : 0
```

Figure 50-6. Adding a Standalone with a Lower MAC Address but Higher Priority to a Stack—After

```
------------------------------STANDALONE AFTER CONNECTION--------------------------------
Standalone#00:18:27: %STKUNIT2-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
00:18:27: %STKUNIT2-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
00:18:40: %STKUNIT2-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 0 down - card removed
00:18:40: %STKUNIT2-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 1 down - card removed
00:19:30: %STKUNIT2-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit i
s present
00:19:30: %STKUNIT2-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
00:19:30: %STKUNIT2-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
[remaining bootup messages omitted]
-----------------------------------STACK AFTER CONNECTION--------------------------------
Stack#3w1d15h: %STKUNIT1-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is not
present
3w1d15h: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present

Going for reboot. Reason is Stack merge
3w1d15h: %STKUNIT1-M:CP %CHMGR-2-STACK_UNIT_DOWN: Stack-unit 0 down - card removed
[bootup messages omitted]
Stack#show system brief

Stack MAC : 00:01:e8:d5:ef:81

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------------
  0   Member       online         S50V        S50V        7.8.1.0     52
  1   Standby      online         S50N        S50N        7.8.1.0     52
  2   Management   online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
```

Management Access on S-Series Stacks

You can access the stack via the console port or VTY line.

- **Console access**: You may access the stack through the console port of the stack manager only. Like a standby RPM, the console port of the standby unit does not provide management capability; only a limited number of commands are available. Member units provide a severely limited set of commands, as shown in Figure 50-7.
- **Remote access**: You may access the stack with SNMP, SSH, or Telnet through any enabled, Layer 3 interface on any stack unit. There is no dedicated management port or management route table.

Figure 50-7. Accessing Non-Master Units on a Stack via the Console Port

```
------------------------------CONSOLE ACCESS ON THE STANDBY---------------------------------
Stack(standby)>?
disable                         Turn off privileged commands
enable                          Turn on privileged commands
exit                            Exit from the EXEC
show                            Show running system information
ssh-peer-stack-unit             Open a SSH connection to the peer Stack-unit
telnet-peer-stack-unit          Open a telnet connection to the peer Stack-unit
terminal                        Set terminal line parameters
Stack(standby)>show ?
calendar                        Display the hardware calendar
clock                           Display the system clock
command-history                 CLI command history
redundancy                      Current Stack unit HA status
version                         Software version
------------------------------CONSOLE ACCESS ON A MEMBER-----------------------------------
Stack(stack-member-0)#?
reset-self              Reset this unit alone
show                    Show running system information
```

Important Points to Remember

- YYou may stack up to eight S25/S50 systems in combination.
- You may stack up to 12 S60 systems
- You may stack up to 3 S4810 systems
- You may not stack different S-Series systems together (except the S25/S50)
- You may not connect 12G and 24G stack ports.
- All stack units must have the same version of FTOS.

S-Series Stacking Installation Tasks

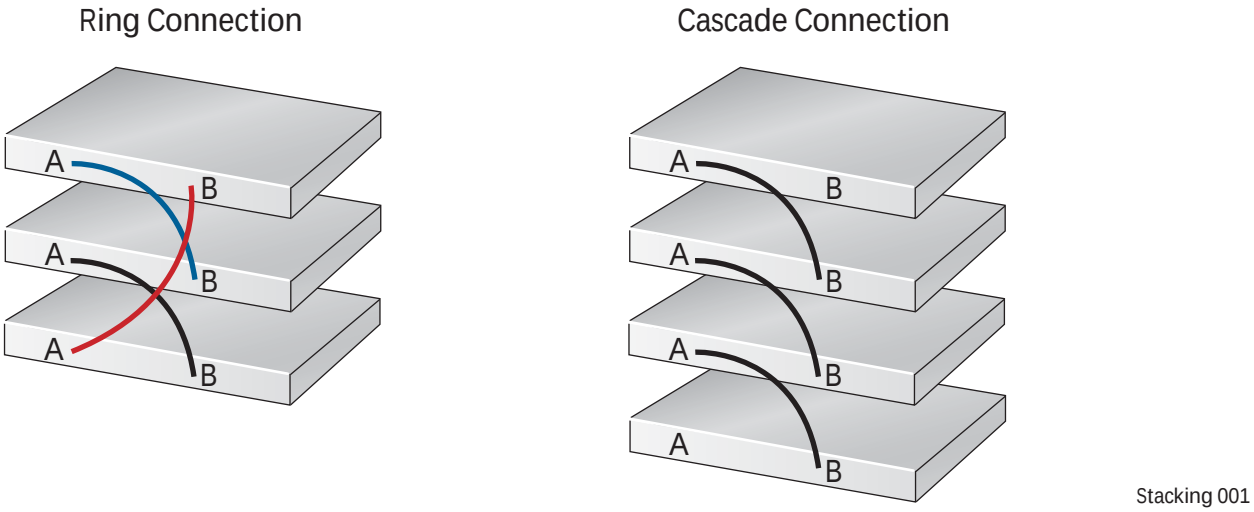
- Create an S-Series Stack on page 1026
- Add a Unit to an S-Series Stack on page 1029
- Remove a Unit from an S-Series Stack on page 1032
- Merge Two S-Series Stacks on page 1034
- Split an S-Series Stack on page 1035

Create an S-Series Stack

Stacking modules are pluggable units in the back of the unit that switch traffic between units in a stack. Units are connected using bi-directional stacking cables; if you stacking modules have two ports, it does not matter if you connect port A to B, or A to A, or B to B. Install stacking modules before powering the unit. If you install a stacking module while the unit is online, FTOS does not register the new hardware; in this case, you must reload the unit.

You may stack any combination of S-Series models that have the same FTOS version. Figure 50-8 shows two common stacking topologies, ring and cascade (also called daisy-chain). A ring topology provides some performance gains and stack integrity.

Figure 50-8. Common S-Series Stacking Topologies



Facing the rear of an S-Series unit, stack-port are numbered from left to right, beginning with the highest Ethernet port number (*n*) plus 1. For example, for a 48-port unit with two 12-Gigabyte stacking modules, the stack-ports are 49, 50, 51, and 52, starting from the left.

To add a unit to an existing stack:

Step	Task	Command Syntax	Command Mode
1	Verify that each unit has the same FTOS version prior to stacking them together.	**show version**	EXEC Privilege
2	Pre-configure unit numbers for each unit so that the stacking is deterministic upon boot up.	**stack-unit renumber**	EXEC Privilege
3	Configure the switch priority for each unit to make management unit selection deterministic.	**stack-unit priority**	CONFIGURATION
4	Connect the units using stacking cables.		
5	Power the stack one unit at a time. Start with the management unit, then the standby, followed by each of the members in order of their assigned stack number (or the position in the stack you want each unit to take). Allow each unit to completely boot, and verify that the unit is detected by the stack manager, and then power the next unit.	**show system brief**	EXEC Privilege

To display the status of the stacking ports, including the topology:

Task	Command Syntax	Command Mode
Display the stacking ports.	**show system stack-ports**	EXEC Privilege

Figure 50-9 shows a daisy-chain topology. Figure 50-10 shows the same stack converted to a ring by connecting stack-port 2/51 to 0/51; you may rearrange the stacking cables without triggering a unit reset, so long as the stack manager is never disconnected from the stack.

Figure 50-9. Displaying the S-Series Stacking Topology

```
Stack#show system stack-ports
Topology: Daisy chain
Interface  Connection    Link Speed      Admin     Link     Trunk
                           (Gb/s)        Status    Status   Group
-----------------------------------------------------------------
 0/51                      12             up        down
 0/52        1/50          12             up        up
 1/49        2/52          12             up        up
 1/50        0/52          12             up        up
 2/51                      12             up        down
 2/52        1/49          12             up        up
```

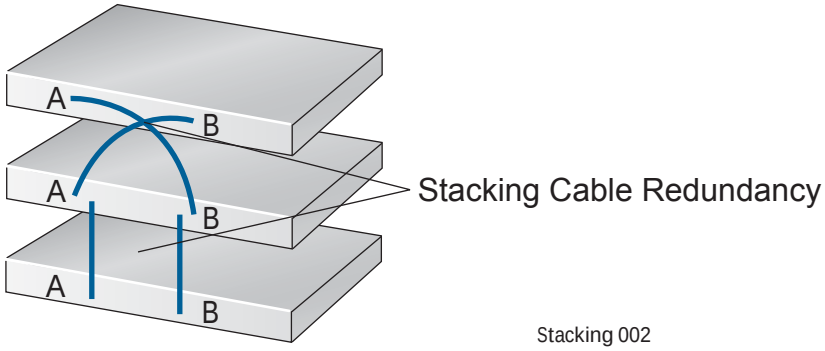
Figure 50-10. Displaying the S-Series Stacking Topology

```
FTOS#show system stack-ports
Topology: Ring
Interface  Connection    Link Speed      Admin     Link     Trunk
                           (Gb/s)        Status    Status   Group
-----------------------------------------------------------------
 0/51        2/51          12             up        up
 0/52        1/50          12             up        up
 1/49        2/52          12             up        up
 1/50        0/52          12             up        up
 2/51        0/51          12             up        up
 2/52        1/49          12             up        up
```

Stacking Cable Redundancy

You can connect two units with two stacking cables as shown in, in case of a stacking port, module, or cable failure. Removal of only one of the cables does not trigger a reset.

Figure 50-11. Stacking Cable Redundancy



LED Status Indicators on an S-Series Stack

The stack unit is displayed in an LED panel on the front of each switch. Each panel displays the stack unit ID — from 0 through 7 — and:

- **A** for the management unit
- **B**for the standby management unit
- **0** for stack members

Add a Unit to an S-Series Stack

If you are adding units to a stack, you can either:

- allow FTOS to automatically assign the new unit a position in the stack, or
- manually determine each units position in the stack by configuring each unit to correspond with the stack before connecting it

Three configurable system variables affect how a new unit joins a stack: priority, stack number, and provision.

- Depending on which has the higher priority, either the standalone unit or the entire stack reloads (excluding the new unit). If the new unit has the higher priority, it becomes the new stack manager and the stack reloads, as shown in Figure 50-3, Figure 50-4, Figure 50-5, and Figure 50-6.
- If you add a unit that has a stack number that conflicts with the stack, the stack assigns the first available stack number, as shown in Figure 50-12 and Figure 50-13.
- If the stack has a provision for the stack-number that will be assigned to the new unit, the provision must match the unit type, or FTOS generates a type mismatch error, as show in Figure 50-14 and Figure 50-15.

After the new unit loads, it synchronizes its running and startup configurations with the stack.

To manually assign a new unit a position in the stack:

Step	Task	Command Syntax	Command Mode
1	While the unit is unpowered, install stacking modules in the new unit.		
2	On the stack, determine the next available stack-unit number, and the management priorityity of the management unit.	**show system brief** **show system stack-unit**	EXEC Privilege
3	Create a virtual unit and assign it the next available stack-unit number.	**stack-unit provision**	CONFIGURATION
4	On the new unit, number it the next available stack-unit number.	**stack-unit renumber**	EXEC Privilege
5	(OPTIONAL) On the new unit, assign a management priority based on whether you want the new unit to be the stack manager.	**stack-unit priority**	CONFIGURATION
6	Connect the new unit to the stack using stacking cables.		

Figure 50-12. Adding a Stack Unit with a Conflicting Stack Number—Before

```
-----------------------STANDALONE BEFORE CONNECTION---------------------------------
Standalone#show system brief
Stack MAC : 00:01:e8:d5:ef:81
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------------
  0   Member       not present    S50V
  1   Management   online         S50V        S50V        7.8.1.0     52
  2   Member       not present
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
----------------------------STACK BEFORE CONNECTION---------------------------------
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------------
  0   Member       not present
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
```

Figure 50-13. Adding a Stack Unit with a Conflicting Stack Number—After

```
-----------------------STANDALONE AFTER CONNECTION--------------------------------
00:08:45: %STKUNIT1-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is present
00:08:45: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
00:08:47: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
Going for reboot. Reason is Stack merge
[bootup messages omitted]
Stack(stack-member-0)#
----------------------------STACK AFTER CONNECTION--------------------------------
Stack#21:27:22: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
21:27:39: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 0 down - card removed
21:28:24: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
21:28:33: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 0 (type S50V, 52 ports)
21:28:33: %S50V:0 %CHMGR-0-PS_UP: Power supply 0 in unit 0 is up
21:28:34: %STKUNIT1-M:CP %CHMGR-5-STACKUNITUP: Stack unit 0 is up
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------
  0   Member       online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
```

Figure 50-14. Adding a Stack Unit with a Conflicting Stack Provision—Before

```
-----------------------STANDALONE BEFORE CONNECTION-------------------------------
Standalone#show system brief
Stack MAC : 00:01:e8:d5:ef:81
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------
  0   Management   online         S50V        S50V        7.8.1.0     52
  1   Member       not present    S50N
  2   Member       not present    S50V
  3   Member       not present    S50V
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
----------------------------STACK BEFORE CONNECTION-------------------------------
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------
  0   Member       not present    S25N
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
```

Figure 50-15. Adding a Stack Unit with a Conflicting Stack Number—After

```
-----------------------STANDALONE AFTER CONNECTION--------------------------------
01:38:34: %STKUNIT0-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is present
01:38:34: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
01:38:34: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 presentGoing for reboot.
Reason is Stack merge
Going for reboot. Reason is Stack merge
[bootup messages omitted]
Stack(stack-member-0)#
----------------------------STACK AFTER CONNECTION--------------------------------
23:11:25: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
23:11:40: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 0 down - card removed
23:12:25: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
23:12:34: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 0 (type S50V, 52 ports)
23:12:34: %STKUNIT1-M:CP %CHMGR-3-STACKUNIT_MISMATCH: Mismatch: Stack unit 0 is type S50V -
type S25N required
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------
  0   Member       type mismatch  S25N        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
```

Remove a Unit from an S-Series Stack

The running-configuration and startup-configuration are synchronized on all stack units. A stack member that is disconnected from the stack maintain this configuration.

To remove a stack member from the stack, disconnect the stacking cables from the unit. You may do this at any time, whether the unit is powered or unpowered, online or offline. Note that if you remove a unit in the middle of the stack, the stack will be split into multiple parts, and each will form a new stack according to the stacking algorithm described throughout this chapter.

Figure 50-16. Removing a Stack Member—Before

```
---------------------------STANDALONE BEFORE DISCONNECTION---------------------------------
Standalone(stack-member-2)#?
reset-self              Reset this unit alone
show                    Show running system information
Standalone(stack-member-2)#show ?
version                 Software version
--------------------------------STACK BEFORE DISCONNECTION---------------------------------
Stack#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------------
  0   Standby      online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Member       online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
```

Figure 50-17. Removing a Stack Member—After

```
--------------------------STANDALONE AFTER DISCONNECTION---------------------------------
Standalone(stack-member-2)#
                          Going for reboot. Reason is Stack split
[bootup messages omitted]
Stack#show system brief

Stack MAC : 00:01:e8:d5:ef:81

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-----------------------------------------------------------------------------
  0   Member       not present    S50V
  1   Member       not present    S50N
  2   Management   online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
[output omitted]
--------------------------------STACK AFTER DISCONNECTION--------------------------------
Stack#3w1d15h: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 2 down - card removed
3w1d15h: %STKUNIT1-M:CP %IFMGR-1-DEL_PORT: Removed port: Gi 2/1-48
3w1d15h: %STKUNIT0-S:CP %IFMGR-1-DEL_PORT: Removed port: Gi 2/1-48
Stack#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-----------------------------------------------------------------------------
  0   Standby      online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Member       not present    S50V
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
```

Merge Two S-Series Stacks

You may merge two stacks while they are powered and online. To merge two stacks, connect one stack to the other using stacking cables. You may not connect 12G and 24G stack ports.

- FTOS selects a primary stack manager from the two existing mangers.
- FTOS resets all the units in the losing stack, and they all become stack members.
- If there is no unit numbering conflict, the stack members retain their previous unit numbers. Otherwise, the stack manager assigns new unit numbers, based on the order that they come online.
- The stack manager overwrites the startup and running config on the losing stack members with its own.

Split an S-Series Stack

To split a stack, unplug the desired stacking cables.You may do this at any time, whether the stack is powered or unpowered, and the units are online or offline. Each portion of the split stack retains the startup and running configuration of the original stack.

For a parent stack that is split into two child stacks, A and B, each with multiple units:

- If one of the new stacks receives the primary and the secondary management units, it is unaffected by the split.
- If one of the new stacks receives only the primary management unit, that units remains the stack manager, and FTOS elects a new secondary management unit.
- If one of the new stacks receives only the secondary management unit, it becomes the primary management, and FTOS elects a new secondary management unit.
- If one of the new stacks receives neither the primary nor the secondary management unit, the stack is reset so that a new election can take place.

S-Series Stacking Configuration Tasks

Assign Unit Numbers to Units in an S-Series Stack

Each unit in the stack has a stack number that is either assigned by you or FTOS. Units are numbered from 0 to 7. Stack numbers are stored in NVRAM and are preserved upon reload.

Task	Command Syntax	Command Mode
Assign a stack-number to a unit.	**stack-unit renumber**	EXEC Privilege

Note: Renumbering the stack manager triggers a failover, as shown in Message 1.

Message 1 Renumbering the Stack Manager

```
Renumbering master unit will reload the stack. Proceed to renumber [confirm yes/no]: yes
```

Create a Virtual Stack Unit on an S-Series Stack

Use virtual stack units to configure ports on the stack before adding a new unit, or to prevent FTOS from assigning a particular stack-number.

Task	Command Syntax	Command Mode
Create a virtual stack unit.	**stack-unit provision**	CONFIGURATION

Display Information about an S-Series Stack

Task	Command Syntax	Command Mode
Display for stack-identity, status, and hardware information on every unit in a stack (Figure 50-18).	**show system**	EXEC Privilege
Display most of the information in **show system**, but in a more convenient tabular form (Figure 50-19).	**show system brief**	EXEC Privilege
Display the same information in **show system**, but only for the specified unit (Figure 50-19).	**show system stack-unit**	EXEC Privilege
Display topology and stack link status for the entire stack. The available options separate the **show system stack-port** output into topology information from link status information (Figure 50-19).	**show system stack-ports** [**status** | **topology**]	EXEC Privilege

Figure 50-18. Displaying Information about an S-Series Stack—show system

```
FTOS#show system

Stack MAC : 00:01:e8:d5:f9:6f

--  Unit 0 --
Unit Type       : Member Unit
Status          : online
Next Boot       : online
Required Type   : S50V - 48-port E/FE/GE with POE (SB)
Current Type    : S50V - 48-port E/FE/GE with POE (SB)
Master priority : 0
Hardware Rev    : 2.0
Num Ports       : 52
Up Time         : 30 min, 7 sec
FTOS Version    : 7.8.1.0
Jumbo Capable   : yes
POE Capable     : yes
Burned In MAC   : 00:01:e8:d5:ef:81
No Of MACs      : 3

--  Module 0 --
Status          : not present

--  Module 1 --
Status          : online
Module Type     : S50-01-12G-2S     - 2-port 12G Stacking (SB)
Num Ports       : 2
Hot Pluggable   : no

--  Power Supplies  --
Unit   Bay   Status       Type
--------------------------------------------------------------------------------
  0     0     up            AC
  0     1     absent

--  Fan  Status  --
Unit   TrayStatus  Speed    Fan0    Fan1    Fan2    Fan3    Fan4    Fan5
-------------------------------------------------------------------------------
 0     up          low      up      up      up      up      up      up
```

Figure 50-19. Displaying Information about an S-Series Stack—show system brief

```
FTOS#show system brief

Stack MAC : 00:01:e8:d5:f9:6f

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------------
  0   Member       online         S50V        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present

--  Module Info  --
Unit  Module No   Status        Module Type        Ports
-------------------------------------------------------------------------------
  0   0           not present   No Module           0
  0   1           online        S50-01-12G-2S       2
  1   0           online        S50-01-12G-2S       2
  1   1           not present   No Module           0
  2   0           not present   No Module           0
  2   1           online        S50-01-12G-2S       2

--  Power Supplies  --
Unit   Bay   Status       Type
-------------------------------------------------------------------------------
  0     0     up            AC
  0     1     absent
  1     0     absent
  1     1     up            DC
  2     0     up            AC
  2     1     absent

--  Fan  Status  --
Unit   TrayStatus  Speed    Fan0    Fan1    Fan2    Fan3    Fan4    Fan5
-------------------------------------------------------------------------------
 0     up          low      up      up      up      up      up      up
 1     up          low      up      up      up      up      up      up
 2     up          low      up      up      up      up      up      up
```

Figure 50-20. Displaying Information about an S-Series Stack—show system stack-ports

```
FTOS#show system stack-ports
Topology: Daisy chain
Interface  Connection    Link Speed      Admin     Link     Trunk
                           (Gb/s)        Status    Status   Group
-----------------------------------------------------------------
 0/51                      12             up        down
 0/52        2/51          12             up        up
 1/49        2/52          12             up        up
 1/50                      12             up        down
 2/51        0/52          12             up        up
 2/52        1/49          12             up        up
```

Influence Management Unit Selection on an S-Series Stack

Stack Priority is the system variable that FTOS uses to determine which units in the stack will be the primary and secondary management units. If multiple units tie for highest priority, then the unit with the highest MAC address prevails.

If management was determined by priority only, a change in management occurs when:

- you powered down, or offline the management unit, or a failover occurs
- you disconnect the management unit from the stack

Task	Command Syntax	Command Mode
Influence the selection of the stack management units. The unit with the numerically highest priority is elected the primary management unit, and the unit with the second highest priority is the secondary management unit. Default: 0 Range: 1-14	**stack-unit priority**	CONFIGURATION

Manage Redundancy on an S-Series Stack

Task	Command Syntax	Command Mode
Reset the current management unit, and make the secondary management unit the new primary. A new secondary is elected, and when the former stack manager comes back online, it becomes a member unit.	**redundancy force-failover stack-unit**	EXEC Privilege
Prevent the stack manager from rebooting after a failover. This command does not affect a forced failover, manual reset, or a stack-link disconnect.	**redundancy disable-auto-reboot stack-unit**	CONFIGURATION
Display redundancy information.	**show redundancy**	EXEC Privilege

Reset a Unit on an S-Series Stack

You may reset any stack unit except for the master (Message 2).

Message 2 Master Reset Disallowed

```
% Error: Reset of master unit is not allowed.
```

Task	Command Syntax	Command Mode
Reload a stack-unit	**reset stack-unit** *0-7*	EXEC Privilege
Reload a member unit, from the unit itself	**reset-self**	EXEC Privilege
Reset a stack-unit when the unit is in a problem state.	**reset stack-unit** *0-7* **hard**	EXEC Privilege

Monitor an S-Series Stack with SNMP

S-Series supports the following tables in *f10-ss-chassis.mib* for stack management through SNMP:

- chStackUnitTable
- chSysStackPortTable

Troubleshoot an S-Series Stack

Recover from Stack Link Flaps

S-Series Stack Link Integrity Monitoring enables units to monitor their own stack ports, and disable any stack port that flaps five times within 10 seconds. FTOS displays console messages the local and remote members of a flapping link, and on the primary and secondary management units as KERN-2-INT messages if the flapping port belongs to either of these units.

In Figure 50-21, a stack-port on the manager flaps. The remote member, Member 2, displays a console message, and the manager and standby display KERN-2-INT messages.

To re-enable the downed stack-port, power cycle the offending unit.

Figure 50-21. Recovering from a Stack Link Flapping Error

```
-------------------------------------MANAGMENT UNIT-------------------------------------
Error: Stack Port 50 has flapped 5 times within 10 seconds.Shutting down this st
ack port now.
Error: Please check the stack cable/module and power-cycle the stack.
10:55:20: %STKUNIT1-M:CP %KERN-2-INT: Error: Stack Port 50 has flapped 5 times w
ithin 10 seconds.Shutting down this stack port now.
10:55:20: %STKUNIT1-M:CP %KERN-2-INT: Error: Please check the stack cable/module
 and power-cycle the stack.
-------------------------------------STANDBY UNIT-------------------------------------
10:55:18: %STKUNIT1-M:CP %KERN-2-INT: Error: Stack Port 50 has flapped 5 times within 10
seonds.Shutting down this stack port now.
10:55:18: %STKUNIT1-M:CP %KERN-2-INT: Error: Please check the stack cable/module
 and power-cycle the stack.
-------------------------------------MEMBER 2-------------------------------------
Error: Stack Port 51 has flapped 5 times within 10 seconds.Shutting down this stack port now.
Error: Please check the stack cable/module and power-cycle the stack.
```

Recover from a Card Problem State on an S-Series Stack

If a unit added to a stack has a different FTOS version, the unit does not come online, and FTOS cites a card problem error, as shown in Figure 50-22. To recover, disconnect the new unit from the stack, change the FTOS version to match the stack, and then reconnect it to the stack.

Figure 50-22. Recovering from a Card Problem Error on an S-Series Stack

```
F
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-----------------------------------------------------------------------------
  0   Member       card problem   S25N        unknown     7.7.1.1     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
```

Recover from a Card Mismatch State on an S-Series Stack

A card mismatch occurs if the stack has a provision for the lowest available stack number which does not match the model of a newly added unit (Figure 50-23). To recover, disconnect the new unit. Then, either:

- remove the provision from the stack, then reconnect the standalone unit, or
- renumber the standalone unit with another available stack number on the stack.

Figure 50-23. Recovering from a Card Mismatch State on an S-Series Stack

```
---------------------------------STANDALONE UNIT BEFORE-----------------------------------
Standalone#show system brief
Stack MAC : 00:01:e8:d5:ef:81
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------
  0   Management   online         S50V        S50V        7.8.1.0     52
  1   Member       not present    S50N
  2   Member       not present    S50V
  3   Member       not present    S50V
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
-------------------------------------STACK BEFORE-----------------------------------------
Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------
  0   Member       not present    S25N
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
---------------------------------STANDALONE UNIT AFTER-------------------------------------
01:38:34: %STKUNIT0-M:CP %POLLMGR-2-ALT_STACK_UNIT_STATE: Alternate Stack-unit is present
01:38:34: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
Going for reboot. Reason is Stack merge
01:38:34: %STKUNIT0-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
-------------------------------------STACK AFTER------------------------------------------
23:11:25: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
23:11:40: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 0 down - card removed
23:12:25: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 0 present
23:12:34: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 0 (type S50V, 52 ports)
23:12:34: %STKUNIT1-M:CP %CHMGR-3-STACKUNIT_MISMATCH: Mismatch: Stack unit 0 is type S50V -
type S25N required

Stack#show system brief
Stack MAC : 00:01:e8:d5:f9:6f
--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------
  0   Member       type mismatch  S25N        S50V        7.8.1.0     52
  1   Management   online         S50N        S50N        7.8.1.0     52
  2   Standby      online         S50V        S50V        7.8.1.0     52
  3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present
```

51

Broadcast Storm Control

Broadcast Storm Control is supported on platforms: C E S

This chapter contains the following configuration topics:

Storm Control Overview

FTOS Storm Control is a preventative measure against unexpectedly high rates of broadcast or multicast packets; these traffic bursts are called *storms*. If the rate of these packets on ingress or egress exceeds a user-defined threshold, FTOS, when configured, can suppress forwarding for these packets until the packet rate falls back to the configured limit.

Situations that Can Lead to Packet Storms

- **Layer 2 Broadcasts—**A switch might not have an entry in its MAC address table that matches a packet's destination MAC. In this case, the switch floods the packet on the VLAN. These packets are called unknown-packets; they cause unnecessary extra traffic and can reduce network performance.
- **Layer 3 Broadcast Packets**—There are two types of Layer 3 broadcast packets: the all-hosts broadcast, the IP address of which is 255.255.255.255, and the subnet broadcast address, the address of which has the host portion of the address set to all ones; for example, 10.11.1.255/24 is the broadcast address for the 10.11.1.0 network. Some protocols utilize broadcasts more than others and so storm control might be useful to prevent congestion.
- **Multicast Packets—**Multicast packets are packets that use a special range of MAC and IP addresses to send packets to a group of hosts, rather than a single host. Some multicast applications can cause excessive bandwidth consumption, and storm control can be used (in conjunction with IGMP Snooping) to limit multicast traffic.

Implementation Information

- Storm Control is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.
- The percentage of storm control is calculated based on the advertised rate of the line card, not by the speed setting of the interface.
- Do not apply per-VLAN QoS on an interface that has Storm Control enabled either on an interface or globally.
- On E-Series, bi-directional traffic (unknown-unicast and broadcast) along with egress Storm Control causes the configured traffic rates to be split between the involved ports. The percentage of traffic that each port receives after the split is not predictable. Ports are affected whether they are in the same or different port-pipes or line cards.

Broadcast Storm Control

FTOS offers Layer 3 and Layer 2 broadcast storm control.

Layer 3 Broadcast Storm Control

Layer 3 Storm Control suppresses all-hosts and subnet broadcasts if they exceed a user-defined packet rate.

You can enable Storm Control for Layer 3 broadcasts from INTERFACE mode, CONFIGURATION mode, or both. Each option has a different result.

- From INTERFACE mode: Storm Control limits ingress broadcast traffic on a single interface.
- From CONFIGURATION mode:
  - On the E-Series, Storm Control limits ingress and egress broadcast traffic on all interfaces.
  - On the C-Series and S-Series, Storm Control limits only ingress broadcast traffic, but still on all interfaces.
- From INTERFACE and CONFIGURATION mode: the INTERFACE mode configuration overrides the CONFIGURATION mode configuration.

FTOS Behavior: On the E-Series, when broadcast Storm Control is enabled on an interface or globally on ingress and DSCP marking is enabled and set to DSCP=1 is configured for data traffic, the traffic is queued to Queue 1 instead of Queue 0.

Enable Broadcast Storm Control on an Interface

Enabling Storm control on an interface affects only ingress broadcasts.

Task	Command Syntax	Command Mode
On the E-Series, suppress Layer 3 all-hosts and subnet broadcasts on ingress if they exceed a user-defined limit.	**storm-control broadcast** *percentage partial-percentage* [**in** | **out**]	INTERFACE
On the C-Series and S-Series, suppress Layer 3 all-host and subnet broadcasts on ingress if they exceed a user-defined limit.	**storm-control broadcast** *packets-per-second* **in**	INTERFACE

Enable Broadcast Storm Control on all Interfaces

The result of enabling Storm Control globally varies by platform:

- On the E-Series, enabling Storm Control from CONFIGURATION mode limits ingress and egress broadcast traffic on all interfaces.
- On the C-Series and S-Series, enabling Storm Control from CONFIGURATION mode limits only ingress broadcast traffic, but still on all interfaces.

Task	Command Syntax	Command Mode
On the E-Series, suppress Layer 3 all-hosts and subnet broadcasts on *ingress and egress* if they exceed a user-defined limit.	**storm-control broadcast** *percentage partial-percentage* [**in** | **out**]	CONFIGURATION
On the C-Series and S-Series, suppress Layer 3 all-host and subnet broadcasts on ingress if they exceed a user-defined limit.	**storm-control broadcast** *packets-per-second* **in**	CONFIGURATION

Layer 2 Broadcast Storm Control

Unknown-unicast packets are those for which the switch has no entry in its MAC address table for the packet destination MAC. In this case the switch broadcasts (floods) these packets on the VLAN. This extra traffic unnecessary and can reduce performance.

FTOS Behavior: On the E-Series, if unknown-unicast Storm Control is enabled on an interface or globally on the ingress, and DSCP marking is enabled and set to DSCP=2 for data traffic, the traffic is queued to Queue 2 instead of Queue 0.

Task	Command Syntax	Command Mode
On the E-Series, suppress unknown-unicast packets if they exceed a user-defined limit.	**storm-control unknown-unicast** *percentage partial-percentage* [**in** | **out**]	CONFIGURATION
On the C-Series and S-Series, unknown-unicast packets on ingress if they exceed a user-defined limit.	**storm-control unknown-unicast** *packets-per-second* **in**	CONFIGURATION

Multicast Storm Control

Multicast Storm Control is supported only on platforms: C S

Task	Command Syntax	Command Mode
Suppress multicast packets if they exceed a user-defined limit.	**storm-control multicast** *packets-per-second* **in**	CONFIGURATION

Storm Control Show Commands

The **show storm-control** commands display the current storm control configuration of the entire Dell Force10 platform. These show commands are accessed from within either the EXEC or EXEC Privilege mode, and consist of the following three commands:

- **show storm-control broadcast**
- **show storm-control multicast**
- **show storm-control unknown-unicast**

These show commands can be entered with or without the *interface* option. Without the *interface* option, storm control configuration information for the entire Dell Force10 platform is displayed.

To display the storm control configuration of a particular interface, use the *interface* option along with the *interface-type* keyword and the slot and port information.

The following is a list of the *interface-type* keywords:

- **Fastethernet** — for a fast Ethernet interface
- **GigabitEthernet** — for a 1-Gigabit Ethernet interface
- **TenGigabitEthernet** — for a 10-Gigabit Ethernet interface

The following example uses the **show storm-control broadcast** command to display the storm control configuration on a Gigabit Ethernet interface on port 11, in slot 11 of an E-Series platform.

```
FTOS#show storm-control broadcast gigabitethernet 11/11

 Broadcast storm control configuration

Interface       Direction       Percentage       Wred Profile
-----------------------------------------------------------------
Gi 11/11         Ingress               5.6

Gi 11/11         Egress                5.6                -
FTOS#
```

The following example displays the output of the **show storm-control broadcast** command on a C-Series platform.

```
FTOS#show storm-control broadcast gigabitethernet 3/24

 Broadcast storm control configuration

Interface       Direction       Packets/Second
-----------------------------------------------
Gi 3/24          Ingress               1000

FTOS#
```

The following example displays the output from the **show storm-control multicast** command on a S-Series platform

```
FTOS#show storm-control multicast gigabitethernet 1/0

 Multicast storm control configuration

Interface       Direction       Packets/Second
-----------------------------------------------
Gi 1/0           Ingress               5

FTOS#
```

The following example displays the output from the **show storm-control unknown-unicast** command on a C-Series platform

```
FTOS#show storm-control unknown-unicast gigabitethernet 11/1

 Unknown-unicast storm control configuration

Interface       Direction       Percentage       Wred Profile
-----------------------------------------------------------------
Gi 11/1          Ingress               5.9        -

Gi 11/1          Egress                5.7        w8

FTOS#
```

52

Spanning Tree Protocol

Spanning Tree Protocol is supported on platforms: C E S

STP is supported on the E-Series ExaScale platform with FTOS 8.1.1.2 and later.

Protocol Overview

Spanning Tree Protocol (STP) is a Layer 2 protocol—specified by IEEE 802.1d—that eliminates loops in a bridged topology by enabling only a single path through the network. By eliminating loops, the protocol improves scalability in a large network and enables you to implement redundant paths, which can be activated upon the failure of active paths. Layer 2 loops, which can occur in a network due to poor network design and without enabling protocols like xSTP, can cause unnecessarily high switch CPU utilization and memory consumption.

FTOS supports three other variations of Spanning Tree, as shown here:

Table 52-1. FTOS Supported Spanning Tree Protocols

Dell Force10 Term	IEEE Specification
Spanning Tree Protocol	802.1d
Rapid Spanning Tree Protocol	802.1w
Multiple Spanning Tree Protocol	802.1s
Per-VLAN Spanning Tree Plus	Third Party

Configuring Spanning Tree

Configuring Spanning Tree is a two-step process:

1. Configure interfaces for Layer 2. See page 1051.
2. Enable Spanning Tree Protocol. See page 1052.

Related Configuration Tasks

- Adding an Interface to the Spanning Tree Group on page 1054
- Removing an Interface from the Spanning Tree Group on page 1054
- Modifying Global Parameters on page 1055
- Modifying Interface STP Parameters on page 1056
- Enabling PortFast on page 1056
- Preventing Network Disruptions with BPDU Guard on page 1057
- STP Root Selection on page 1059
- SNMP Traps for Root Elections and Topology Changes on page 1063
- Configuring Spanning Trees as Hitless on page 1064

Important Points to Remember

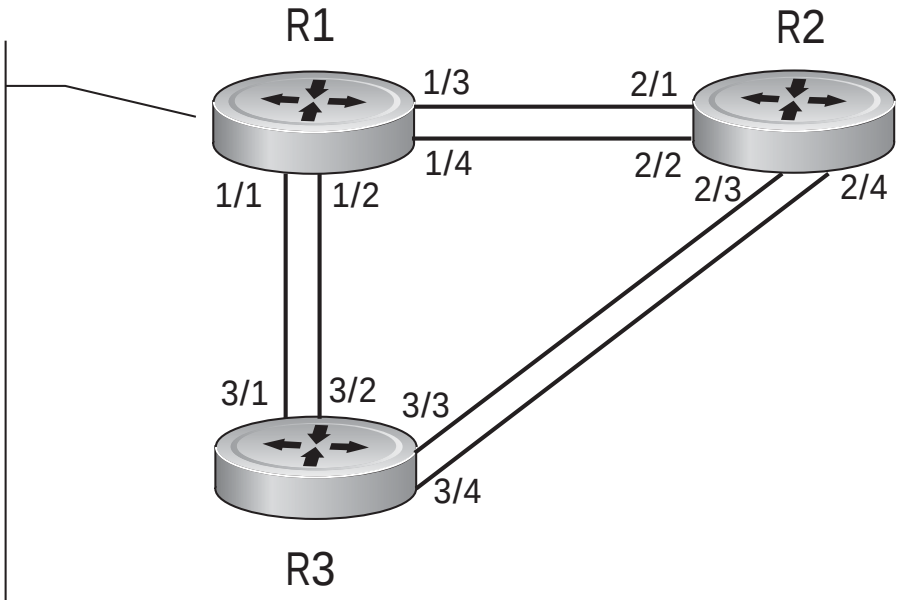
- Spanning Tree Protocol (STP) is disabled by default.
- FTOS supports only one Spanning Tree instance (0). For multiple instances, you must enable MSTP, or PVST+. You may only enable one flavor of Spanning Tree at any one time.
- All ports in VLANs and all enabled interfaces in Layer 2 mode are automatically added to the Spanning Tree topology at the time you enable the protocol.
- To add interfaces to the Spanning Tree topology after STP is enabled, enable the port and configure it for Layer 2 using the command switchport.
- The IEEE Standard 802.1D allows eight bits for port ID and eight bits for priority. However, the eight bits for port ID provide port IDs for only 256 ports and the C-Series can contain 336 ports. To accommodate the increased number of ports, FTOS uses four bits from priority field in the port ID field.This implementation affects the Bridge MIB (RFC 1493), and you must interpret objects such as *dot1dStpPortDesignatedPort* object by using the first four bits as the priority and the last 12 bits as the port ID.

Configuring Interfaces for Layer 2 Mode

All interfaces on all switches that will participate in Spanning Tree must be in Layer 2 mode and enabled.

Figure 52-1. Example of Configuring Interfaces for Layer 2 Mode

```
R1(conf)# int range gi 1/1 - 4
R1(conf-if-gi-1/1-4)# switchport
R1(conf-if-gi-1/1-4)# no shutdown
R1(conf-if-gi-1/1-4)#show config
!
interface GigabitEthernet 1/1
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/2
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/3
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 1/4
 no ip address
 switchport
 no shutdown
```



To configure the interfaces for Layer 2 and then enable them:

Step	Task	Command Syntax	Command Mode
1	If the interface has been assigned an IP address, remove it.	**no ip address**	INTERFACE
2	Place the interface in Layer 2 mode.	**switchport**	INTERFACE
3	Enable the interface.	**no shutdown**	INTERFACE

Verify that an interface is in Layer 2 mode and enabled using the **show config** command from INTERFACE mode.

Figure 52-2. Verifying Layer 2 Configuration

```
FTOS(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
 no ip address
 switchport
no shutdown
FTOS(conf-if-gi-1/1)#
```

Indicates that the interface is in Layer 2 mode

Enabling Spanning Tree Protocol Globally

Spanning Tree Protocol must be enabled globally; it is not enabled by default.

To enable Spanning Tree globally for all Layer 2 interfaces:

Step	Task	Command Syntax	Command Mode
1	Enter the PROTOCOL SPANNING TREE mode.	**protocol spanning-tree 0**	CONFIGURATION
2	Enable Spanning Tree.	**no disable**	PROTOCOL SPANNING TREE

Note: To disable STP globally for all Layer 2 interfaces, enter the **disable** command from PROTOCOL SPANNING TREE mode.

Verify that Spanning Tree is enabled using the **show config** command from PROTOCOL SPANNING TREE mode.

Figure 52-3. Verifying STP is Enabled

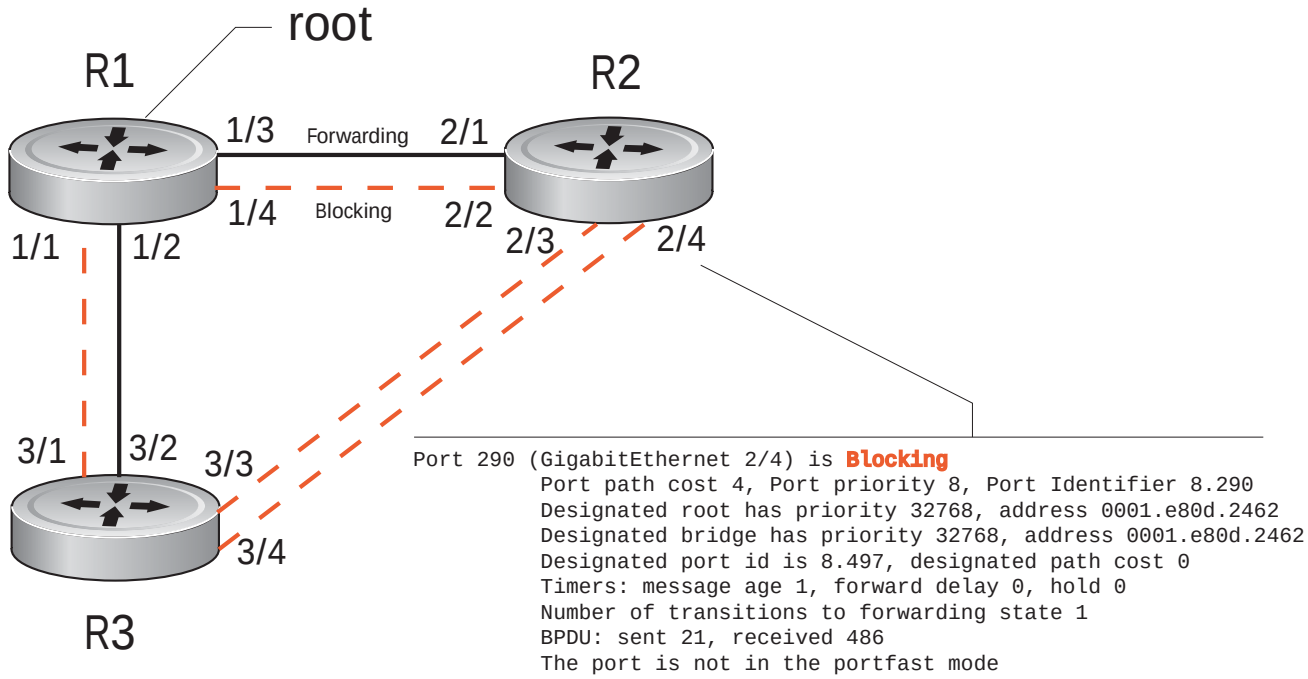
```
FTOS(conf)#protocol spanning-tree 0
FTOS(config-span)#show config
!
protocol spanning-tree 0
 no disable
FTOS#
```

Indicates that Spanning Tree is enabled

When you enable Spanning Tree, all physical, VLAN, and port-channel interfaces that are enabled and in Layer 2 mode are automatically part of the Spanning Tree topology.

- Only one path from any bridge to any other bridge participating in STP is enabled.
- Bridges block a redundant path by disabling one of the link ports.

Figure 52-4. Spanning Tree Enabled Globally



View the Spanning Tree configuration and the interfaces that are participating in STP using the **show spanning-tree 0** command from EXEC privilege mode. If a physical interface is part of a port channel, only the port channel is listed in the command output.

Figure 52-5. show spanning-tree 0 Command Example

```
R2#show spanning-tree 0
     Executing IEEE compatible Spanning Tree Protocol
          Bridge Identifier has priority 32768, address 0001.e826.ddb7
          Configured hello time 2, max age 20, forward delay 15
          Current root has priority 32768, address 0001.e80d.2462
          Root Port is 289 (GigabitEthernet 2/1), cost of root path is 4
          Topology change flag not set, detected flag not set
          Number of topology changes 3 last change occurred 0:16:11 ago
                  from GigabitEthernet 2/3
          Timers: hold 1, topology change 35
                  hello 2, max age 20, forward delay 15
          Times:  hello 0, topology change 0, notification 0, aging Normal

     Port 289 (GigabitEthernet 2/1) is Forwarding
          Port path cost 4, Port priority 8, Port Identifier 8.289
          Designated root has priority 32768, address 0001.e80d.2462
          Designated bridge has priority 32768, address 0001.e80d.2462
          Designated port id is 8.496, designated path cost 0
          Timers: message age 1, forward delay 0, hold 0
          Number of transitions to forwarding state 1
          BPDU: sent 21, received 486
          The port is not in the portfast mode

     Port 290 (GigabitEthernet 2/2) is Blocking
          Port path cost 4, Port priority 8, Port Identifier 8.290
--More--
```

Confirm that a port is participating in Spanning Tree using the **show spanning-tree 0 brief** command from EXEC privilege mode.

Figure 52-6. show spanning-tree brief Command Example

```
FTOS#show spanning-tree 0 brief
     Executing IEEE compatible Spanning Tree Protocol
          Root ID  Priority 32768, Address 0001.e80d.2462
          We are the root of the spanning tree
          Root Bridge hello time 2, max age 20, forward delay 15
          Bridge ID  Priority 32768, Address 0001.e80d.2462
          Configured hello time 2, max age 20, forward delay 15
Interface                                   Designated
 Name           PortID Prio Cost Sts Cost     Bridge ID              PortID
-------------- ------ ---- ---- --- -----    -----------------      ------
Gi 1/1          8.496    8    4 DIS     0    32768 0001.e80d.2462  8.496
Gi 1/2          8.497    8    4 DIS     0    32768 0001.e80d.2462  8.497
Gi 1/3          8.513    8    4 FWD     0    32768 0001.e80d.2462  8.513
Gi 1/4          8.514    8    4 FWD     0    32768 0001.e80d.2462  8.514
FTOS#
```

Adding an Interface to the Spanning Tree Group

To add a Layer 2 interface to the Spanning Tree topology:

Task	Command Syntax	Command Mode
Enable Spanning Tree on a Layer 2 interface.	**spanning-tree 0**	INTERFACE

Removing an Interface from the Spanning Tree Group

To remove a Layer 2 interface from the Spanning Tree topology:

Task	Command Syntax	Command Mode
Disable Spanning Tree on a Layer 2 interface.	**no spanning-tree 0**	INTERFACE

In FTOS versions prior to 7.6.1.0, the command **no spanning tree** disables Spanning Tree on the interface, however, BPDUs are still forwarded to the RPM, where they are dropped. Beginning in FTOS version 7.6.1.0, the command **no spanning tree** disables Spanning Tree on the interface, and incoming BPDUs are dropped at the line card instead of at the RPM, which frees processing resources. This behavior is called Layer 2 BPDU filtering and is available for STP, RSTP, PVST+, and MSTP.

Modifying Global Parameters

You can modify Spanning Tree parameters. The root bridge sets the values for forward-delay, hello-time, and max-age and overwrites the values set on other bridges participating in Spanning Tree.

Note: Dell Force10 recommends that only experienced network administrators change the Spanning Tree parameters. Poorly planned modification of the Spanning Tree parameters can negatively impact network performance.

Table 52-2 displays the default values for Spanning Tree.

Table 52-2. STP Default Values

STP Parameter		Default Value
Forward Delay		15 seconds
Hello Time		2 seconds
Max Age		20 seconds
Port Cost	100-Mb/s Ethernet interfaces	19
	1-Gigabit Ethernet interfaces	4
	10-Gigabit Ethernet interfaces	2
	Port Channel with 100 Mb/s Ethernet interfaces	18
	Port Channel with 1-Gigabit Ethernet interfaces	3
	Port Channel with 10-Gigabit Ethernet interfaces	1
Port Priority		8

To change STP global parameters:

Task	Command Syntax	Command Mode
Change the forward-delay parameter (the wait time before the interface enters the *forwarding* state). • Range: 4 to 30 • Default: 15 seconds	**forward-delay** *seconds*	PROTOCOL SPANNING TREE
Change the hello-time parameter (the BPDU transmission interval). **Note:** With large configurations (especially those with more ports) Dell Force10 recommends that you increase the hello-time. Range: 1 to 10 Default: 2 seconds	**hello-time** *seconds*	PROTOCOL SPANNING TREE
Change the max-age parameter (the refresh interval for configuration information that is generated by recomputing the Spanning Tree topology). Range: 6 to 40 Default: 20 seconds	**max-age** *seconds*	PROTOCOL SPANNING TREE

View the current values for global parameters using the **show spanning-tree 0** command from EXEC privilege mode. See Figure 52-5.

Modifying Interface STP Parameters

You can set the port cost and port priority values of interfaces in Layer 2 mode.

- **Port cost** is a value that is based on the interface type. The greater the port cost, the less likely the port will be selected to be a forwarding port.
- **Port priority** influences the likelihood that a port will be selected to be a forwarding port in case that several ports have the same port cost.

The default values are listed in Table 52-2.

To change the port cost or priority of an interface:

Task	Command Syntax	Command Mode
Change the port cost of an interface. Range: 0 to 65535 Default: see Table 52-2.	**spanning-tree 0 cost** *cost*	INTERFACE
Change the port priority of an interface. Range: 0 to 15 Default: 8	**spanning-tree 0 priority** *priority-value*	INTERFACE

View the current values for interface parameters using the **show spanning-tree 0** command from EXEC privilege mode. See Figure 52-5.

Enabling PortFast

The PortFast feature enables interfaces to transition to a forwarding state and start to transmit traffic approximately 30 seconds sooner.

Interfaces forward frames by default until they receive a BPDU that indicates that they should behave otherwise; they do not go through the Learning and Listening states. The **bpduguard shutdown-on-violation** option causes the interface hardware to shut down when it receives a BPDU. When only **bpduguard** is configured, although the interface is placed in an Error Disabled state when receiving the BPDU, the physical interface remains up and spanning-tree will drop packets in the hardware after a BPDU violation. BPDUs are dropped in the software after receiving the BPDU violation.

△ **Caution:** Enable PortFast only on links connecting to an end station. PortFast can cause loops if it is enabled on an interface connected to a network.

To enable PortFast on an interface:

Task	Command Syntax	Command Mode
Enable PortFast on an interface.	**spanning-tree** *stp-id* **portfast** [**bpduguard** [**shutdown-on-violation**]]	INTERFACE

Verify that PortFast is enabled on a port using the **show spanning-tree** command from the EXEC privilege mode or the **show config** command from INTERFACE mode; Dell Force10 recommends using the **show config** command, as shown in Figure 52-7.

Figure 52-7. PortFast Enabled on Interface

```
FTOS#(conf-if-gi-1/1)#show conf
!
interface GigabitEthernet 1/1
 no ip address
 switchport
 spanning-tree 0 portfast          <-- Indicates that the interface is in PortFast mode
 no shutdown
FTOS#(conf-if-gi-1/1)#
```

Preventing Network Disruptions with BPDU Guard

The Portfast (and Edgeport, in the case of RSTP, PVST+, and MSTP) feature should be configured on ports that connect to end stations. End stations do not generate BPDUs, so ports configured with Portfast/ Edgport (edgeports) do not expect to receive BPDUs. If an edge port does receive a BPDU, it likely means that it is connected to another part of the network, which can negatively effect the STP topology.

The BPDU Guard feature blocks an edge port upon receiving a BPDU to prevent network disruptions, and FTOS displays Message 1.

△ **Caution:** Do not enable Portfast BPDU guard and loop guard at the same time on a port. Enabling both features may result in a port that remains in a blocking state and prevents traffic from flowing through it. For example, when Portfast BPDU guard and loop guard are both configured:

- If a BPDU is received from a remote device, BPDU guard places the port in an err-disabled blocking state and no traffic is forwarded on the port.
- If no BPDU is received from a remote device, loop guard places the port in a loop-inconsistent blocking state and no traffic is forwarded on the port.

Enable BPDU Guard using the option **bpduguard** when enabling PortFast or EdgePort. Configure the **bpduguard shutdown-on-violation** option to cause the interface hardware to shut down when it receives a BPDU. Otherwise with only the option enabled, although the interface is placed in an Error Disabled state when receiving the BPDU, the physical interface remains up and spanning-tree will only drop *packets* after a BPDU violation.

Figure 52-8 shows a scenario in which an edgeport might unintentionally receive a BPDU. The port on the Dell Force10 system is configured with Portfast. If the switch is connected to the hub, the BPDUs that the switch generates might trigger an undesirable topology change. If BPDU Guard is enabled, when the edge port receives the BPDU, the BPDU will be dropped, the port will be blocked, and a console message will be generated.

Message 1 BPDU Guard Error

```
3w3d0h: %RPM0-P:RP2 %SPANMGR-5-BPDU_GUARD_RX_ERROR: Received Spanning Tree BPDU on BPDU guard
port. Disable GigabitEthernet 3/41.
```

Note: Note that *unless* the **shutdown-on-violation** option is enabled, spanning-tree only *drops packets* after a BPDU violation; the physical interface remains up, as shown below.

```
FTOS(conf-if-gi-0/7)#do show spanning-tree rstp brief
Executing IEEE compatible Spanning Tree Protocol
Root ID    Priority 32768, Address 0001.e805.fb07
Root Bridge hello time 2, max age 20, forward delay 15
Bridge ID    Priority 32768, Address 0001.e85d.0e90
Configured hello time 2, max age 20, forward delay 15

Interface                                        Designated
 Name      PortID   Prio Cost    Sts Cost        Bridge ID         PortID
---------- -------- ---- ------- --- ------- ------------------- --------
Gi 0/6     128.263  128  20000   FWD 20000   32768 0001.e805.fb07 128.653
Gi 0/7     128.264  128  20000   EDS 20000   32768 0001.e85d.0e90 128.264

Interface
 Name      Role   PortID   Prio Cost    Sts Cost    Link-type Edge
---------- ------ -------- ---- ------- --- ------- --------- ----
Gi 0/6     Root   128.263  128  20000   FWD 20000   P2P       No
Gi 0/7     ErrDis 128.264  128  20000   EDS 20000   P2P       No
FTOS(conf-if-gi-0/7)#do show ip int br gi 0/7
Interface                IP-Address      OK  Method Status                Protocol
GigabitEthernet 0/7      unassigned      YES Manual up                    up
```

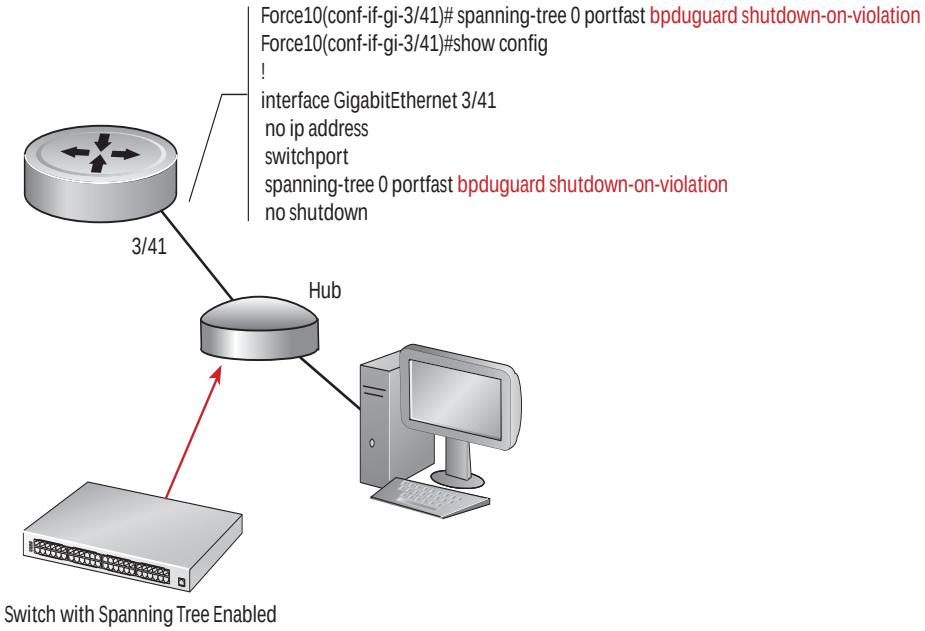
FTOS Behavior: Regarding **bpduguard shutdown-on-violation** behavior:

1. If the interface to be shutdown is a port channel then all the member ports are disabled in the hardware.
2. When a physical port is added to a port channel already in error disable state, the new member port will also be disabled in the hardware.
3. When a physical port is removed from a port channel in error disable state, the error disabled state is cleared on this physical port (the physical port will be enabled in the hardware).
4. The **reset linecard** command does not clear the error disabled state of the port or the hardware disabled state. The interface continues to be disables in the hardware.

   The error disabled state can be cleared with any of the following methods:

   - Perform an **shutdown** command on the interface.
   - Disable the **shutdown-on-violation** command on the interface (**no spanning-tree** *stp-id* **portfast** [**bpduguard |** [**shutdown-on-violation**]]).
   - Disable spanning tree on the interface (**no spanning-tree** in INTERFACE mode).
   - Disabling global spanning tree (**no spanning-tree** in CONFIGURATION mode).

Figure 52-8. Enabling BPDU Guard



To verify the Portfast BPDU loop guard configuration on a port or port-channel interface, enter the **show spanning-tree 0 guard [interface** *interface*] command in global configuration mode.

FTOS Behavior: BPDU Guard and BPDU filtering (see Removing an Interface from the Spanning Tree Group on page 1054) both block BPDUs, but are two separate features.
BPDU Guard:

- is used on edgeports and blocks all traffic on edgeport if it receives a BPDU
- drops the BPDU after it reaches the RPM and generates a console message

BPDU Filtering:

- disables Spanning Tree on an interface
- drops all BPDUs at the line card without generating a console message

STP Root Selection

Although the Spanning Tree Protocol determines the root bridge, you can assign a lower priority to one bridge to increase the likelihood that it will be selected as the root bridge. You can also configure a bridge to be the root or secondary root.

To change the bridge priority or specify that a bridge is the root or secondary root:

Task	Command Syntax	Command Mode
Assign a number as the bridge priority or designate it as the root or secondary root. *priority-value* range: 0 to 65535. The lower the number assigned, the more likely this bridge will become the root bridge. The default is 32768. • The primary option specifies a bridge priority of 8192. • The secondary option specifies a bridge priority of 16384.	**bridge-priority** {*priority-value* | **primary** | **secondary**}	PROTOCOL SPANNING TREE

View only the root information using the **show spanning-tree root** command (see Figure 52-9) from EXEC privilege mode.

Figure 52-9. show spanning-tree root Command Example

```
FTOS#show spanning-tree 0 root
          Root ID  Priority 32768, Address 0001.e80d.2462
          We are the root of the spanning tree
          Root Bridge hello time 2, max age 20, forward delay 15
FTOS#
```

STP Root Guard

STP Root Guard is supported only on platforms: C E S

Use the STP Root Guard feature in a Layer 2 network to avoid bridging loops. In STP, the switch in the network with the lowest priority (as determined by STP or set with the **bridge-priority** command) is selected as the root bridge. If two switches have the same priority, the switch with the lower MAC address is selected as the root. All other switches in the network use the root bridge as the reference used to calculate the shortest forwarding path.

Because any switch in an STP network with a lower priority can become the root bridge, the forwarding topology may not be stable. The location of the root bridge can change, resulting in unpredictable network behavior. The STP root guard feature ensures that the position of the root bridge does not change.

Root Guard Scenario

For example, in Figure 52-10 (STP topology 1 upper left) Switch A is the root bridge in the network core. Switch C functions as an access switch connected to an external device. The link between Switch C and Switch B is in a blocking state. The flow of STP BPDUs is shown in the illustration.

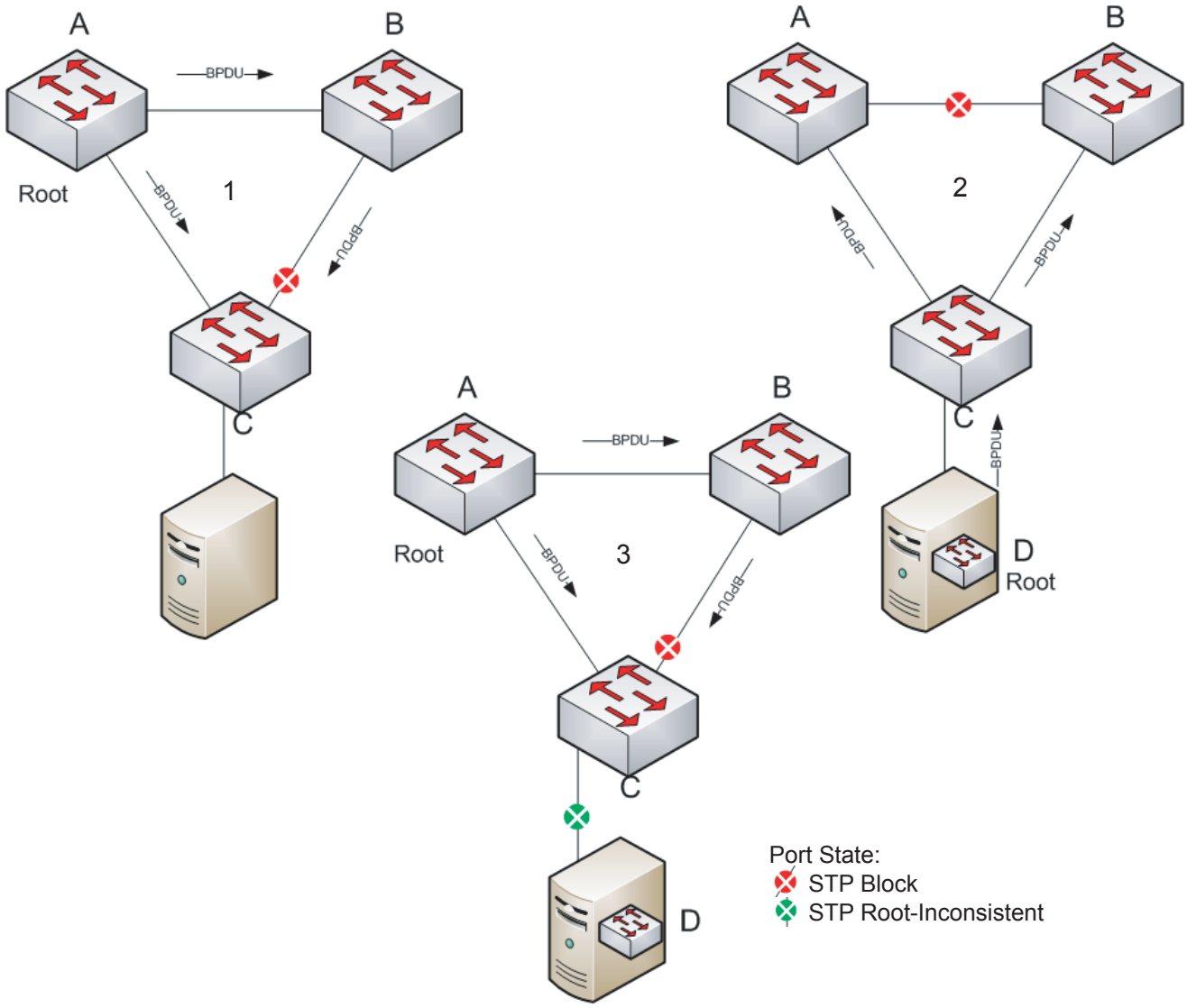
In STP topology 2 (Figure 52-10 upper right), STP is enabled on device D on which a software bridge application is started to connect to the network. Because the priority of the bridge in device D is lower than the root bridge in Switch A, device D is elected as root, causing the link between Switches A and B to enter a blocking state. Network traffic then begins to flow in the directions indicated by the BPDU arrows in the topology. If the links between Switches C and A or Switches C and B cannot handle the increased traffic flow, frames may be dropped.

In STP topology 3 (Figure 52-10 lower middle), if the root guard feature is enabled on the STP port on Switch C that connects to device D, and device D sends a superior BPDU that would trigger the election of device D as the new root bridge, the BPDU is ignored and the port on Switch C transitions from a forwarding to a root-inconsistent state (shown by the green X icon). As a result, Switch A becomes the root bridge.

All incoming and outgoing traffic is blocked on an STP port in a root-inconsistent state. After the timeout period, the Switch C port automatically transitions to a forwarding state as soon as device D stops sending BPDUs that advertise a lower priority.

If you enable a root guard on all STP ports on the links where the root bridge should not appear, you can ensure a stable STP network topology and avoid bridging loops.

Figure 52-10. STP Root Guard Prevents Bridging Loops



Root Guard Configuration

You enable STP root guard on a per-port or per-port-channel basis.

FTOS Behavior: The following conditions apply to a port enabled with STP root guard:

- Root guard is supported on any STP-enabled port or port-channel interface except when used as a stacking port.
- Root guard is supported on a port in any Spanning Tree mode:
  - Spanning Tree Protocol (STP)
  - Rapid Spanning Tree Protocol (RSTP)
  - Multiple Spanning Tree Protocol (MSTP)
  - Per-VLAN Spanning Tree Plus (PVST+)
- When enabled on a port, root guard applies to all VLANs configured on the port.
- Root guard and loop guard cannot be enabled at the same time on an STP port. For example, if you configure root guard on a port on which loop guard is already configured, the following error message is displayed:
  `% Error: LoopGuard is configured. Cannot configure RootGuard.`
- When used in an MSTP network, if root guard blocks a boundary port in the CIST, the port is also blocked in all other MST instances.

To enable the root guard on an STP-enabled port or port-channel interface in instance 0, enter the **spanning-tree 0 rootguard** command:

Task	Command Syntax	Command Mode
Enable root guard on a port or port-channel interface. **0**: Enables root guard on an STP-enabled port assigned to instance 0. **mstp**: Enables root guard on an MSTP-enabled port. **rstp**: Enables root guard on an RSTP-enabled port. **pvst**: Enables root guard on a PVST-enabled port.	**spanning-tree** {**0** | **mstp** | **rstp** | **pvst**} **rootguard**	INTERFACE INTERFACE PORT-CHANNEL

To disable STP root guard on a port or port-channel interface, enter the **no spanning-tree 0 rootguard** command in an interface configuration mode.

To verify the STP root guard configuration on a port or port-channel interface, enter the **show spanning-tree 0 guard [interface** *interface*] command in global configuration mode.

SNMP Traps for Root Elections and Topology Changes

- Enable SNMP traps for Spanning Tree state changes using the command **snmp-server enable traps stp**.
- Enable SNMP traps for MSTP using the command **snmp-server enable traps xstp**.

Configuring Spanning Trees as Hitless

Configuring Spanning Trees as Hitless is supported only on platforms: C E

You can configure Spanning Tree (STP), Rapid Spanning Tree (RSTP), Multiple Spanning Tree (MSTP), and Per-Vlan Spanning Tree (PVST+) to be hitless (all or none must be configured as hitless). When configured as hitless, critical protocol state information is synchronized between RPMs so that RPM failover is seamless, and no topology change is triggered.

Configure all Spanning Tree types to be hitless using the command **redundancy protocol xstp** from CONFIGURATION mode, as shown in Figure 52-11.

Figure 52-11. Configuring all Spanning Tree Types to be Hitless

```
FTOS(conf)#redundancy protocol xstp
FTOS#show running-config redundancy
!
redundancy protocol xstp
FTOS#
```

STP Loop Guard

STP Loop Guard is supported only on platforms: C E S

Loop Guard Scenario

The STP Loop Guard feature provides protection against Layer 2 forwarding loops (STP loops) caused by a hardware failure, such as a cable failure or an interface fault. When a cable or interface fails, a participating STP link may become unidirectional (STP requires links to be bidirectional) and an STP port does not receive BPDUs. When an STP blocking port does not receive BPDUs, it transitions to a forwarding state. This condition can create a loop in the network.

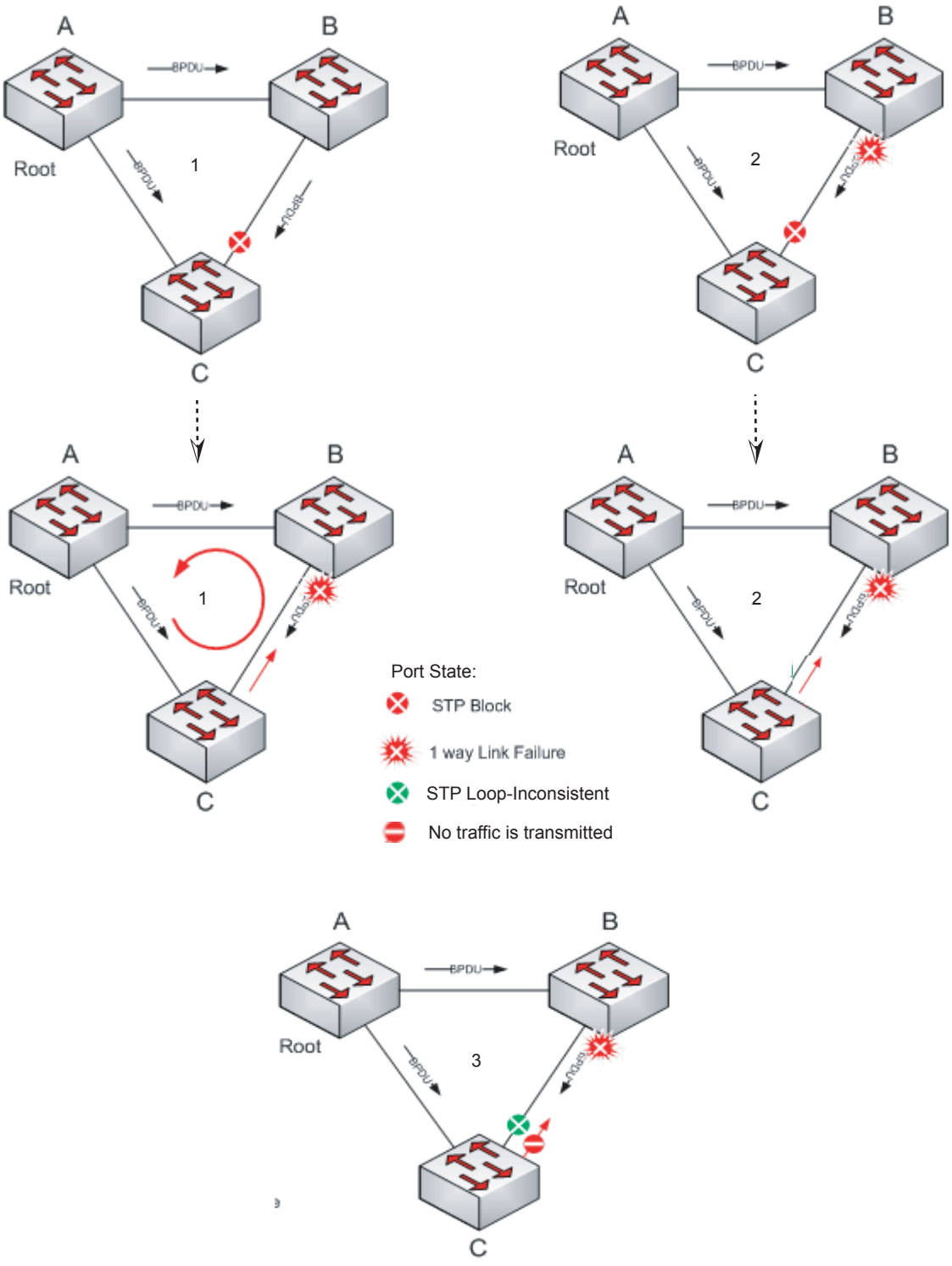
For example, in Figure 52-12 (STP topology 1 - upper left), Switch A is the root switch and Switch B normally transmits BPDUs to Switch C. The link between Switch C and Switch B is in a blocking state. However, if there is a unidirectional link failure (STP topology 1 - lower left), Switch C does not receive BPDUs from Switch B. When the **max-age** timer expires, the STP port on Switch C becomes unblocked and transitions to forwarding state. A loop is created as both Switch A and Switch C transmit traffic to Switch B.

Note that in Figure 52-12 (STP topology 2 - upper right), a loop can also be created if the forwarding port on Switch B becomes busy and does not forward BPDUs within the configured **forward-delay** time. As a result, the blocking port on Switch C transitions to a forwarding state, and both Switch A and Switch C transmit traffic to Switch B (STP topology 2 - lower right).

As shown in STP topology 3 (Figure 52-12 bottom middle), after you enable loop guard on an STP port or port-channel on Switch C, if no BPDUs are received and the **max-age** timer expires, the port transitions from a blocked state to a loop-inconsistent state (instead of to a forwarding state). Loop guard blocks the STP port so that no traffic is transmitted and no loop is created.

As soon as a BPDU is received on an STP port in a loop-inconsistent state, the port returns to a blocking state. If you disable STP loop guard on a port in a loop-inconsistent state, the port transitions to an STP blocking state and restarts the **max-age** timer.

Figure 52-12. STP Loop Guard Prevents Forwarding Loops



Loop Guard Configuration

You enable STP loop guard on a per-port or per-port channel basis.

FTOS Behavior: The following conditions apply to a port enabled with loop guard:

- Loop guard is supported on any STP-enabled port or port-channel interface.
- Loop guard is supported on a port or port-channel in any Spanning Tree mode:
  - Spanning Tree Protocol (STP)
  - Rapid Spanning Tree Protocol (RSTP)
  - Multiple Spanning Tree Protocol (MSTP)
  - Per-VLAN Spanning Tree Plus (PVST+)
- Root guard and loop guard cannot be enabled at the same time on an STP port. For example, if you configure loop guard on a port on which root guard is already configured, the following error message is displayed:
  `% Error: RootGuard is configured. Cannot configure LoopGuard.`
- **C-Series and E-Series only**: Loop guard is supported on a C-Series or E-Series switch configured for hitless STP (see Configuring Spanning Trees as Hitless on page 1064).
- Enabling Portfast BPDU guard and loop guard at the same time on a port results in a port that remains in a blocking state and prevents traffic from flowing through it. For example, when Portfast BPDU guard and loop guard are both configured:

  - If a BPDU is received from a remote device, BPDU guard places the port in an err-disabled blocking state and no traffic is forwarded on the port.

  - If no BPDU is received from a remote device, loop guard places the port in a loop-inconsistent blocking state and no traffic is forwarded on the port.
- When used in a PVST+ network, STP loop guard is performed per-port or per-port channel at a VLAN level. If no BPDUs are received on a VLAN interface, the port or port-channel transitions to a loop-inconsistent (blocking) state only for this VLAN.

To enable a loop guard on an STP-enabled port or port-channel interface, enter the **spanning-tree 0 loopguard** command:

Task	Command Syntax	Command Mode
Enable loop guard on a port or port-channel interface. **0**: Enables loop guard on an STP-enabled port assigned to instance 0. **mstp**: Enables loop guard on an MSTP-enabled port. **rstp**: Enables loop guard on an RSTP-enabled port. **pvst**: Enables loop guard on a PVST-enabled port.	**spanning-tree** {**0** | **mstp** | **rstp** | **pvst**} **loopguard**	INTERFACE INTERFACE PORT-CHANNEL

To disable STP loop guard on a port or port-channel interface, enter the **no spanning-tree 0 loopguard** command in an INTERFACE configuration mode.

To verify the STP loop guard configuration on a port or port-channel interface, enter the **show spanning-tree 0 guard [interface** *interface*] command in global configuration mode.

Displaying STP Guard Configuration

To verify the STP guard configured on port or port-channel interfaces, enter the **show spanning-tree 0 guard** [**interface** *interface*] command.

Figure 52-13 shows an example for an STP network (instance 0) in which:

- Root guard is enabled on a port that is in a root-inconsistent state.
- Loop guard is enabled on a port that is in a listening state.
- BPDU guard is enabled on a port that is shut down (Error Disabled state) after receiving a BPDU.

Figure 52-13. Displaying STP Guard Configuration

```
FTOS#show spanning-tree 0 guard
Interface
Name      Instance    Sts          Guard type
---------  --------   ---------     ----------
Gi 0/1     0          INCON(Root)  Rootguard
Gi 0/2     0          LIS          Loopguard
Gi 0/3     0          EDS (Shut)   Bpduguard
```

53

System Time and Date

Chapter 53, System Time and Date settings, and Network Time Protocol are supported on platforms: C E S

Time and Date and NTP are supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later.

System times and dates can be set and maintained through the Network Time Protocol (NTP). They are also set through FTOS CLIs and hardware settings.

This chapter includes the following sections:

Network Time Protocol

Network Time Protocol (NTP) synchronizes timekeeping among a set of distributed time servers and clients. The protocol also coordinates time distribution in a large, diverse network with a variety of interfaces. In NTP, servers maintain the time and NTP clients synchronize with a time-serving host. NTP clients choose from among several NTP servers to determine which offers the best available source of time and the most reliable transmission of information.

NTP is a fault-tolerant protocol that will automatically select the best of several available time sources to synchronize to. Multiple candidates can be combined to minimize the accumulated error. Temporarily or permanently insane time sources will be detected and avoided.

Dell Force10 recommends configuring NTP for the most accurate time. In FTOS, other time sources can be configured (the hardware clock and the software clock).

NTP is designed to produce three products: clock offset, roundtrip delay, and dispersion, all of which are relative to a selected reference clock.

- **Clock offset** represents the amount to adjust the local clock to bring it into correspondence with the reference clock.
- **Roundtrip delay** provides the capability to launch a message to arrive at the reference clock at a specified time.
- **Dispersion** represents the maximum error of the local clock relative to the reference clock.

Since most host time servers will synchronize via another peer time server, there are two components in each of these three products, those determined by the peer relative to the primary reference source of standard time and those measured by the host relative to the peer.

Each of these components are maintained separately in the protocol in order to facilitate error control and management of the subnet itself. They provide not only precision measurements of offset and delay, but also definitive maximum error bounds, so that the user interface can determine not only the time, but the quality of the time as well.

In what may be the most common client/server model a client sends an NTP message to one or more servers and processes the replies as received. The server interchanges addresses and ports, overwrites certain fields in the message, recalculates the checksum and returns the message immediately. Information included in the NTP message allows the client to determine the server time with respect to local time and adjust the local clock accordingly. In addition, the message includes information to calculate the expected timekeeping accuracy and reliability, as well as select the best from possibly several servers.

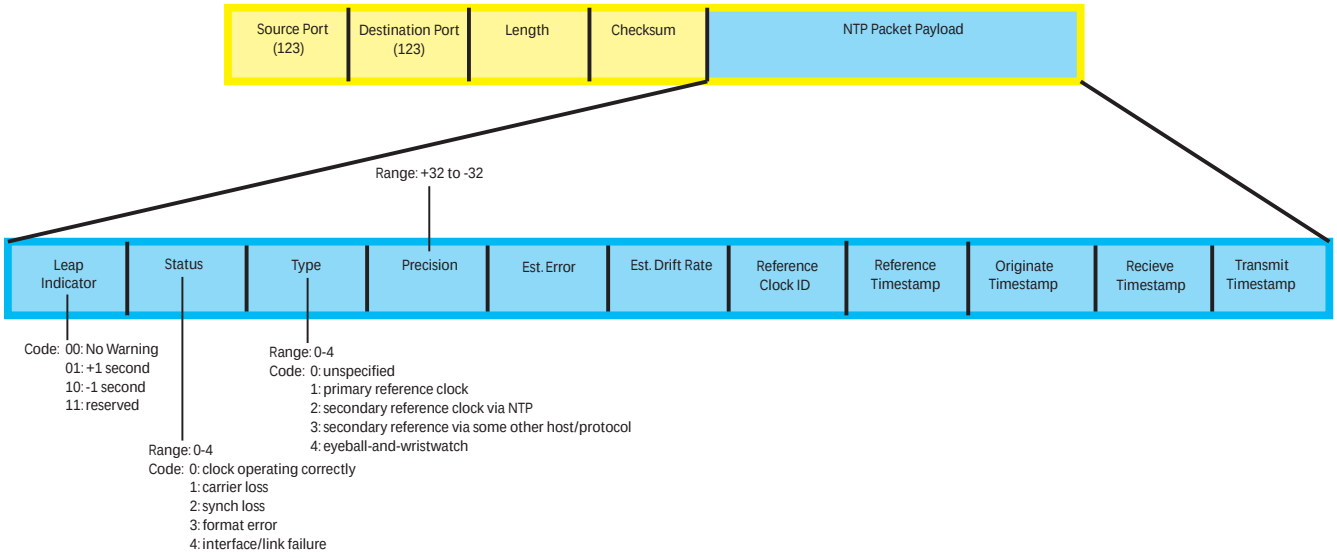
Following conventions established by the telephone industry [BEL86], the accuracy of each server is defined by a number called the stratum, with the topmost level (primary servers) assigned as one and each level downwards (secondary servers) in the hierarchy assigned as one greater than the preceding level.

FTOS synchronizes with a time-serving host to get the correct time. You can set FTOS to poll specific NTP time-serving hosts for the current time. From those time-serving hosts, the system chooses one NTP host with which to synchronize and serve as a client to the NTP host. As soon as a host-client relationship is established, the networking device propagates the time information throughout its local network.

Protocol Overview

NTP message to one or more servers and processes the replies as received. The server interchanges addresses and ports, fills in or overwrites certain fields in the message, recalculates the checksum and returns it immediately. Information included in the NTP message allows each client/server peer to determine the timekeeping characteristics of its other peers, including the expected accuracies of their clocks. Using this information each peer is able to select the best time from possibly several other clocks, update the local clock and estimate its accuracy.

Figure 53-1. NTP Fields



Implementation Information

- Dell Force10 systems can only be an NTP client.

Configuring Network Time Protocol

Configuring NTP is a one-step process:

1. Enable NTP. See page 1072.

Related Configuration Tasks

- Configure NTP broadcasts on page 1073
- Set the Hardware Clock with the Time Derived from NTP on page 1073
- Set the Hardware Clock with the Time Derived from NTP on page 1073
- Disable NTP on an interface on page 1073
- Configure a source IP address for NTP packets on page 1074 (optional)

Enable NTP

NTP is disabled by default. To enable it, specify an NTP server to which the Dell Force10 system will synchronize. Enter the command multiple times to specify multiple servers. You may specify an unlimited number of servers at the expense of CPU resources.

Task	Command	Command Mode
Specify the NTP server to which the Dell Force10 system will synchronize. You may specify an IPv4 or IPv6 address, or hostname that resolves to an IPv4 or IPv6 address.	**ntp server** {*hostname* | *ipv4-address* | *ipv6-address*} [**key** *keyid*] [**prefer**] [**version** *number*]	CONFIGURATION

Display the system clock state with respect to NTP using the command **show ntp status** from EXEC Privilege mode, as shown in Figure 53-2.

Figure 53-2. Displaying the System Clock State with respect to NTP

```
FTOS(conf)#do show ntp status
Clock is synchronized, stratum 2, reference is 192.168.1.1
frequency is -369.623 ppm, stability is 53.319 ppm, precision is 4294967279
reference time is CD63BCC2.0CBBD000 (16:54:26.049 UTC Thu Mar 12 2009)
clock offset is 997.529984 msec, root delay is 0.00098 sec
root dispersion is 10.04271 sec, peer dispersion is 10032.715 msec
peer mode is client
```

Display the calculated NTP synchronization variables received from the server that the system will use to synchronize its clock using the command **show ntp associations** from EXEC Privilege mode, as shown in Figure 53-3.

Figure 53-3. Displaying the Calculated NTP Synchronization Variables

```
R6_E300(conf)#do show ntp associations
   remote        ref clock     st when poll reach   delay   offset    disp
==============================================================================
#192.168.1.1     .LOCL.           1   16   16   76     0.98   -2.470  879.23
* master (synced), # master (unsynced), + selected, - candidate
```

Set the Hardware Clock with the Time Derived from NTP

Task	Command	Command Mode
Periodically update the system hardware clock with the time value derived from NTP.	**ntp update-calendar**	CONFIGURATION

Figure 53-4. Displaying the Calculated NTP Synchronization Variables

```
R5/R8(conf)#do show calendar
06:31:02 UTC Mon Mar 13 1989
R5/R8(conf)#ntp update-calendar 1
R5/R8(conf)#do show calendar
06:31:26 UTC Mon Mar 13 1989
R5/R8(conf)#do show calendar
12:24:11 UTC Thu Mar 12 2009
```

Configure NTP broadcasts

With FTOS, you can receive broadcasts of time information. You can set interfaces within the system to receive NTP information through broadcast.

To configure an interface to receive NTP broadcasts, use the following commands in the INTERFACE mode:

Task	Command	Command
Set the interface to receive NTP packets.	**ntp broadcast client**	INTERFACE

Table 53-1.

```
2w1d11h : NTP: Maximum Slew:-0.000470, Remainder = -0.496884
```

Disable NTP on an interface

By default, NTP is enabled on all active interfaces. If you disable NTP on an interface, FTOS drops any NTP packets sent to that interface.

To disable NTP on an interface, use the following command in the INTERFACE mode:

Command Syntax	Command Mode	Purpose
ntp disable	INTERFACE	Disable NTP on the interface.

To view whether NTP is configured on the interface, use the **show config** command in the INTERFACE mode. If **ntp disable** is not listed in the **show config** command output, then NTP is enabled. (The **show config** command displays only non-default configuration information.)

Configure a source IP address for NTP packets

By default, the source address of NTP packets is the IP address of the interface used to reach the network. You can configure one interface's IP address to be included in all NTP packets.

To configure an IP address as the source address of NTP packets, use the following command in the CONFIGURATION mode:

Command Syntax	Command Mode	Purpose
ntp source *interface*	CONFIGURATION	Enter the following keywords and slot/port or number information: • For a 1-Gigabit Ethernet interface, enter the keyword **GigabitEthernet** followed by the slot/port information. • For a loopback interface, enter the keyword **loopback** followed by a number between 0 and 16383. • For a port channel interface, enter the keyword **lag** followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale. • For a SONET interface, enter the keyword **sonet** followed by the slot/port information. • For a 10-Gigabit Ethernet interface, enter the keyword **TenGigabitEthernet** followed by the slot/port information. • For a VLAN interface, enter the keyword **vlan** followed by a number from 1 to 4094. E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

To view the configuration, use the **show running-config ntp** command (Figure 38) in the EXEC privilege mode.

Configure NTP authentication

NTP authentication and the corresponding trusted key provide a reliable means of exchanging NTP packets with trusted time sources. NTP authentication begins when the first NTP packet is created following the configuration of keys. NTP authentication in FTOS uses the MD5 algorithm and the key is embedded in the synchronization packet that is sent to an NTP time source.

FTOS Behavior: FTOS versions 8.2.1.0 and later use an encryption algorithm to store the authentication key that is different from previous FTOS versions; beginning in version 8.2.1.0, FTOS uses DES encryption to store the key in the startup-config when you enter the command **ntp authentication-key**. Therefore, if your system boots with a startup-configuration from an FTOS versions prior to 8.2.1.0 in which you have configured **ntp authentication-key**, the system cannot correctly decrypt the key, and cannot authenticate NTP packets. In this case you must re-enter this command and save the running-config to the startup-config.

To configure NTP authentication, use these commands in the following sequence in the CONFIGURATION mode:

Step	Command Syntax	Command Mode	Purpose
1	**ntp authenticate**	CONFIGURATION	Enable NTP authentication.
2	**ntp authentication-key** *number* **md5** *key*	CONFIGURATION	Set an authentication key. Configure the following parameters: *number:* Range 1 to 4294967295. This *number* must be the same as the *number* in the **ntp trusted-key** command. *key:* Enter a text string. This text string is encrypted.
3	**ntp trusted-key** *number*	CONFIGURATION	Define a trusted key. Configure a number from 1 to 4294967295. The *number* must be the same as the *number* used in the **ntp authentication-key** command.

To view the NTP configuration, use the **show running-config ntp** command (Figure 40) in the EXEC privilege mode. Figure 53-5 shows an encrypted authentication key. All keys are encrypted.

Figure 53-5. show running-config ntp Command Example

Command Syntax	Command Mode	Purpose
ntp server *ip-address* [**key** *keyid*] [**prefer**] [**version** *number*]	CONFIGURATION	Configure an NTP server. Configure the IP address of a server and the following optional parameters: • **key** *keyid:* Configure a text string as the key exchanged between the NTP server and client. • **prefer:** Enter the keyword to set this NTP server as the preferred server. • **version** *number:* Enter a number 1 to 3 as the NTP version.

```
R6_E300(conf)#1w6d23h : NTP: xmit packet to 192.168.1.1:
 leap 0, mode 3, version 3, stratum 2, ppoll 1024
 rtdel 0219 (8.193970), rtdsp AF928 (10973.266602), refid C0A80101 (192.168.1.1)
 ref CD7F4F63.6BE8F000 (14:51:15.421 UTC Thu Apr 2 2009)
 org CD7F4F63.68000000 (14:51:15.406 UTC Thu Apr 2 2009)
 rec CD7F4F63.6BE8F000 (14:51:15.421 UTC Thu Apr 2 2009)
 xmt CD7F5368.D0535000 (15:8:24.813 UTC Thu Apr 2 2009)
1w6d23h : NTP: rcv packet from 192.168.1.1
 leap 0, mode 4, version 3, stratum 1, ppoll 1024
 rtdel 0000 (0.000000), rtdsp AF587 (10959.090820), refid 4C4F434C (76.79.67.76)
 ref CD7E14FD.43F7CED9 (16:29:49.265 UTC Wed Apr 1 2009)
 org CD7F5368.D0535000 (15:8:24.813 UTC Thu Apr 2 2009)
 rec CD7F5368.D0000000 (15:8:24.812 UTC Thu Apr 2 2009)
 xmt CD7F5368.D0000000 (15:8:24.812 UTC Thu Apr 2 2009)
 inp CD7F5368.D1974000 (15:8:24.818 UTC Thu Apr 2 2009)

rtdel-root delay
rtdsp - round trip dispersion
refid - reference id
org -
rec - (last?) receive timestamp
xmt - transmit timestamp

mode - 3 client, 4 server
stratum - 1 primary reference clock, 2 secondary reference clock (via NTP)
version - NTP version 3
leap -
```

- Leap Indicator (sys.leap, peer.leap, pkt.leap): This is a two-bit code warning of an impending leap second to be inserted in the NTP time scale. The bits are set before 23:59 on the day of insertion and reset after 00:00 on the following day. This causes the number of seconds (rollover interval) in the day of insertion to be increased or decreased by one. In the case of primary servers the bits are set by operator intervention, while in the case of secondary servers the bits are set by the protocol. The two bits, bit 0 and bit 1, respectively, are coded as follows:
- Poll Interval: integer indicating the minimum interval between transmitted messages, in seconds as a power of two. For instance, a value of six indicates a minimum interval of 64 seconds.
- Precision: integer indicating the precision of the various clocks, in seconds to the nearest power of two. The value must be rounded to the next larger power of two; for instance, a 50-Hz (20 ms) or 60-Hz (16.67ms) power-frequency clock would be assigned the value -5 (31.25 ms), while a 1000-Hz (1 ms) crystal-controlled clock would be assigned the value -9 (1.95 ms).

- Root Delay (sys.rootdelay, peer.rootdelay, pkt.rootdelay): This is a signed fixed-point number indicating the total roundtrip delay to the primary reference source at the root of the synchronization subnet, in seconds. Note that this variable can take on both positive and negative values, depending on clock precision and skew.
- Root Dispersion (sys.rootdispersion, peer.rootdispersion, pkt.rootdispersion): This is a signed fixed-point number indicating the maximum error relative to the primary reference source at the root of the synchronization subnet, in seconds. Only positive values greater than zero are possible.
- Reference Clock Identifier (sys.refid, peer.refid, pkt.refid): This is a 32-bit code identifying the particular reference clock. In the case of stratum 0 (unspecified) or stratum 1 (primary reference source), this is a four-octet, left-justified, zero-padded ASCII string, for example (see Appendix A for comprehensive list): the case of stratum 2 and greater (secondary reference) this is the four-octet Internet address of the peer selected for synchronization.
- Reference Timestamp (sys.reftime, peer.reftime, pkt.reftime): This is the local time, in timestamp format, when the local clock was last updated. If the local clock has never been synchronized, the value is zero.
- **Originate Timestamp**: The departure time on the server of its last NTP message. If the server becomes unreachable, the value is set to zero.
- **Receive Timestamp**: The arrival time on the client of the last NTP message from the server. If the server becomes unreachable, the value is set to zero.
- **Transmit Timestamp**: The departure time on the server of the current NTP message from the sender.
- Filter dispersion is the error in calculating the minimum delay from a set of sample data from a peer.

FTOS Time and Date

The time and date can be set using the FTOS CLI.

Configuring time and date settings

The following list includes the configuration tasks for setting the system time:

- Set the time and date for the switch hardware clock
- Set the time and date for the switch software clock
- Set the timezone
- Set daylight savings time
  - Set Daylight Saving Time Once
  - Set Recurring Daylight Saving Time

Set the time and date for the switch hardware clock

Command Syntax	Command Mode	Purpose
calendar set *time month day year*	EXEC Privilege	Set the hardware clock to the current time and date. *time:* Enter the time in hours:minutes:seconds. For the hour variable, use the 24-hour format, for example, 17:15:00 is 5:15 pm. *month:* Enter the name of one of the 12 months in English. You can enter the name of a day to change the order of the display to *time day month year*. *day:* Enter the number of the day. Range: 1 to 31. You can enter the name of a month to change the order of the display to *time day month year* *year:* Enter a four-digit number as the year. Range: 1993 to 2035.

```
FTOS#calendar set 08:55:00 september 18 2009
FTOS#
```

Set the time and date for the switch software clock

You can change the order of the *month* and *day* parameters to enter the time and date as *time day month year*. You cannot delete the software clock.

The software clock runs only when the software is up. The clock restarts, based on the hardware clock, when the switch reboots.

Command Syntax	Command Mode	Purpose
clock set *time month day year*	EXEC Privilege	Set the system software clock to the current time and date. *time:* Enter the time in hours:minutes:seconds. For the hour variable, use the 24-hour format, for example, 17:15:00 is 5:15 pm. *month:* Enter the name of one of the 12 months in English. You can enter the name of a day to change the order of the display to *time day month year*. *day:* Enter the number of the day. Range: 1 to 31. You can enter the name of a month to change the order of the display to *time day month year* *year:* Enter a four-digit number as the year. Range: 1993 to 2035.

```
FTOS#clock set 16:20:00 19 september 2009
FTOS#
```

Set the timezone

Coordinated Universal Time (UTC) is the time standard based on the International Atomic Time standard, commonly known as Greenwich Mean time. When determining system time, you must include the differentiator between UTC and your local timezone. For example, San Jose, CA is the Pacific Timezone with a UTC offset of -8.

Command Syntax	Command Mode	Purpose
clock timezone *timezone-name offset*	CONFIGURATION	Set the clock to the appropriate timezone. *timezone-name:* Enter the name of the timezone. Do not use spaces. *offset:* Enter one of the following: • a number from 1 to 23 as the number of hours in addition to UTC for the timezone. • a minus sign (-) followed by a number from 1 to 23 as the number of hours

```
FTOS#conf
FTOS(conf)#clock timezone Pacific -8
FTOS(conf)#01:40:19: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Timezone
configuration changed from "UTC 0 hrs 0 mins" to "Pacific -8 hrs
0 mins"
```

Set daylight savings time

FTOS supports setting the system to daylight savings time once or on a recurring basis every year.

Set Daylight Saving Time Once

Set a date (and time zone) on which to convert the switch to daylight savings time on a one-time basis.

Command Syntax	Command Mode	Purpose
clock summer-time *time-zone* **date** *start-month start-day start-year start-time end-month end-day end-year end-time* [*offset*]	CONFIGURATION	Set the clock to the appropriate time zone and daylight savings time. *time-zone:* Enter the three-letter name for the time zone. This name is displayed in the show clock output. *start-month:* Enter the name of one of the 12 months in English. You can enter the name of a day to change the order of the display to *time day month year* *start-day:* Enter the number of the day. Range: 1 to 31. You can enter the name of a month to change the order of the display to *time day month year*. *start-year:* Enter a four-digit number as the year. Range: 1993 to 2035 *start-time:* Enter the time in hours:minutes. For the hour variable, use the 24-hour format, example, 17:15 is 5:15 pm. *end-month:* Enter the name of one of the 12 months in English. You can enter the name of a day to change the order of the display to *time day month year*. *end-day:* Enter the number of the day. Range: 1 to 31. You can enter the name of a month to change the order of the display to *time day month year*. *end-year:* Enter a four-digit number as the year. Range: 1993 to 2035. *end-time:* Enter the time in hours:minutes. For the hour variable, use the 24-hour format, example, 17:15 is 5:15 pm. *offset:* (OPTIONAL) Enter the number of minutes to add during the summer-time period. Range: 1 to1440. Default: 60 minutes

Command Syntax	Command Mode	Purpose

```
FTOS(conf)#clock summer-time pacific date Mar 14 2009 00:00 Nov 7 2009 00:00
FTOS(conf)#02:02:13: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Summertime configuration changed from
"none" to "Summer time starts 00:00:00 Pacific Sat Mar 14 2009;Summer time ends 00:00:00 pacific
Sat Nov 7 2009"
```

Set Recurring Daylight Saving Time

Set a date (and time zone) on which to convert the switch to daylight savings time on a specific day every year.

If you have already set daylight savings for a one-time setting, you can set that date and time as the recurring setting with the **clock summer-time** *time-zone* **recurring** command.

Command Syntax	Command Mode	Purpose
clock summer-time *time-zone* **recurring** *start-week start-day start-month start-time end-week end-day end-month end-time* [*offset*]	CONFIGURATION	Set the clock to the appropriate timezone and adjust to daylight savings time every year. *time-zone:* Enter the three-letter name for the time zone. This name is displayed in the show clock output. *start-week:* (OPTIONAL) Enter one of the following as the week that daylight savings begins and then enter values for *start-day* through *end-time*: • *week-number:* Enter a number from 1-4 as the number of the week in the month to start daylight savings time. • **first:** Enter this keyword to start daylight savings time in the first week of the month. • **last:** Enter this keyword to start daylight savings time in the last week of the month. *start-month:* Enter the name of one of the 12 months in English. You can enter the name of a day to change the order of the display to *time day month year* *start-day:* Enter the number of the day. Range: 1 to 31. You can enter the name of a month to change the order of the display to *time day month year*.

Command Syntax	Command Mode	Purpose
		start-year: Enter a four-digit number as the year. Range: 1993 to 2035 *start-time:* Enter the time in hours:minutes. For the hour variable, use the 24-hour format, example, 17:15 is 5:15 pm. *end-week: If you entered a start-week,* Enter the one of the following as the week that daylight savings ends: • *week-number:* enter a number from 1-4 as the number of the week to end daylight savings time. • **first:** enter the keyword first to end daylight savings time in the first week of the month. • **last:** enter the keyword last to end daylight savings time in the last week of the month. *end-month:* Enter the name of one of the 12 months in English. You can enter the name of a day to change the order of the display to *time day month year*. *end-day:* Enter the number of the day. Range: 1 to 31. You can enter the name of a month to change the order of the display to *time day month year*. *end-year:* Enter a four-digit number as the year. Range: 1993 to 2035. *end-time:* Enter the time in hours:minutes. For the hour variable, use the 24-hour format, example, 17:15 is 5:15 pm. *offset:* (OPTIONAL) Enter the number of minutes to add during the summer-time period. Range: 1 to1440. Default: 60 minutes

```
FTOS(conf)#clock summer-time pacific recurring Mar 14 2009 00:00 Nov 7 2009 00:00 ?
FTOS(conf)#02:02:13: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Summertime configuration changed from
"none" to "Summer time starts 00:00:00 Pacific Sat Mar 14 2009;Summer time ends 00:00:00 pacific
Sat Nov 7 2009"
```

Note: If you enter <CR> after entering the recurring command parameter, and you have already set a one-time daylight saving time/date, the system will use that time and date as the recurring setting.

Command Syntax	Command Mode	Purpose

```
FTOS(conf)#clock summer-time pacific recurring ?
<1-4>                   Week number to start
first                   Week number to start
last                    Week number to start
<cr>
FTOS(conf)#clock summer-time pacific recurring
FTOS(conf)#02:10:57: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Summertime configuration changed from
"Summer time starts 00:00:00 Pacific Sat Mar 14 2009 ; Summer time ends 00:00:00 pacific Sat Nov
7 2009" to "Summer time starts 02:00:00 Pacific Sun Mar 8 2009;Summer time ends 02:00:00 pacific
```

54

Uplink Failure Detection (UFD)

Uplink Failure Detection (UFD) is supported on platform: S (S50 only)

Feature Description

Uplink Failure Detection (UFD) provides detection of the loss of upstream connectivity and, if used with NIC teaming, automatic recovery from a failed link.

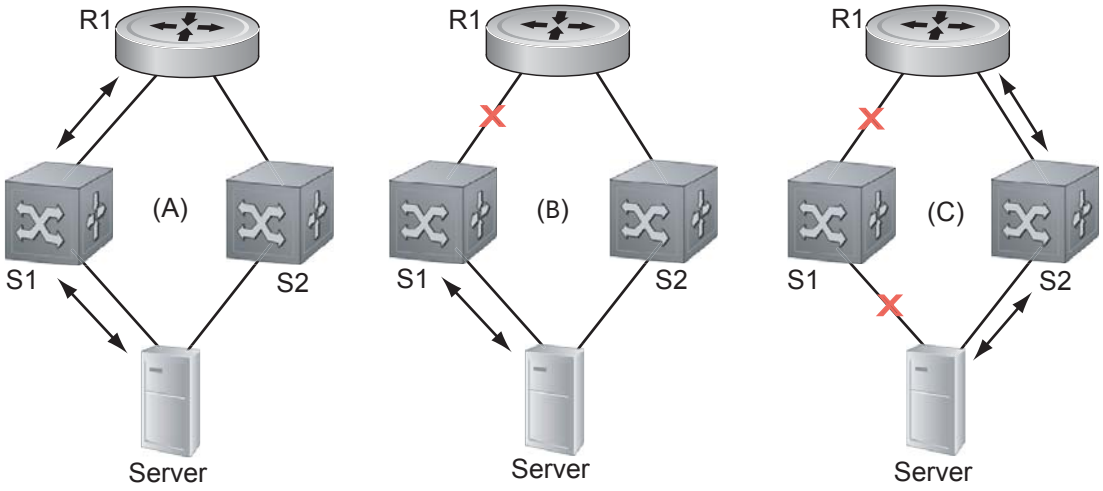
A switch provides upstream connectivity for devices, such as servers. If a switch loses its upstream connectivity, downstream devices also lose their connectivity. However, the devices do not receive a direct indication that upstream connectivity is lost since connectivity to the switch is still operational.

UFD allows a switch to associate downstream interfaces with upstream interfaces. When upstream connectivity fails, the switch disables the downstream links. Failures on the downstream links allow downstream devices to recognize the loss of upstream connectivity.

For example, in Figure 54-1 Switches S1 and S2 both have upstream connectivity to Router R1 and downstream connectivity to the server. UFD operation is shown in Steps A through C:

- In Step A, the server configuration uses the connection to S1 as the primary path. Network traffic flows from the server to S1 and then upstream to R1.
- In Step B, the upstream link between S1 and R1 fails. The server continues to use the link to S1 for its network traffic, but the traffic is not successfully switched through S1 because the upstream link is down.
- In Step C, UFD on S1 disables the link to the server. The server then stops using the link to S1 and switches to using its link to S2 to send traffic upstream to R1.

Figure 54-1. Uplink Failure Detection

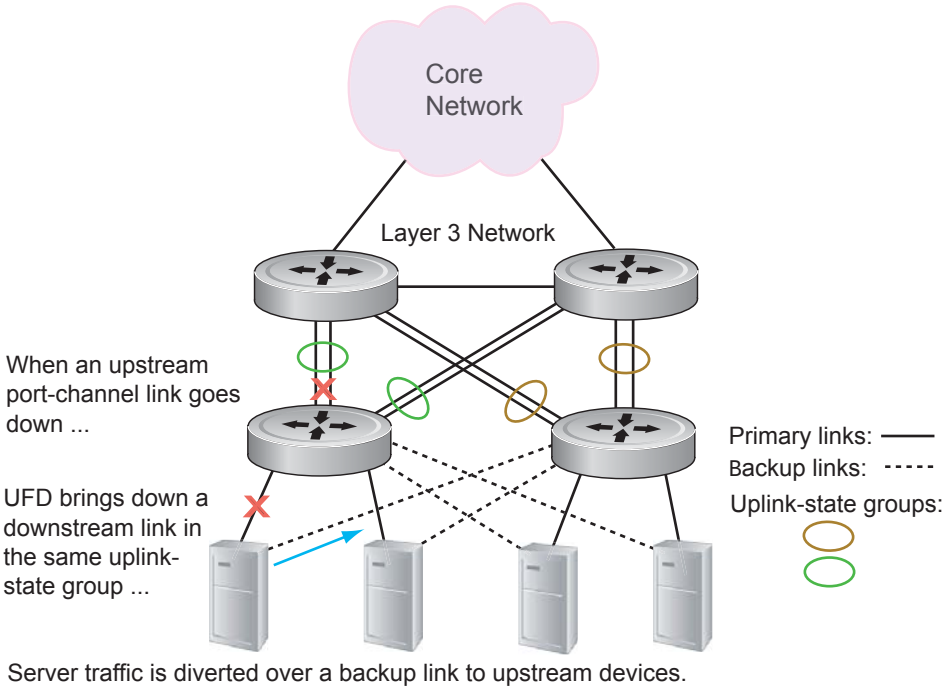


How Uplink Failure Detection Works

UFD creates an association between upstream and downstream interfaces. The association of uplink and downlink interfaces is called an *uplink-state group*. An interface in an uplink-state group can be a physical interface or a port-channel (LAG) aggregation of physical interfaces.

An enabled uplink-state group tracks the state of all assigned upstream interfaces. Failure on an upstream interface results in the automatic disabling of downstream interfaces in the uplink-state group. As a result, downstream devices can execute the protection or recovery procedures they have in place to establish alternate connectivity paths as shown in Figure 54-2.

Figure 54-2. Uplink Failure Detection Example



If only one of the upstream interfaces in an uplink-state group goes down, a specified number of downstream ports associated with the upstream interface are put into a link-down state. This number is user-configurable and is calculated by the ratio of upstream port bandwidth to downstream port bandwidth in the same uplink-state group. This calculation ensures that there are no traffic drops due to insufficient bandwidth on the upstream links to the routers/switches.

By default, if all upstream interfaces in an uplink-state group go down, all downstream interfaces in the same uplink-state group are put into a link-down state.

Using UFD, you can configure the automatic recovery of downstream ports in an uplink-state group when the link status of an upstream port changes. The tracking of upstream link status does not have a major impact on CPU usage.

UFD and NIC Teaming

Uplink Failure Detection on a switch can be used with network adapter teaming on a server (see NIC Teaming on page 569) to implement a rapid failover solution. For example, in Figure 54-2 the switch/router with UFD detects the uplink failure and automatically disables the associated downstream link port to the server. The server with NIC teaming detects the disabled link and automatically switches over to the backup link in order to continue to transmit traffic upstream.

Important Points to Remember

When you configure Uplink Failure Detection, the following conditions apply:

- You can configure up to sixteen uplink-state groups. By default, no uplink-state groups are created.

  An uplink-state group is considered to be operationally *up* if it has at least one upstream interface in the link-up state.

  An uplink-state group is considered to be operationally *down* if it has no upstream interfaces in the link-up state. No uplink-state tracking is performed when a group is disabled or in an operationally down state.

- You can assign physical port or port-channel interfaces to an uplink-state group.

  You can assign an interface to only one uplink-state group. Each interface assigned to an uplink-state group must be configured as either an upstream or downstream interface, but not both.

  You can assign individual member ports of a port channel to the group. An uplink-state group can contain either the member ports of a port channel or the port channel itself, but not both.

  If you assign a port channel as an upstream interface, the port channel interface enters a link-down state when the number of port-channel member interfaces in a link-up state drops below the configured Minimum Number of Members parameter.

- If one of the upstream interfaces in an uplink-state group goes down, either a user-configurable set of downstream ports or all the downstream ports in the group are put in an operationally down state with an UFD Disabled error. The order in which downstream ports are disabled is from the lowest numbered port to the highest.

  If one of the upstream interfaces in an uplink-state group that was down comes up, the set of UFD-disabled downstream ports (which were previously disabled due to this upstream port going down) are brought up and the UFD Disabled error is cleared.

- If an uplink-state group is disabled, the downstream interfaces are not disabled regardless of the state of the upstream interfaces.

  If an uplink-state group has no upstream interfaces assigned, downstream interfaces cannot be disabled when an upstream link goes down.

- To enable the debug messages for events related to a specified uplink-state group or all groups, enter the **debug uplink-state-group** [*group-id*] command, where *group-id* is 1 to 16.

  To turn off debugging event messages, enter the **no debug uplink-state-group** [*group-id*] command.

  For an example of debug log messages, see Message 1.

Configuring Uplink Failure Detection

To configure Uplink Failure Detection, follow these steps:

Step	Command Syntax and Mode	Description
1	**uplink-state-group** *group-id* Command Mode: CONFIGURATION	Creates an uplink-state group and enabling the tracking of upstream links on the switch/router. Valid *group-id* values are 1 to 16. To delete an uplink-state group, enter the **no uplink-state-group** *group-id* command.
2	**{upstream | downstream}** *interface* Command Mode: UPLINK-STATE-GROUP	Assigns a port or port-channel to the uplink-state group as an upstream or downstream interface. For *interface*, enter one of the following interface types: Fast Ethernet: **fastethernet** {*slot*/*port* | *slot*/*port-range*} 1-Gigabit Ethernet: **gigabitethernet** {*slot*/*port* |*slot*/*port-range*} 10-Gigabit Ethernet: **tengigabitethernet** {*slot*/*port* |*slot*/*port-range*} Port channel: **port-channel** {1-512 | *port-channel-range*} Where *port-range* and *port-channel-range* specify a range of ports separated by a dash (-) and/or individual ports/port channels in any order; for example: `upstream gigabitethernet 1/1-2,5,9,11-12` `downstream port-channel 1-3,5` A comma is required to separate each port and port-range entry. To delete an interface from the group, enter the **no {upstream | downstream}** *interface* command.
3	**downstream disable links {***number* | **all}** Command Mode: UPLINK-STATE-GROUP	Configures the number of downstream links in the uplink-state group that will be disabled (Oper Down state) if one upstream link in the group goes down. *number* specifies the number of downstream links to be brought down. Range: 1 to 1024. **all** brings down all downstream links in the group. Default: No downstream links are disabled when an upstream link goes down. **Note**: Downstream interfaces in an uplink-state group are put into a link-down state with an UFD-Disabled error message only when all upstream interfaces in the group go down. To revert to the default setting, enter the **no downstream disable links** command.
4	**downstream auto-recover** Command Mode: UPLINK-STATE-GROUP	(Optional) Enables auto-recovery so that UFD-disabled downstream ports in the uplink-state group come up when a disabled upstream port in the group comes back up. Default: Auto-recovery of UFD-disabled downstream ports is enabled. To disable auto-recovery, enter the **no downstream auto-recover** command.

Step	Command Syntax and Mode	Description
5	**description** *text* Command Mode: UPLINK-STATE-GROUP	(Optional) Enters a text description of the uplink-state group. Maximum length: 80 alphanumeric characters.
6	**no enable** Command Mode: UPLINK-STATE-GROUP	(Optional) Disables upstream-link tracking without deleting the uplink-state group. Default: Upstream-link tracking is automatically enabled in an uplink-state group. To re-enable upstream-link tracking, enter the **enable** command.

Clearing a UFD-Disabled Interface

Debug log message

You can manually bring up a downstream interface in an uplink-state group that has been disabled by UFD and is in a UFD-disabled error state. To re-enable one or more disabled downstream interfaces and clear the UFD-disabled error state, enter the following command:

Command Syntax	Description
clear ufd-disable {interface *interface* | **uplink-state-group** *group-id* **}** Command Mode: CONFIGURATION	Re-enables a downstream interface on the switch/router that is in a UFD-disabled error state so that it can send and receive traffic. For *interface*, enter one of the following interface types: Fast Ethernet: **fastethernet** {*slot*/*port* | *slot*/*port-range*} 1-Gigabit Ethernet: **gigabitethernet** {*slot*/*port* |*slot*/*port-range*} 10-Gigabit Ethernet: **tengigabitethernet** {*slot*/*port* |*slot*/*port-range*} Port channel: **port-channel** {1-512 | *port-channel-range*} Where *port-range* and *port-channel-range* specify a range of ports separated by a dash (-) and/or individual ports/port channels in any order; for example: `gigabitethernet 1/1-2,5,9,11-12` `port-channel 1-3,5` A comma is required to separate each port and port-range entry. **uplink-state-group** *group-id* re-enables all UFD-disabled downstream interfaces in the group. Valid values are 1 to 16.

Debug log message

Message 1 shows the Syslog messages displayed when you clear the UFD-disabled state from all disabled downstream interfaces in an uplink-state group by entering the **clear ufd-disable uplink-state-group** *group-id* command. All downstream interfaces return to an operationally up state.

Message 1 Syslog Messages before and after entering **clear ufd-disable uplink-state-group** Command

```
02:36:43: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Gi 0/46
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Gi 0/46
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Te
13/0
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Te
13/1
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Te
13/3
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Te
13/5
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 13/0
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 13/1
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 13/3
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 13/5


02:37:29: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Gi 0/47
02:37:29: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Gi 0/47
02:37:29 : UFD: Group:3, UplinkState: DOWN

02:37:29: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed uplink state group state to down: Group 3
02:37:29: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Te
13/6
02:37:29: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 13/6


02:38:31 : UFD: Group:3, UplinkState: UP
02:38:31: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed uplink state group state to up: Group 3

02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD
error-disabled: Te 13/0
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD
error-disabled: Te 13/1
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD
error-disabled: Te 13/3
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD
error-disabled: Te 13/5
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD
error-disabled: Te 13/6
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te 13/0
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te 13/1
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te 13/3
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te 13/5
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Te 13/6
```

Displaying Uplink Failure Detection

To display information on the Uplink Failure Detection feature, enter any of the following **show** commands:

Show Command Syntax	Description
show uplink-state-group [*group-id***] [detail]** Command Mode: EXEC	Displays status information on a specified uplink-state group or all groups. Valid *group-id* values are 1 to 16. **detail** displays additional status information on the upstream and downstream interfaces in each group (see Figure 54-3).
show interfaces *interface* Command Mode: EXEC	Displays the current status of a port or port-channel interface assigned to an uplink-state group. *interface* specifies one of the following interface types: Fast Ethernet: Enter **fastethernet** *slot*/*port.* 1-Gigabit Ethernet: Enter **gigabitethernet** *slot*/*port*. 10-Gigabit Ethernet: Enter **tengigabitethernet** *slot*/*port*. Port channel: Enter **port-channel** {1-512}. If a downstream interface in an uplink-state group has been disabled (Oper Down state) by uplink-state tracking because an upstream port went down, the message error-disabled[UFD] is displayed in the output (see Figure 54-4).
show running-config uplink-state-group [*group-id***]** Command Mode: EXEC Or **show configuration** Command Mode: UPLINK-STATE-GROUP	Displays the current configuration of all uplink-state groups (Figure 54-5) or a specified group (Figure 54-6). Valid *group-id* values are 1 to 16.

Figure 54-3. show uplink-state-group Command Output

```
FTOS# show uplink-state-group

Uplink State Group: 1   Status: Enabled, Up
Uplink State Group: 3   Status: Enabled, Up
Uplink State Group: 5   Status: Enabled, Down
Uplink State Group: 6   Status: Enabled, Up
Uplink State Group: 7   Status: Enabled, Up
Uplink State Group: 16  Status: Disabled, Up


FTOS# show uplink-state-group 16
Uplink State Group: 16  Status: Disabled, Up


FTOS#show uplink-state-group detail
(Up): Interface up   (Dwn): Interface down   (Dis): Interface disabled

Uplink State Group    : 1       Status: Enabled, Up
Upstream Interfaces   :
Downstream Interfaces :

Uplink State Group    : 3       Status: Enabled, Up
Upstream Interfaces   : Gi 0/46(Up) Gi 0/47(Up)
Downstream Interfaces : Te 13/0(Up) Te 13/1(Up) Te 13/3(Up) Te 13/5(Up) Te 13/6(Up)

Uplink State Group    : 5       Status: Enabled, Down
Upstream Interfaces   : Gi 0/0(Dwn) Gi 0/3(Dwn) Gi 0/5(Dwn)
Downstream Interfaces : Te 13/2(Dis) Te 13/4(Dis) Te 13/11(Dis) Te 13/12(Dis) Te 13/13(Dis)
                        Te 13/14(Dis) Te 13/15(Dis)

Uplink State Group    : 6       Status: Enabled, Up
Upstream Interfaces   :
Downstream Interfaces :

Uplink State Group    : 7       Status: Enabled, Up
Upstream Interfaces   :
Downstream Interfaces :

Uplink State Group    : 16      Status: Disabled, Up
Upstream Interfaces   : Gi 0/41(Dwn) Po 8(Dwn)
Downstream Interfaces : Gi 0/40(Dwn)
```

Figure 54-4. show interfaces Command: UFD Output

```
FTOS#show interfaces gigabitethernet 7/45
GigabitEthernet 7/45 is up, line protocol is down (error-disabled[UFD])
Hardware is Force10Eth, address is 00:01:e8:32:7a:47
    Current address is 00:01:e8:32:7a:47
Interface index is 280544512
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 1000 Mbit, Mode auto
Flowcontrol rx off tx off
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:25:46
Queueing strategy: fifo
Input Statistics:
     0 packets, 0 bytes
     0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     0 Multicasts, 0 Broadcasts
     0 runts, 0 giants, 0 throttles
     0 CRC, 0 overrun, 0 discarded
Output Statistics:
     0 packets, 0 bytes, 0 underruns
     0 64-byte pkts, 0 over 64-byte pkts, 0 over 127-byte pkts
     0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
     0 Multicasts, 0 Broadcasts, 0 Unicasts
     0 throttles, 0 discarded, 0 collisions
Rate info (interval 299 seconds):
     Input 00.00 Mbits/sec,          0 packets/sec, 0.00% of line-rate
     Output 00.00 Mbits/sec,          0 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:01:23
```

Figure 54-5. show running-config uplink-state-group Command: UFD Output

```
FTOS#show running-config uplink-state-group
!
no enable
uplink state track 1
downstream GigabitEthernet 0/2, 4, 6, 11-19
upstream TengigabitEthernet 0/48, 52
upstream PortChannel 1
!
uplink state track 2
downstream GigabitEthernet 0/1, 3, 5, 7-10
upstream TengigabitEthernet 0/56, 60
```

Figure 54-6. show configuration Command: UFD Output

```
FTOS(conf-uplink-state-group-16)# show configuration
!
uplink-state-group 16
no enable
description test
downstream disable links all
downstream GigabitEthernet 0/40
upstream GigabitEthernet 0/41
upstream Port-channel 8
```

Sample Configuration: Uplink Failure Detection

Figure 54-7 shows a sample configuration of Uplink Failure Detection on a switch/router in which you:

- Configure uplink-state group 3.
- Add downstream links Gigabitethernet 0/1, 0/2, 0/5, 0/9, 0/11, and 0/12.
- Configure two downstream links to be disabled if an upstream link fails.
- Add upstream links Gigabitethernet 0/3 and 0/4.
- Add a text description for the group.
- Verify the configuration with various **show** commands.

Figure 54-7. Configuring Uplink Failure Detection

```
FTOS(conf)# uplink-state-group 3
00:08:11: %STKUNIT0-M:CP %IFMGR-5-ASTATE_UP: Changed uplink state group Admin state to up:
Group 3
FTOS(conf-uplink-state-group-3)# downstream gigabitethernet 0/1-2,5,9,11-12
FTOS(conf-uplink-state-group-3)# downstream disable links 2
FTOS(conf-uplink-state-group-3)# upstream gigabitethernet 0/3-4
00:10:00: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled:
Gi 0/1
FTOS#
00:10:00: %STKUNIT0-M:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Gi 0/1
FTOS(conf-uplink-state-group-3)# description Testing UFD feature


FTOS(conf-uplink-state-group-3)# show config
!
uplink-state-group 3
description Testing UFD feature
downstream disable links 2
downstream GigabitEthernet 0/1-2,5,9,11-12
upstream GigabitEthernet 0/3-4
FTOS(conf-uplink-state-group-3)#
FTOS(conf-uplink-state-group-3)#exit
FTOS(conf)#exit
FTOS#
00:13:06: %STKUNIT0-M:CP %SYS-5-CONFIG_I: Configured from console by  console


FTOS# show running-config uplink-state-group
!
uplink-state-group 3
description Testing UFD feature
downstream disable links 2
downstream GigabitEthernet 0/1-2,5,9,11-12
upstream GigabitEthernet 0/3-4


FTOS# show uplink-state-group 3

Uplink State Group: 3   Status: Enabled, Up


FTOS# show uplink-state-group detail

(Up): Interface up   (Dwn): Interface down   (Dis): Interface disabled
Uplink State Group    : 3       Status: Enabled, Up
Upstream Interfaces   : Gi 0/3(Up) Gi 0/4(Dwn)
Downstream Interfaces : Gi 0/1(Dis) Gi 0/2(Dwn) Gi 0/5(Dwn) Gi 0/9(Dwn) Gi 0/11(Dwn)
                        Gi 0/12(Dwn)
```

55

Upgrade Procedures

Find the upgrade procedures

Go to the *FTOS Release Notes* for your system type to see all the requirements to upgrade to the desired FTOS version. Follow the procedures in the *FTOS Release Notes* for the software version you wish to upgrade *to*.

Get Help with upgrades

Direct any questions or concerns about FTOS Upgrade Procedures to Dell Force10 Technical Support Center. You can reach Technical Support:

- On the Web: www.force10networks.com/support/
- By email: support@force10networks.com
- By phone: US and Canada: 866.965.5800, International: 408.965.5800

56

VLAN

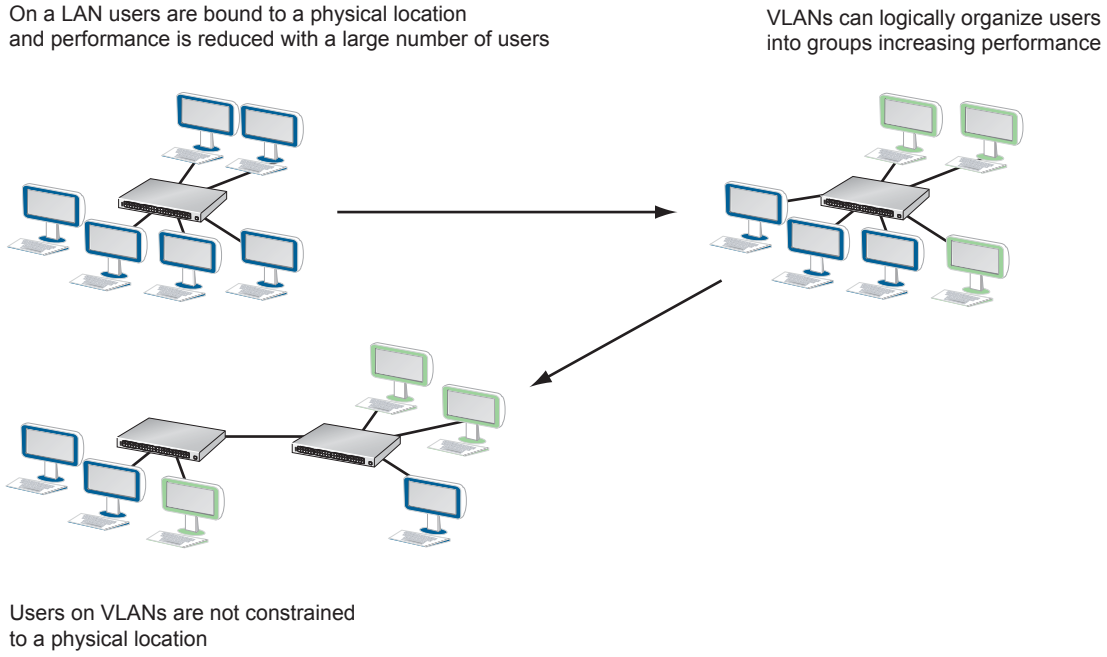
VLANs are supported on platforms: C E S

This chapter contains the following configuration topics:

Virtual LAN Overview

A Local Area Network (LAN) is a collection of devices in the same broadcast domain. As a network increases in size, segmenting a single broadcast domain into multiple domains improves scalability, manageability, and security. However, doing so using physically with separate switches and routers is both static and expensive.

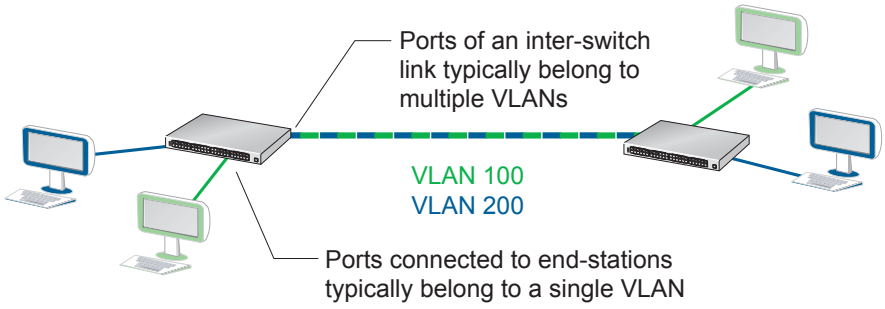
Virtual LANs (VLANs) are a cost-effective method of segmenting and organizing a network. A single switch can be divided into multiple broadcast domains so that devices can be grouped and isolated; each logical segment is virtual LAN. Applying VLANs reduces broadcast traffic, introduces flexibility in the placement of devices on the network, and increases network security by allowing separate policies to be applied to each group.



Port-based VLANs

On FTOS, a VLAN is a user-defined group of ports (there is also the concept of protocol-based VLANs). Ports in different VLANs do not communicate unless routing is configured between them. A port may belong to more than one VLAN. Typically, ports connected to a host belong to only one VLAN, and ports on an inter-switch link belong to more than one VLAN; these ports are sometimes called *trunk* ports.

Figure 56-1. VLAN Membership

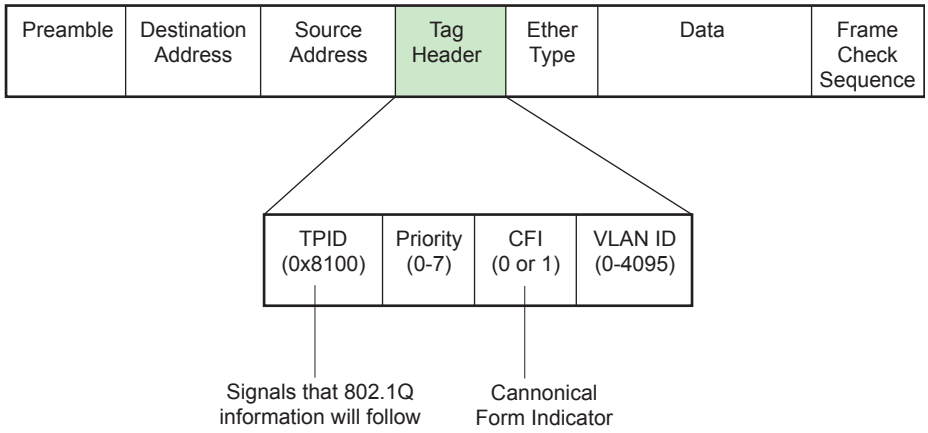


VLAN Tagging

Since a port may belong to more than one VLAN, the switch must be able to identify the VLAN two which a broadcast frame belongs. For this case, IEEE 802.1Q defines a method of marking frames to indicate the VLAN on which the frame originated.

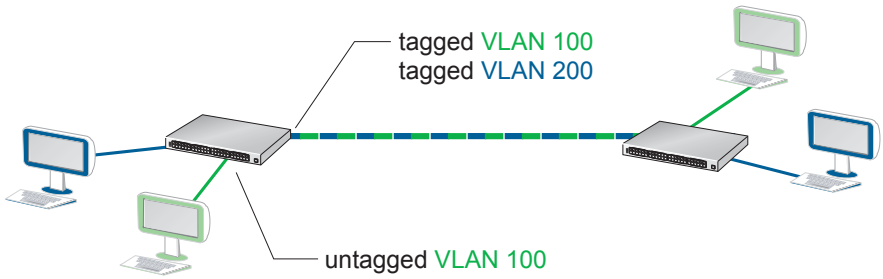
The marker, called a *VLAN tag*, is 4 bytes and is inserted after the source MAC in the Ethernet frame header, as shown in Figure 56-2. The tag is preserved as the frame moves through the network so that intermediate switches can forward the frame appropriately.

Figure 56-2. 802.1Q VLAN Tag



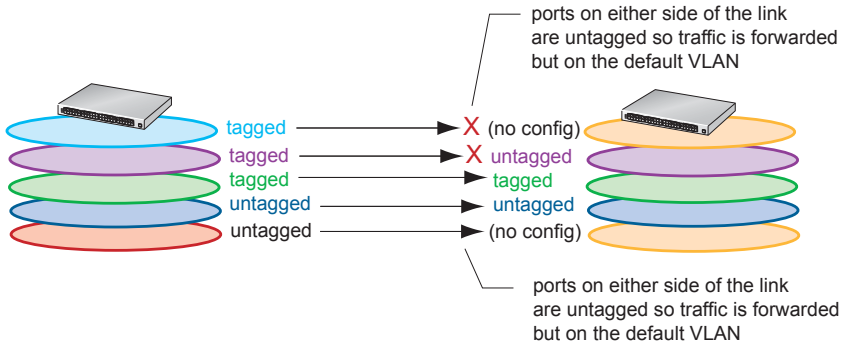
Ports that belong to more than one VLAN insert VLAN tags into frames and so they are called *tagged ports*. Ports that belong to a single VLAN, do not insert VLAN tags into frames and are called *untagged ports*. When you add a port to a VLAN, you must specify whether the port should be tagged or untagged.

Figure 56-3. Tagged and Untagged Ports



Ports on either side of the link must have the same tagged/untagged designation, and if tagged, must belong to the same VLAN. Else, the frame is dropped.

Figure 56-4. Switch Behavior for Tagged/Untagged Port Mismatch



Default VLAN

The Default VLAN and is part of the system startup configuration, and is by default VLAN 1. You may make another VLAN the default VLAN. The default VLAN cannot be deleted, disabled, or configured (you cannot assign it an IP address), and only untagged interfaces can belong to it.

When an interface is configured as a switchport automatically places it in the default VLAN as an untagged interfaces. All switchports must belong to at least one VLAN, so to remove a switchport from the default VLAN, you must place it as tagged or untagged in some other VLAN, or remove the **switchport** configuration.

Implementation Information

- FTOS supports up to 4093 port-based VLANs plus 1 Default VLAN.
- E-Series ExaScale FTOS versions earlier than 8.2.1.0 for the E-Series ExaScale support 2094 VLANs.

Configuring VLANs

Configuring a VLAN is a two-step process:

1. Create a VLAN. See page 1103.
2. Add a switchport as a tagged or untagged member port. See page 1104.
3. Optionally, assign an IP address to a VLAN to enable routing between VLANs. See page 1105.

Related Configuration Tasks

- Use a Native VLAN on Trunk Ports on page 1106
- Change the Default VLAN ID on page 1107

- Set the Null VLAN as the Default VLAN on page 1107
- Enable VLAN Interface Counters on page 1108

Related Protocols and Topics

The following protocols and topics are premised on VLANs, and contain more information about the utility of VLANs:

- 802.1X
- Chapter 17, GARP VLAN Registration Protocol.
- Chapter 46, Service Provider Bridging
- Chapter 40, Per-VLAN Spanning Tree Plus.

Create a VLAN

A VLAN is created when you assign it a VLAN ID.

Task	Command Syntax	Command Mode
Create a VLAN.	**interface vlan** *vlan-id*	CONFIGURATION
Display all VLANs.	**show vlan** *vlan-id*	EXEC Privilege

```
FTOS#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status    Q Ports
*   1      Inactive  U So 9/4-11
    2      Active    U Gi 0/1,18
    3      Active    U Gi 0/2,19
    4      Active    T Gi 0/3,20
    5      Active    U Po 1
    6      Active    U Gi 0/12
                     U So 9/0
```

A VLAN is active only if the VLAN contains interfaces and those interfaces are up. VLAN 1 is inactive because it contains the interfaces that are not up. When you delete a VLAN (**no interface vlan** *vlan-id*), any interfaces assigned to that VLAN are reassigned to the default VLAN as untagged.

Assign Interfaces to VLANs

A port may either be an *untagged* member of a single VLAN, or a *tagged* member of perhaps multiple VLANs.

- **Untagged Ports —** ports that do not append an 802.1Q VLAN tag to frames on egress, and do not accept tagged frames on ingress (tagged frames are dropped). Untagged ports must be connected to VLAN-unaware devices.
- **Tagged Ports** — ports that append an 802.1Q tag to frames on egress, and accept only tagged frames on ingress (untagged frames are dropped). Tagged ports must be connected to VLAN-aware devices.

When you place configure an enabled port as a switchport, the port is placed in the default VLAN. To remove a switchport from the default VLAN, remove the **switchport** configuration. To move the port to another VLAN, add it to the desired VLAN as either a tagged or untagged member.

Step	Task	Command Syntax	Command Mode
1	Assign a switchport to a VLAN. `FTOS(conf)#int vlan 4` `FTOS(conf-if-vlan)#tagged po 1` `FTOS(conf-if-vlan)#show conf` `!` `interface Vlan 4` `no ip address` `tagged Port-channel 1`	[**tagged** | **untagged**] *interface*	INTERFACE VLAN
2	Display all switchports and the VLANs of which they are members.	**show vlan**	EXEC Privilege

Step	Task	Command Syntax	Command Mode

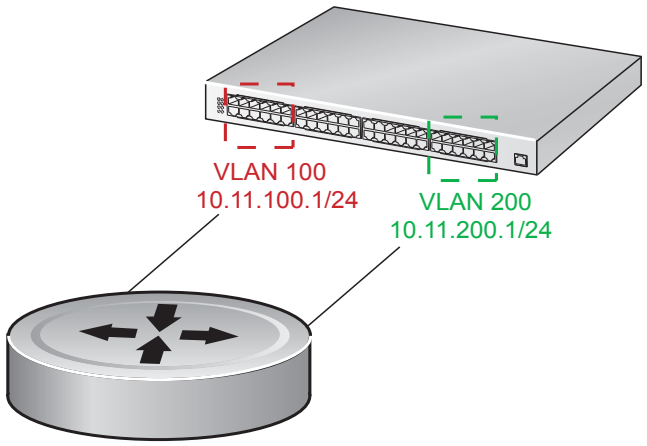
```
FTOS#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status    Q Ports
*   1      Inactive
    2      Active    T Po1(So 0/0-1)
                     T Gi 3/0
    3      Active    T Po1(So 0/0-1)
                     T Gi 3/1
    4      Active    T Po1(So 0/0-1)
FTOS(conf)#int vlan 4
FTOS(conf-if-vlan)#untagged gi 3/2
FTOS(conf-if-vlan)#show config
!
interface Vlan 4
 no ip address
 untagged GigabitEthernet 3/2
FTOS#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status    Q Ports
*   1      Inactive
    2      Active    T Po1(So 0/0-1)
                     T Gi 3/0
    3      Active    T Po1(So 0/0-1)
                     T Gi 3/1
    4      Active    U Gi 3/2
```

Enable Routing between VLANs

Each VLAN is a broadcast domain. For devices in two different broadcast domains to communicate, traffic must be routed, and so in this case each VLAN needs an IP address.

Note: The **shutdown** command marks a physical interface as unavailable for traffic. Disabling a VLAN or a port-channel results in a different behavior. When a VLAN is disabled, the Layer 3 functions within that VLAN are disabled, but Layer 2 traffic continues to flow.

Figure 56-5. Communicating between VLANs



Task	Command Syntax	Command Mode
Assign an IP address to a VLAN interface.	**ip address** *address/mask*	INTERFACE VLAN

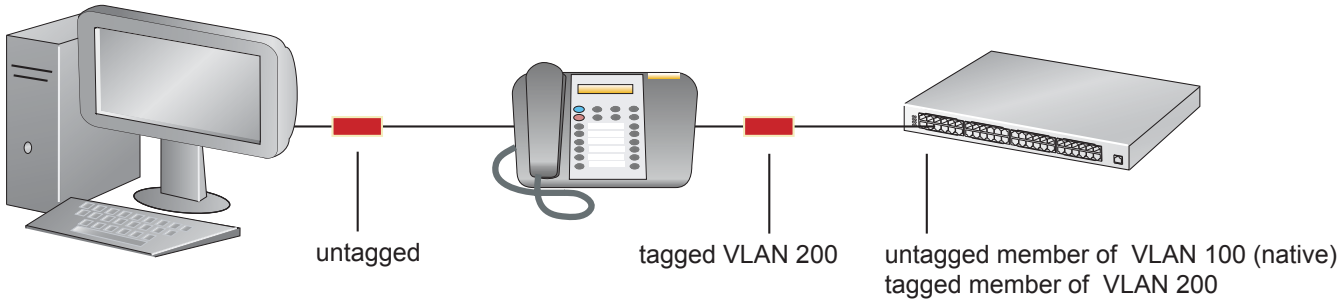
Use a Native VLAN on Trunk Ports

Traditionally, a port may either be an untagged member of a single VLAN or a tagged member of multiple VLANs. However, FTOS allows you to make a port an untagged member and a tagged member of VLANs, concurrently.

Ports that are an untagged and tagged member concurrently are called hybrid ports; physical ports and port-channels may be hybrid ports. On a hybrid port, the VLAN of which the port is an untagged member is the *native VLAN*.

A Native VLAN is useful on trunk ports, which receive both tagged and untagged traffic (a trunk port is a port that carries traffic for one or more VLANs on the switch). The classic example is a VOIP phone and a PC connected to the same port of a switch, where the VOIP phone generates packets tagged with VLAN ID = VOICE VLAN, and the PC generates untagged packets.

Figure 56-6. Using Native VLANs with PC/VOIP Phone



To configure a port so that it has a native VLAN:

Step	Task	Command	Command Mode
1	Remove any Layer 2 or Layer 3 configurations from the interface.		INTERFACE
	If the port has any configurations on it when you enter the command **portmode hybrid**, FTOS rejects the configuration, citing the following message: `% Error: Port is in Layer-2 mode <slot/port>.`		
2	Configure the interface for hybrid mode.	**portmode hybrid**	INTERFACE
3	Configure the interface for switchport mode.	**switchport**	INTERFACE
4	Add the interface as a member of one or more VLANs.	[**tagged** | **untagged**]	VLAN INTERFACE

Change the Default VLAN ID

By default, VLAN 1 is the Default VLAN. You can make another VLAN the default (which then enables you to configure VLAN 1).

Task	Command Syntax	Command Mode
Make a VLAN other than VLAN 1 the default VLAN.	**default vlan-id**	CONFIGURATION

Set the Null VLAN as the Default VLAN

In a Carrier Ethernet for Metro Service environment, service providers who perform frequent reconfigurations for customers with changing requirements occasionally enable multiple interfaces, each connected to a different customer, before the interfaces are fully configured. This presents a vulnerability because both interfaces are initially placed in the native VLAN, VLAN 1, and for that period customers are able to access each other's networks. FTOS has a Null VLAN to eliminate this vulnerability. When you enable the Null VLAN, all ports are placed into by it default, so that even if you activate the physical ports of multiple customers, no traffic is allowed to traverse the links until each port is place in another VLAN.

Task	Command Syntax	Command Mode
Disable the default VLAN, so that all ports belong to the Null VLAN until configured as a member of another VLAN.	**default-vlan disable** Default: the default VLAN is enabled (**no default-vlan disable**).	CONFIGURATION

Enable VLAN Interface Counters

Use a Native VLAN on Trunk Ports is available only on platform: E X

Note: VLAN egress counters might be higher than expected because source-suppression drops are counted.

Task	Command Syntax	Command Mode
Configure ingress, egress or both counters for VLAN interfaces.	**enable vlan-counter** [**ingress** | **egress** | **all**]	CONFIGURATION

57

Virtual Routing and Forwarding (VRF)

Virtual Routing and Forwarding (VRF) (VRF) is supported only on platform: E

VRF allows a physical router to partition itself into multiple Virtual Routers (VRs). The control and data plane are isolated in each VR so that traffic does NOT flow across VRs.Virtual Routing and Forwarding (VRF) allows multiple instances of a routing table to co-exist within the same router at the same time.

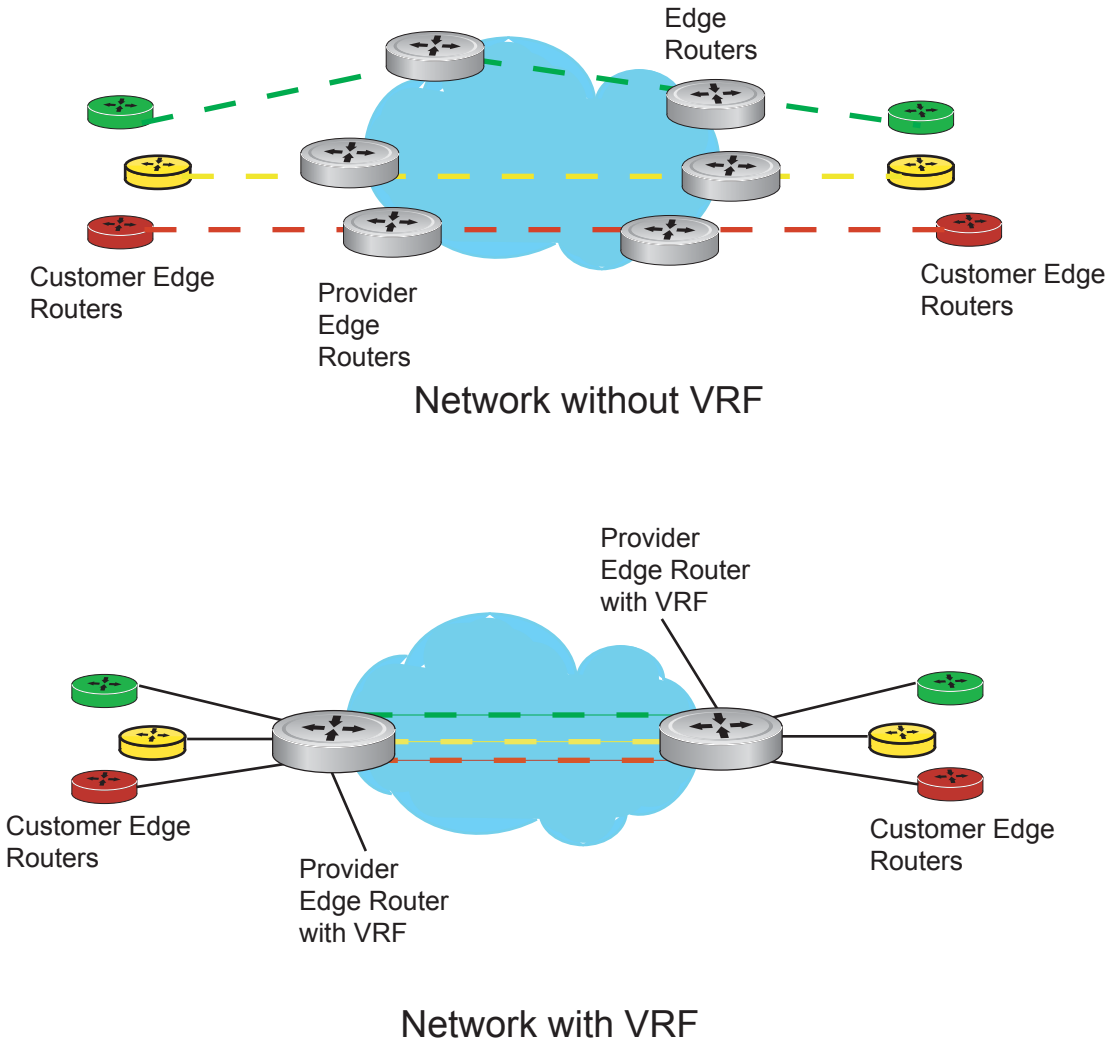
VRF improves functionality by allowing network paths to be segmented without using multiple devices. Using VRF also increases network security and can eliminate the need for encryption and authentication due to traffic segmentation. Internet service providers (ISPs) often take advantage of VRF to create separate virtual private networks (VPNs) for customers; VRF is also referred to as VPN routing and forwarding.

VRF is implemented in a network device by having a distinct Forwarding Information Base (FIB) per VRF instance. A network device has the ability to configure different virtual routers, so that each has its own FIB that is not accessible to any other virtual router instance on the same device.

VRF acts like a logical router; while a physical router may include many routing tables, a VRF instance uses only a single routing table. VRF uses a forwarding table that designates the next hop for each data packet, a list of devices that may be called upon to forward the packet, and a set of rules and routing protocols that govern how the packet is forwarded. These VRF forwarding tables prevent traffic from being forwarded outside a specific VRF path and also keep out traffic that should remain outside the VRF path.

VRF uses interfaces to distinguish routes for different VRF instances. Interfaces in a VRF can be either physical (Ethernet port or port channel) or logical (VLANs). Starting in release 8.4.1.0, you can configure identical or overlapping IP subnets on different interfaces if each interface belongs to a different VRF instance.

Figure 57-1. VRF Network Example



VRF Configuration Notes

On E-Series routers, Dell Force10 VRF supports up to 15 VRF instances: 1 to 14 and the default VRF (0).

Although there is no restriction on the number of VLANs that can be assigned to a VRF instance, the total number of routes supported in VRF is limited by the size of the IPv4 FIB table in the CAM.

VRF is implemented in a network device by using Forwarding Information Bases (FIBs). Each VRF uses one FIB.

A network device may have the ability to configure different virtual routers, where each one has its own FIB that is not accessible to any other virtual router instance on the same device.

Only Layer 3 interfaces can belong to a VRF. VRF is supported on following types of interface:

- physical Ethernet interfaces
- physical Sonet interfaces
- port-channel interfaces (static & dynamic using LACP)
- VLAN interfaces
- loopback interfaces

VRF supports route redistribution between routing protocols (including static routes) only when the routes are within the same VRF.

FTOS uses both the VRF name and VRF ID to manage VRF instances. The VRF name and VRF ID number are assigned using the **ip vrf** command. The VRF ID is displayed in **show ip vrf** command output.

The VRF ID is not exchanged between routers. VRF IDs are local to a router.

VRF supports some routing protocols only on the default VRF (*default-vrf*) instance. Table 57-1 displays the software features supported in VRF and whether they are supported on all VRF instances or only the default VRF.

Table 57-1.

Feature/Capability	Supported?	Note
Configuration rollback for commands introduced or modified	Yes	
LLDP protocol on the port	Yes	
802.1x protocol on the VLAN port	Yes	Supported only for default-VRF
OSPF, RIP, ISIS, BGP on physical and logical interfaces	Yes	OSPF supported on all VRF ports. Others supported only on default-VRF ports
Dynamic Port-channel (LACP) on VLAN port or a Layer 3 port	Yes	
Static Port-channel as VLAN port or a Layer 3 port	Yes	
Port-monitoring	Yes	Mirroring port (MG) has to be in default-VRF
BFD on physical and logical interfaces	Yes	Supported on default-VRF ports only
PVST, MSTP, RSTP and 802.1D STP for VLANs	Yes	
FRRP (if applicable) for VLANs	Yes	
Multicast protocols (PIM-SM, PIM-DM, MSDP)	Yes	Supported on default-VRF ports only
Layer 3 (IPv4/IPv6) ACLs, TraceLists, PBR, QoS on VLANs	Yes	ACLs supported on all VRF VLAN ports. TraceLists are common for entire line card (except on ExaScale). PBR supported on default-VRF only. QoS not supported on VLANs.

Table 57-1.

Feature/Capability	Supported?	Note
Layer 3 (IPv4/IPv6) ACLs, TraceLists, PBR, QoS on physical interfaces and LAGs		ACLs supported on all VRF ports. TraceLists are common for entire line card (except on ExaScale). PBR supported on default-VRF only. QoS supported on all VRF ports.
IPv4 ARP and IPv6 Neighbor Discovery	Yes	ARP is VRF-aware. IPv6 is supported only for default-VRF.
Layer 2 ACLs on VLANs	Yes	
FEFD	Yes	
Layer 2 QoS	Yes	
Support for storm-control (broadcast and unknown-unicast)	Yes	
sFlow	Yes	Extended-gateway information supported for default-VRF only
VRRP on physical and logical interfaces	Yes	Supported on all VRF instances, including the default-VRF
Secondary IP Addresses	Yes	
Following IPv6 capabilities		
Basic	Yes	Supported on default-VRF only
OSPFv3	Yes	Supported on default-VRF only
ISIS	Yes	Supported on default-VRF only
BGP	Yes	Supported on default-VRF only
ACL	Yes	ACL supported on all VRF ports
Multicast	Yes	Supported on default-VRF only
NDP	Yes	Supported on default-VRF only
RAD	Yes	Supported on default-VRF only
Ingress/Egress Storm-Control (per-interface/global)	Yes	

CAM Profiles

Layer 3 CAM resources are shared among all VRF instances. To ensure that each VRF instance has sufficient CAM space:

- On an E-Series Terascale platform, use the **cam-profile ipv4-vrf** or **cam-profile ipv4-v6-vrf** command and reload the system command to activate the VRF CAM profile for IPv4 or IPv6.

- On an E-Series Exascale platform, use the **cam-profile** command to set the CAM size. Then select and enable VRF microcode for use with the VRF CAM-profile template, and reload the system to activate the profile. You can set the CAM size to 40M (default) which supports both IPv4 and IPv6 or 10M which supports only IPv4.

Note: Any physical (port or port channel) or VLAN interface assigned to a VRF uses two CAM entries instead of one for each route and host entry.

Table 57-2 and Table 57-3 each show the required CAM settings for IPv4 and IPv6.

Note: VRF is supported in single CAM cards only.

Table 57-2. IPv4-VRF CAM Profile (Single CAM card)

CAM Profile Table	Allocation (K)
L2FIB	32K
L2ACL	3K
IPv4FIB	160K
IPv4ACL	2K
IPv4Flow	12K
EgL2ACL	1K
EgIPv4ACL	12K
Reserved	2K
IPv6FIB	0K
IPv6ACL	0K
IPv6Flow	0K
EgIPv6ACL	0K

Note: When configuring the IPv6 CAM profile, the CAM tables that are carved within l2acl and ipv4Flow tables remain at default values. For more information on the CAM and CAM profiling, refer to Chapter 11, “Content Addressable Memory,” on page 281

Table 57-3. IPv4-v6-VRF CAM Profiles (Single CAM card)

CAM Profile Table	Allocation (K)
L2FIB	32K
L2ACL	3K
IPv4FIB	64K
IPv4ACL	1K
IPv4Flow	12K
EgL2ACL	1K

Table 57-3. IPv4-v6-VRF CAM Profiles (Single CAM card)

CAM Profile Table	Allocation (K)
EgIPv4ACL	11K
Reserved	2K
IPv6FIB	18K
IPv6ACL	4K
IPv6Flow	3K
EgIPv6ACL	1K

DHCP

DHCP requests are not forwarded across VRF instances. The DHCP client and server must be on the same VRF instance.

IP addressing

Starting in release 8.4.1.0, you can configure identical or overlapping IP subnets on different interfaces if each interface belongs to a different VRF instance. In previous releases, VRF did not support the same IP address on multiple interfaces in different VRF instances.

VRF Configuration

Note: Starting in FTOS 8.4.2.1, when VRF microcode is loaded on an E-Series ExaScale or TeraScale router, the **ip vrf** [**default-vlan** | *vrf-name*] command is deprecated, and is replaced by the **ip vrf** *vrf-name vrf-id* command. The **ip vrf-vlan-block**, **start-vlan-id default-vrf**, and **start-vlan-id** *vlan-start-id* commands are also deprecated.

The VRF configuration tasks are:

1. Load the VRF CAM Profile
2. Enable VRF
3. Assign an Interface to a VRF

You can also:

- View VRF instance information
- Connect an OSPF process to a VRF instance
- Configure VRRP on a VRF Interface

Load the VRF CAM Profile

On an **E-series Terascale** platform, select the IPv4 or IPv6 CAM profile used to support VRF and reload the system to activate the profile.

Step	Task	Command Syntax	Command Mode
1	Select the appropriate CAM profile for your system.	**cam-profile [ipv4-vrf microcode ipv4-vrf | ipv4-v6-vrf microcode ipv4-v6-vrf]** *default-vrf | vrf-name*	CONFIGURATION
2	Reload the system to implement the new CAM profile.	**reload**	EXEC

On an **E-series Exascale** platform, configure the CAM size used to support VRF. Then enable VRF microcode for use with the CAM profile and reload the system to activate the profile. You can set the CAM size to 40M (default) which supports both IPv4 and IPv6 or 10M which supports only IPv4.

Step	Task	Command Syntax	Command Mode
1	Configure the CAM size (10M or 40M) to be used by your system.	**cam-profile** *name* **[10M-cam]** Default: 40M	CONFIGURATION
2	Select and enable VRF microcode for use with the VRF CAM profile.	**microcode vrf enable**	CAM-PROFILE
3	Reload the system to activate the new CAM profile.	**reload**	EXEC

Enable VRF

VRF is enabled by default when VRF microcode is loaded on an E-Series ExaScale or TeraScale router. On an E-Series router, Dell Force10 VRF supports up to 15 VRF instances: 1 to 14 and the default VRF (0).

A VRF name is not exchanged between routers. VRF IDs are local to a router. The following features and functionality are supported only on the default VRF (0) instance:

- ISIS
- BGP
- RIP
- IPv6
- Multicast
- Static ARP

Task	Command Syntax	Command Mode
Create a non-default VRF instance by specifying a name and VRF ID number, and enter VRF configuration mode. The default VRF 0 is automatically configured when a router with VRF loaded in CAM boots up.	**ip vrf** *vrf-name vrf-id* VRF ID range: 1 to 14 and 0 (default VRF)	CONFIGURATION

Note: Starting in FTOS 8.4.2.1, when VRF microcode is loaded on a E-Series ExaScale or TeraScale router, the **ip vrf [default-vlan |** *vrf-name]* command is deprecated, and is replaced by the **ip vrf** *vrf-name vrf-id* command. The **ip vrf-vlan-block**, **start-vlan-id default-vrf**, and **start-vlan-id** *vlan-start-id* commands are also deprecated.

Assign an Interface to a VRF

You must enter the **ip vrf forwarding** command before you configure the IP address or any other setting on an interface.

Note: Starting in release 8.4.1.0, you can configure an IP address or subnet on a physical or VLAN interface that overlaps the same IP address or subnet configured on another interface only if the interfaces are assigned to different VRFs. If two interfaces are assigned to the same VRF, you cannot configure overlapping IP subnets or the same IP address on them.

When you assign a VLAN interface to a VRF instance, the following conditions apply:

- VLANs assigned to the same VRF have the same MAC address. VLANs assigned to different VRFs have different MAC addresses. The last four bits of a VLAN's MAC address correspond to the VRF ID configured with the **ip vrf** command.
- On a switch port on which multiple VLANs are assigned to different VRFs, the source MAC address in packets routed on a VRF may not be the same as the MAC address distributed in ARP requests. As a result, security applications running on neighboring routers that check the source MAC address in incoming packets may find that the address does not match the ARP-learned MAC address.
- You can assign a static ARP only to a VLAN that is mapped to the default VRF (0) instance.
- By default, all VLANs (4096) are associated with the default VRF until you reassign them to a non-default VRF with the **ip vrf forwarding** command. You can assign up to 4096 VLANs to a non-default VRF instance.
- VLANs used with VRF must be Layer 3 VLANs. Layer 2 VLANs can be configured for non-VRF use. Refer to Chapter 56, "VLAN," on page 1099 for complete information.
- All VLAN member ports must be removed from a VLAN that you move from one VRF instance to another.

Task	Command Syntax	Command Mode
Assign an interface to a VRF instance.	**ip vrf forwarding** *vrf-name*	INTERFACE

View VRF instance information

To display information about VRF configuration, enter the **show ip vrf** command.

Task	Command Syntax	Command Mode
Display the interfaces assigned to a VRF instance. To display information on all VRF instances (including the default VRF 0), do not enter a value for *vrf-name*.	**show ip vrf** *[vrf-name]*	EXEC

Connect an OSPF process to a VRF instance

OSPF routes are supported on all VRF instances. Refer to Chapter 32, Open Shortest Path First (OSPFv2 and OSPFv3) for complete OSPF configuration information.

Assign an OSPF process to a VRF instance . Return to CONFIGURATION mode to enable the OSPF process. The OSPF Process ID is the identifying number assigned to the OSPF process, and the Router ID is the IP address associated with the OSPF process..

Task	Command Syntax	Command Mode
Enable the OSPFv2 process globally for a VRF instance. Enter the VRF key word and instance name to tie the OSPF instance to the VRF. All network commands under this OSPF instance are subsequently tied to the VRF instance. *process-id range*: 0-65535	**router ospf** *process-id* **vrf** *vrf name*	CONFIGURATION

Once the OSPF process and the VRF are tied together, the OSPF Process ID cannot be used again in the system.

Configure VRRP on a VRF Interface

Starting in release 8.4.1.0, you can configure the VRRP feature on interfaces that belong to a VRF instance. In previous releases, VRRP was not supported on interfaces that were configured for a non-default VRF.

In a virtualized network that consists of multiple VRFs, various overlay networks can exist on a shared physical infrastructure. Nodes (hosts and servers) that are part of the VRFs can be configured with IP static routes for reaching specific destinations through a given gateway in a VRF. VRRP provides high availability and protection for next-hop static routes by eliminating a single point of failure in the default static routed network. For more information, refer to Chapter 58, “Virtual Router Redundancy Protocol (VRRP),” on page 1127.

Sample VRF Configuration

The following configuration illustrates a typical VRF set up.

Figure 57-2. Set up OSPF and static routes

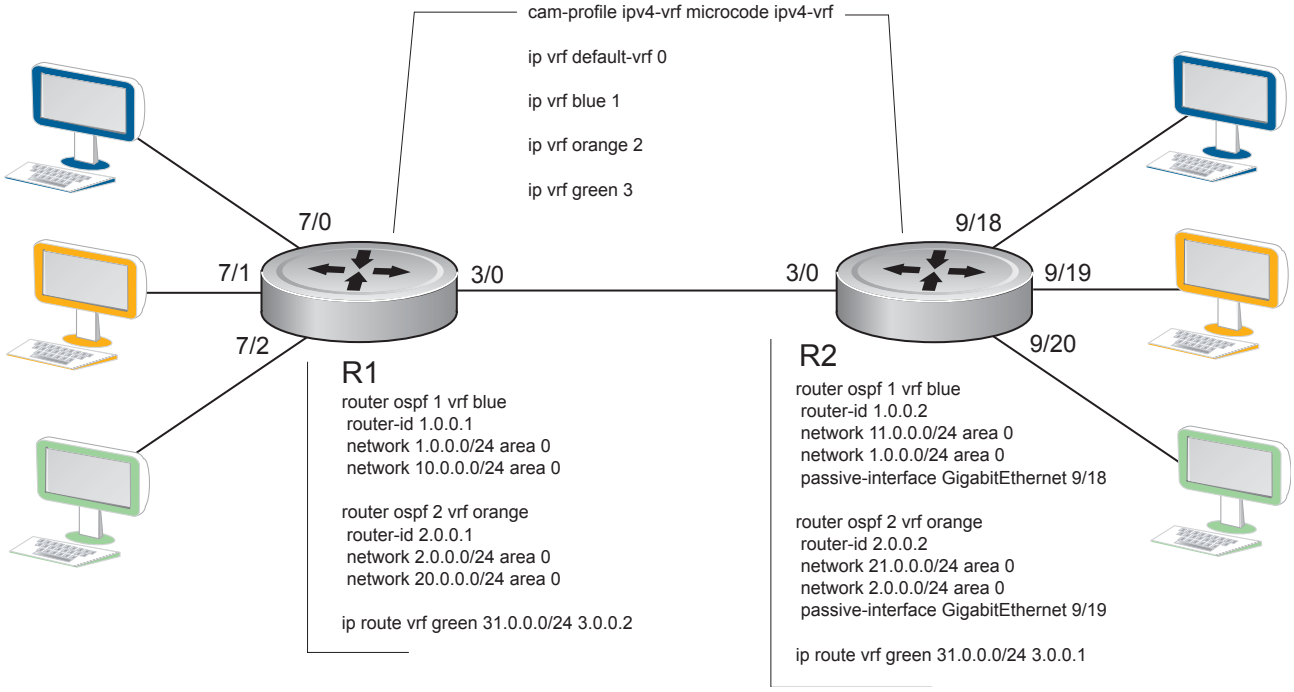
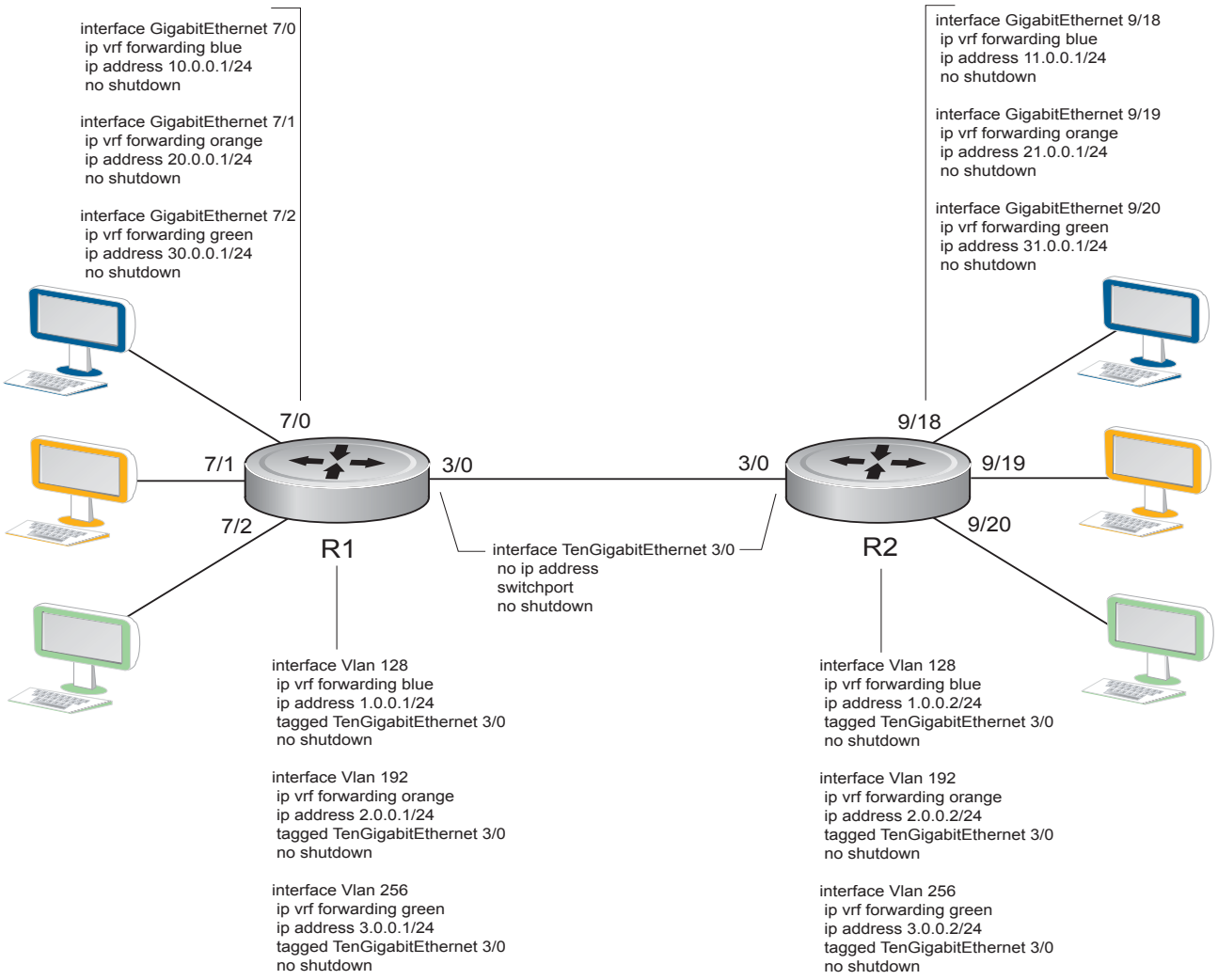


Figure 57-3. Set up VRF interfaces



The following example relates to the configuration shown in Figure 57-2 and Figure 57-3.

ROUTER 1

```
cam-profile ipv4-vrf microcode ipv4-vrf
!
ip vrf default-vrf 0
!
ip vrf blue 1
!
ip vrf orange 2
!
ip vrf green 3
!
interface TenGigabitEthernet 3/0
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 7/0
 ip vrf forwarding blue
 ip address 10.0.0.1/24
 no shutdown
!
interface GigabitEthernet 7/1
 ip vrf forwarding orange
 ip address 20.0.0.1/24
 no shutdown
!
interface GigabitEthernet 7/2
 ip vrf forwarding green
 ip address 30.0.0.1/24
 no shutdown
!
interface Vlan 128
 ip vrf forwarding blue
 ip address 1.0.0.1/24
 tagged TenGigabitEthernet 3/0
 no shutdown
!
interface Vlan 192
 ip vrf forwarding orange
 ip address 2.0.0.1/24
 tagged TenGigabitEthernet 3/0
 no shutdown
!
interface Vlan 256
 ip vrf forwarding green
 ip address 3.0.0.1/24
 tagged TenGigabitEthernet 3/0
 no shutdown
!
```

------continued next page -------

ROUTER 1 continued

```
router ospf 1 vrf blue
 router-id 1.0.0.1
 network 1.0.0.0/24 area 0
 network 10.0.0.0/24 area 0
!
router ospf 2 vrf orange
 router-id 2.0.0.1
 network 2.0.0.0/24 area 0
 network 20.0.0.0/24 area 0
!
ip route vrf green 31.0.0.0/24 3.0.0.2!
```

==========

ROUTER 2

```
cam-profile ipv4-vrf microcode ipv4-vrf
!
ip vrf default-vrf 0
!
ip vrf blue 1
!
ip vrf orange 2
!
ip vrf green 3
!
interface TenGigabitEthernet 3/0
 no ip address
 switchport
 no shutdown
!
interface GigabitEthernet 9/18
 ip vrf forwarding blue
 ip address 11.0.0.1/24
 no shutdown
!
interface GigabitEthernet 9/19
 ip vrf forwarding orange
 ip address 21.0.0.1/24
 no shutdown
!
interface GigabitEthernet 9/20
 ip vrf forwarding green
 ip address 31.0.0.1/24
 no shutdown
!
interface Vlan 128
 ip vrf forwarding blue
 ip address 1.0.0.2/24
 tagged TenGigabitEthernet 3/0
 no shutdown
```

------continued next page -------

ROUTER 2 continued

```
interface Vlan 192
 ip vrf forwarding orange
 ip address 2.0.0.2/24
 tagged TenGigabitEthernet 3/0
 no shutdown
!
interface Vlan 256
 ip vrf forwarding green
 ip address 3.0.0.2/24
 tagged TenGigabitEthernet 3/0
 no shutdown
!
router ospf 1 vrf blue
 router-id 1.0.0.2
 network 11.0.0.0/24 area 0
 network 1.0.0.0/24 area 0
 passive-interface GigabitEthernet 9/18
!
router ospf 2 vrf orange
 router-id 2.0.0.2
 network 21.0.0.0/24 area 0
 network 2.0.0.0/24 area 0
 passive-interface GigabitEthernet 9/19
!
ip route vrf green30.0.0.0/24 3.0.0.1
!
================================================================================
```

The following shows the output of the **show** commands on Router 1.

ROUTER 1

```
FTOS#show ip vrf
VRF-Name                        VRF-ID Interfaces

default-vrf                     0      Gi 2/0-89,
                                       Te 3/0-3,
                                       Gi 4/0-89,
                                       Gi 5/0-89,
                                       Gi 7/3-47,
                                       Gi 9/0-47,
                                       Gi 10/0-47,
                                       Gi 11/0-47,
                                       Gi 12/0-47,
                                       Gi 13/0-47,
                                       Ma 0/0,
                                       Ma 1/0,
                                       Nu 0,
                                       Vl 1
blue                            1      Gi 7/0,
                                       Vl 128
orange                          2      Gi 7/1,
                                       Vl 192
green                           3      Gi 7/2,
                                       Vl 256
```

------continued next page -------

ROUTER 1 continued

```
FTOS#show ip ospf 1 neighbor
Neighbor ID    Pri      State        Dead Time Address        Interface Area
1.0.0.2         1       FULL/DR      00:00:32 1.0.0.2         Vl 128    0

FTOS#sh ip ospf 2 neighbor
Neighbor ID    Pri      State        Dead Time Address        Interface Area
2.0.0.2         1       FULL/DR      00:00:37 2.0.0.2         Vl 192    0
FTOS#show ip route vrf blue

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set

       Destination        Gateway                     Dist/Metric Last Change
       -----------        -------                     ----------- -----------
  C    1.0.0.0/24         Direct, Vl 128                    0/0    00:20:48
  C    10.0.0.0/24        Direct, Gi 7/0                    0/0    00:10:06
  O    11.0.0.0/24        via 1.0.0.2, Vl 128             110/2    00:11:13

FTOS#show ip route vrf orange

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set

       Destination        Gateway                     Dist/Metric Last Change
       -----------        -------                     ----------- -----------
  C    2.0.0.0/24         Direct, Vl 192                    0/0    00:20:55
  C    20.0.0.0/24        Direct, Gi 7/1                    0/0    00:10:05
  O    21.0.0.0/24        via 2.0.0.2, Vl 192             110/2    00:10:41

FTOS#show ip route vrf green

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set

       Destination        Gateway                     Dist/Metric Last Change
       -----------        -------                     ----------- -----------
  C    3.0.0.0/24         Direct, Vl 256                    0/0    00:20:52
  C    30.0.0.0/24        Direct, Gi 7/2                    0/0    00:09:45
  S    31.0.0.0/24        via 3.0.0.2, Vl 256               1/0    00:09:06
================================================================================
```

The following shows the output of the **show** commands on Router 2.

ROUTER 2

```
FTOS#show ip vrf
VRF-Name                        VRF-ID Interfaces

default-vrf                     0      Gi 1/0-89,
                                       Te 3/0-3,
                                       Gi 4/0-89,
                                       Gi 5/0-89,
                                       Gi 6/0-89,
                                       Gi 9/0-17,21-47,
                                       Gi 11/0-47,
                                       Gi 12/0-47,
                                       Gi 13/0-47,
                                       Ma 0/0,
                                       Ma 1/0,
                                       Nu 0,
                                       Vl 1
blue                            1      Gi 9/18,
                                       Vl 128
orange                             2      Gi 9/19,
                                       Vl 192
green                           3      Gi 9/20,
                                       Vl 256


FTOS#show ip ospf 1 neighbor
Neighbor ID    Pri      State        Dead Time Address        Interface Area
1.0.0.1         1       FULL/BDR     00:00:36 1.0.0.1         Vl 128    0
!
FTOS#sh ip ospf 2 neighbor
Neighbor ID    Pri      State        Dead Time Address        Interface Area
2.0.0.1         1       FULL/BDR     00:00:33 2.0.0.1         Vl 192    0
!
FTOS#show ip route vrf blue

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set

       Destination        Gateway                     Dist/Metric Last Change
       -----------        -------                     ----------- -----------
  C    1.0.0.0/24         Direct, Vl 128                    0/0    00:27:21
  O    10.0.0.0/24        via 1.0.0.1, Vl 128             110/2    00:14:24
  C    11.0.0.0/24        Direct, Gi 9/18                   0/0    00:19:46
```

------continued next page -------

ROUTER 2 continued

```
FForce10#show ip route vrf orange

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set

       Destination        Gateway                    Dist/Metric Last Change
       -----------        -------                    ----------- -----------
  C    2.0.0.0/24         Direct, Vl 192                   0/0    00:26:44
  O    20.0.0.0/24        via 2.0.0.1, Vl 192            110/2    00:14:22
  C    21.0.0.0/24        Direct, Gi 9/19                  0/0    00:20:38


FTOS#show ip route vrf green

Codes: C - connected, S - static, R - RIP,
       B - BGP, IN - internal BGP, EX - external BGP,LO - Locally Originated,
       O - OSPF, IA - OSPF inter area, N1 - OSPF NSSA external type 1,
       N2 - OSPF NSSA external type 2, E1 - OSPF external type 1,
       E2 - OSPF external type 2, i - IS-IS, L1 - IS-IS level-1,
       L2 - IS-IS level-2, IA - IS-IS inter area, * - candidate default,
       > - non-active route, + - summary route

Gateway of last resort is not set

       Destination        Gateway                    Dist/Metric Last Change
       -----------        -------                    ----------- -----------
  C    3.0.0.0/24         Direct, Vl 256                   0/0    00:26:27
  S    30.0.0.0/24        via 3.0.0.1, Vl 256              1/0    00:17:03
  C    31.0.0.0/24        Direct, Gi 9/20                  0/0    00:20:19
FTOS#
```

58

Virtual Router Redundancy Protocol (VRRP)

IPv4 Virtual Router Redundancy Protocol (VRRP) is available on platforms: C E S

IPv6 VRRP (VRRP version 3) is available on platforms: C E S

This chapter covers the following information:

- VRRP Overview
- VRRP Benefits
- VRRP Implementation
- VRRP Configuration
- Sample Configurations

Virtual Router Redundancy Protocol (VRRP) is designed to eliminate a single point of failure in a statically routed network.

VRRP Overview

VRRP specifies a MASTER router that owns the next hop IP and MAC address for end stations on a LAN. The MASTER router is chosen from the virtual routers by an election process and forwards packets sent to the next hop IP address. If the MASTER router fails, VRRP begins the election process to choose a new MASTER router and that new MASTER continues routing traffic.

VRRP uses the Virtual Router Identifier (VRID) to identify each virtual router configured The IP address of the MASTER router is used as the next hop address for all end stations on the LAN. The other routers represented by IP addresses are BACKUP routers.

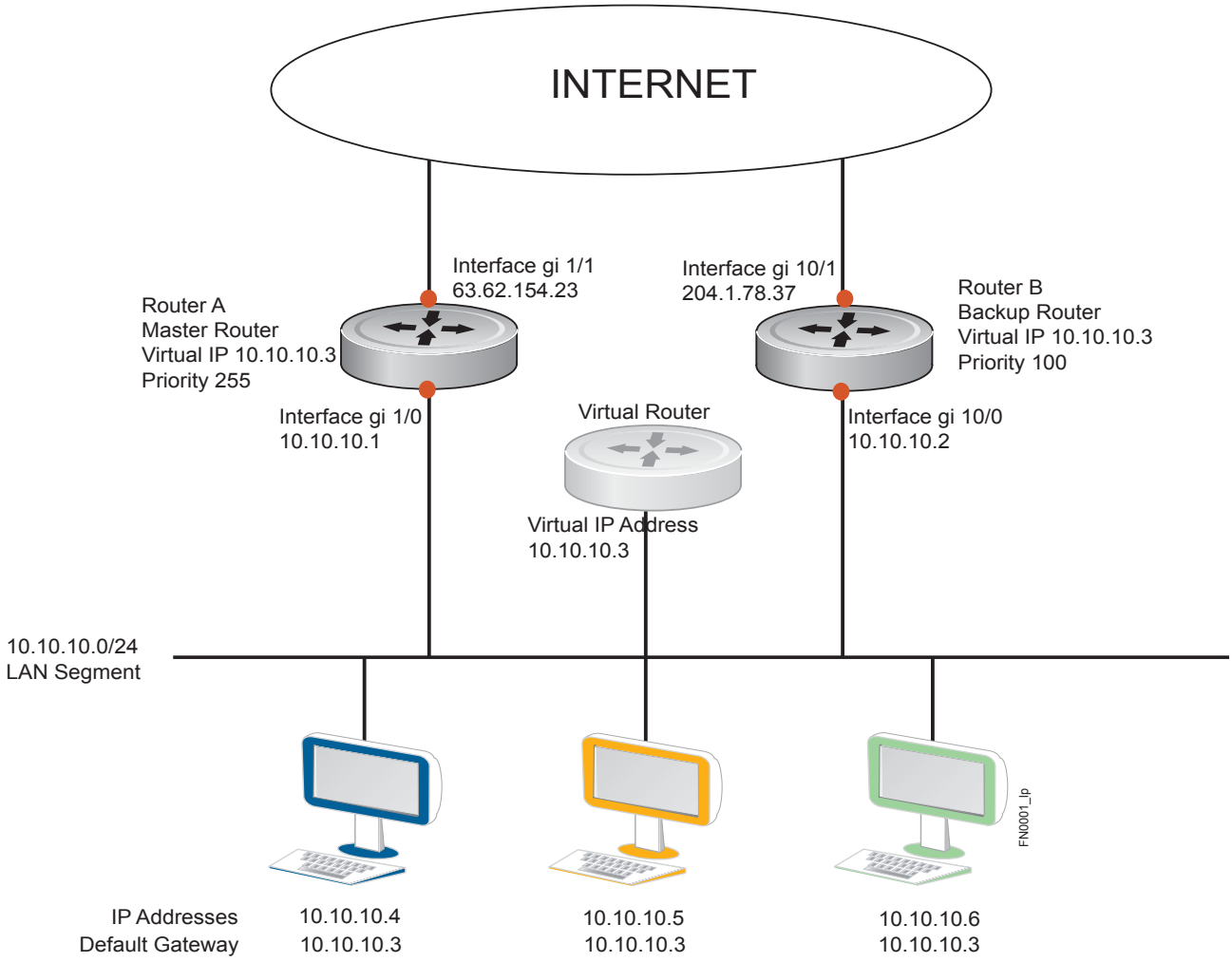
VRRP packets are transmitted with the virtual router MAC address as the source MAC address. The MAC address is in the following format: 00-00-5E-00-01-{VRID}. The first three octets are unchangeable. The next two octets (00-01) indicate the address block assigned to the VRRP protocol, and are unchangeable. The final octet changes depending on the VRRP Virtual Router Identifier and allows for up to 255 VRRP routers on a network.

Figure 58-1 shows a typical network configuration using VRRP. Instead of configuring the hosts on the network 10.10.10.0 with the IP address of either Router A or Router B as their default router; their default router is the IP Address configured on the virtual router. When any host on the LAN segment wants to access the Internet, it sends packets to the IP address of the virtual router.

In Figure 58-1 below, Router A is configured as the MASTER router. It is configured with the IP address of the virtual router and sends any packets addressed to the virtual router through interface GigabitEthernet 1/1 to the Internet. As the BACKUP router, Router B is also configured with the IP address of the virtual router. If for any reason Router A becomes unavailable, VRRP elects a new MASTER Router. Router B assumes the duties of Router A and becomes the MASTER router. At that time, Router B responds to the packets sent to the virtual IP address.

All workstations continue to use the IP address of the virtual router to address packets destined to the Internet. Router B receives and forwards them on interface GigabitEthernet 10/1. Until Router A resumes operation, VRRP allows Router B to provide uninterrupted service to the users on the LAN segment accessing the Internet.

Figure 58-1. Basic VRRP Configuration



For more detailed information on VRRP, refer to RFC 2338, *Virtual Router Redundancy Protocol*.

VRRP Benefits

With VRRP configured on a network, end-station connectivity to the network is not subject to a single point-of-failure. End-station connections to the network are redundant and they are not dependent on IGP protocols to converge or update routing tables.

VRRP Implementation

On E-Series ExaScale and TeraScale routers, VRRP is implemented as follows:

- When VRF microcode is not loaded, VRRP supports an unlimited total number of VRRP groups on a router and up to 255 VRRP groups on an interface (see Table 58-1).
- When VRF microcode is loaded (see Load the VRF CAM Profile on page 1115), VRRP supports an unlimited total number of VRRP groups on a router and up to 15 VRRP groups on an interface (see Table 58-1).

C-Series supports a total of 128 VRRP groups on the switch with varying number of maximum VRRP groups per interface (Table 58-1).

S-Series supports a total of 120 VRRP groups on a switch with FTOS *or* a total of 20 VRRP groups when using SFTOS. The S-Series supports varying number of maximum VRRP groups per interface (Table 58-1).

Within a single VRRP group, up to 12 virtual IP addresses are supported. Virtual IP addresses can belong to the primary or secondary IP address' subnet configured on the interface. You can ping all the virtual IP addresses configured on the Master VRRP router from anywhere in the local subnet.

Though FTOS on E-Series supports unlimited VRRP groups, default VRRP settings may affect the maximum number of groups that can be configured and work efficiently, as a result of hardware throttling VRRP advertisement packets reaching the RP2 processor on the E-Series, the CP on the C-Series, or the FP on the S-Series. To avoid throttling VRRP advertisement packets, Dell Force10 recommends you to increase the VRRP advertisement interval to a value higher than the default value of 1 second. The recommendations are as follows:

Table 58-1. Recommended VRRP Advertise Intervals

	Recommended Advertise Interval			Groups/Interface			
Total VRRP Groups	**E-Series**	**C-Series**	**S-Series**	**E-Series ExaScale**	**E-Series TeraScale**	**C-Series**	**S-Series**
Less than 250	1 second	1 second	1 second	512	255	12	12
Between 250 and 450	2 seconds	2 - 3 seconds	2 - 3 seconds	512	255	24	24
Between 450 and 600	3 seconds	4 seconds	3 - 4 seconds	512	255	36	36
Between 600 and 800	4 seconds	5 seconds	4 seconds	512	255	48	48
Between 800 and 1000	5 seconds	5 seconds	5 seconds	512	255	84	84

Table 58-1. Recommended VRRP Advertise Intervals

	Recommended Advertise Interval			Groups/Interface			
Total VRRP Groups	E-Series	C-Series	S-Series	E-Series ExaScale	E-Series TeraScale	C-Series	S-Series
Between 1000 and 1200	7 seconds	7 seconds	7 seconds	512	255	100	100
Between 1200 and 1500	8 seconds	8 seconds	8 seconds	512	255	120	120

Note: 1500 VRRP groups are supported in FTOS Release 6.3.1.0 and later.

The recommendations in Table 58-1 may vary depending on various factors like ARP broadcasts, IP broadcasts, or STP before changing the advertisement interval. When the number of packets processed by RP2/CP/FP processor increases or decreases based on the dynamics of the network, the advertisement intervals in may increase or decrease accordingly.

CAUTION: Increasing the advertisement interval increases the VRRP Master dead interval, resulting in an increased failover time for Master/Backup election. Take extra caution when increasing the advertisement interval, as the increased dead interval may cause packets to be dropped during that switch-over time.

VRRP version 3

VRRP version 3 defines VRRP for IPv6. The **vrrp-ipv6-group** command is used to create IPv6 VRRP groups, and is similar to the **vrrp-group** command, which creates an IPv4 VRRP group and moves you from INTERFACE mode to a group-specific VRRP command sub-mode.

In the VRRP mode, all VRRP commands are supported for IPv4 and IPv6, except for **authentication-type** which is not supported for IPv6. Also, the following EXEC commands are different for IPv4 and IPv6:

- **clear**:
  - IPv4: **clear counters vrrp**
  - IPv6: **clear counters vrrp ipv6**
- **debug**:
  - IPv4: **debug vrrp**
  - IPv6: **debug vrrp ipv6**
- **show**:
  - IPv4: **show vrrp**
  - IPv6: **show vrrp ipv6**

VRRP Configuration

By default, VRRP is not configured.

The following list specifies the configuration tasks for VRRP:

- Create a Virtual Router on page 1131 (mandatory)
- Assign Virtual IP addresses on page 1132 (mandatory)
- Set VRRP Group (Virtual Router) Priority on page 1135 (optional)
- Configure VRRP Authentication on page 1136 (optional)
- Disable Preempt on page 1137 (optional)
- Change the Advertisement interval on page 1138 (optional)
- Track an Interface or Object on page 1139 (optional)

For a complete listing of all commands related to VRRP, refer to *FTOS Command Line Interface*.

Create a Virtual Router

To enable VRRP, you must create a Virtual Router on a physical or VLAN interface. In FTOS, a VRRP Group is identified by the Virtual Router Identifier (VRID).

Starting in release 8.4.1.0, you can configure a VRRP group on an interface that belongs to a non-default VRF instance.

Prerequisite: The interface on which you create the virtual interface must be enabled and configured with a primary IP address.

To enable a Virtual Router, use the following command in the INTERFACE mode. To delete a VRRP group, use the **no vrrp-group** *vrid* command in the INTERFACE mode.

Task	Command Syntax	Command Mode
Assign an interface (physical or VLAN) to an IPv4 or IPv6 VRRP group.	[**vrrp-group** *vrid* | **vrrp-ipv6-group** *vrid*] VRID range (C-Series and S-Series): 1-255 VRID range (E-Series): 1-255 when VRF microcode is not loaded and 1-15 when VRF microcode is loaded	INTERFACE

Figure 58-2. Command Example: vrrp-group

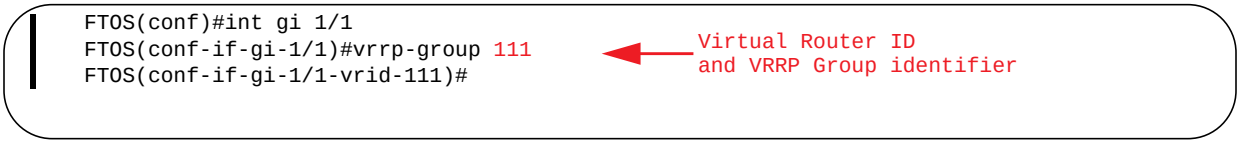


Figure 58-3. Command Example Display: show config for the Interface

```
FTOS(conf-if-gi-1/1)#show conf
!
interface GigabitEthernet 1/1
 ip address 10.10.10.1/24
!
 vrrp-group 111
 no shutdown
FTOS(conf-if-gi-1/1)#
```

Note that the interface has an IP Address and is enabled

Assign Virtual IP addresses

Virtual routers contain virtual IP addresses configured for that VRRP Group (VRID). A VRRP group does not transmit VRRP packets until you assign the Virtual IP address to the VRRP group.

On E-Series ExaScale and TeraScale routers, VRRP is implemented as follows:

- When VRF microcode is not loaded, VRRP supports an unlimited total number of VRRP groups on a router and up to 255 VRRP groups on an interface (see Table 58-1).
- When VRF microcode is loaded (see Load the VRF CAM Profile on page 1115), VRRP supports an unlimited total number of VRRP groups on a router and up to 15 VRRP groups on an interface (see Table 58-1).

C-Series supports a total of 128 VRRP groups on the switch with varying number of maximum VRRP groups per interface (Table 58-1).

S-Series supports a total of 120 VRRP groups on a switch with FTOS *or* a total of 20 VRRP groups when using SFTOS. The S-Series supports varying number of maximum VRRP groups per interface (Table 58-1).

To activate a VRRP Group on an interface (so that VRRP group starts transmitting VRRP packets), configure at least one Virtual IP address in a VRRP group. The Virtual IP address is the IP address of the Virtual Router and does not require the IP address mask.

You can configure up to 12 virtual IP addresses for a VRRP group (VRID). The following configuration rules apply to a virtual IP address:

- A virtual IP address must be in the same subnet as the primary or secondary IP address configured on the interface. Although a single VRRP group can contain virtual IP addresses belonging to multiple IP subnets configured on the interface, Dell Force10 recommends that you configure virtual IP addresses belonging to the *same* IP subnet for any one VRRP group.

For example, an interface (on which VRRP is to be enabled) contains a primary IP address of 50.1.1.1/24 and a secondary IP address of 60.1.1.1/24. The VRRP Group (VRID 1) must contain virtual addresses belonging to *either* subnet 50.1.1.0/24 or subnet 60.1.1.0/24, but not from both subnets (though FTOS allows the same).

- If the virtual IP address and the interface's primary/secondary IP address are the same, the priority on that VRRP group is automatically set to 255. The interface then becomes the MASTER/OWNER router of the VRRP group and the interface's physical MAC address is changed to that of the owner VRRP group's MAC address. (You can also configure a priority for the group even if the group is owned. The configured priority is saved but only applied as the run-time priority when the last virtual address is removed from the group.)
- If multiple VRRP groups are configured on an interface, only one of the VRRP groups can have the primary or secondary IP address of the interface.

Configure a Virtual IP address with these commands in the following sequence in the INTERFACE mode.

Step	Task	Command Syntax	Command Mode
1	Configure an IPv4 or IPv6 VRRP group.	**vrrp-group** *vrid* | **vrrp-ipv6-group** *vrid* VRID range (C-Series and S-Series): 1-255 VRID range (E-Series): 1-255 when VRF microcode is not loaded and 1-15 when VRF microcode is loaded	INTERFACE
2	Configure virtual IP addresses for this VRID.	**virtual-address** *ip-address1* [...*ip-address12*] Range: Up to 12 virtual IP addresses	INTERFACE -VRID

Note: After you enter the **vrrp-group or vrrp-ipv6-group** command, a message similar to the following is displayed to confirm the VRID number used with the VRRP group and displayed in **show vrrp** command output: `The VRID used by the VRRP group is 41.`
For information on how the VRID number changes when VRF microcode is loaded, see the Note in VRRP on a VRF Interface on page 1142.

Figure 58-4. Command Example: virtual-address

```
FTOS(conf-if-gi-1/1-vrid-111)#virtual-address 10.10.10.1
FTOS(conf-if-gi-1/1-vrid-111)#virtual-address 10.10.10.2
FTOS(conf-if-gi-1/1-vrid-111)#virtual-address 10.10.10.3
```

Figure 58-5. Command Example Display: show config for the Interface

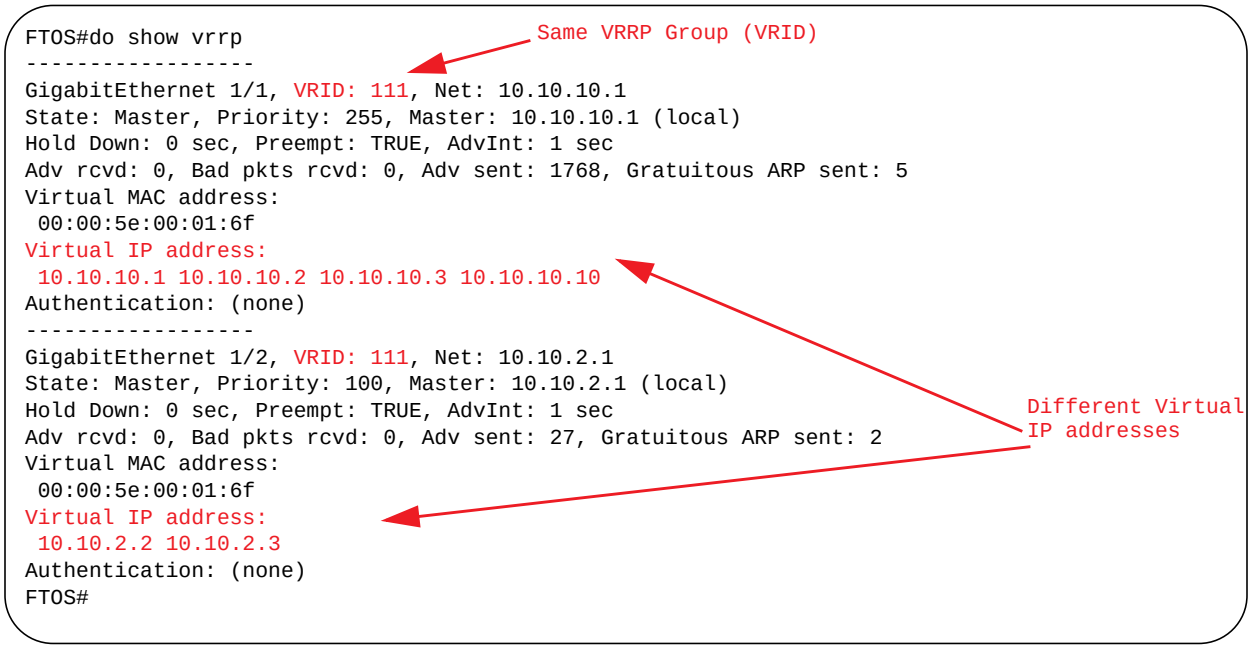
```
FTOS(conf-if-gi-1/1)#show conf
!
interface GigabitEthernet 1/1
 ip address 10.10.10.1/24
!
 vrrp-group 111
  priority 255
  virtual-address 10.10.10.1
  virtual-address 10.10.10.2
  virtual-address 10.10.10.3
!
 vrrp-group 222
 no shutdown
```

Note that the Primary IP address and the Virtual IP addresses are on the same subnet

Figure 58-6 shows the same VRRP group configured on multiple interfaces on different subnets.

Note: **show vrrp** displays all of the active IPv4 groups, and **show ipv6 vrrp** displays all of the active IPv6 groups.

Figure 58-6. Command Example Display: show vrrp




```
FTOS#do show vrrp
------------------
GigabitEthernet 1/1, VRID: 111, Net: 10.10.10.1
State: Master, Priority: 255, Master: 10.10.10.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 1768, Gratuitous ARP sent: 5
Virtual MAC address:
 00:00:5e:00:01:6f
Virtual IP address:
 10.10.10.1 10.10.10.2 10.10.10.3 10.10.10.10
Authentication: (none)
------------------
GigabitEthernet 1/2, VRID: 111, Net: 10.10.2.1
State: Master, Priority: 100, Master: 10.10.2.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 27, Gratuitous ARP sent: 2
Virtual MAC address:
 00:00:5e:00:01:6f
Virtual IP address:
 10.10.2.2 10.10.2.3
Authentication: (none)
FTOS#
```

When the VRRP process completes its initialization, the State field contains either Master or Backup.

Set VRRP Group (Virtual Router) Priority

Setting a Virtual Router priority to 255 ensures that router is the “owner” virtual router for the VRRP group. VRRP elects the MASTER router by choosing the router with the highest priority. The default priority for a Virtual Router is 100. The higher the number, the higher the priority. If the MASTER router fails, VRRP begins the election process to choose a new MASTER router based on the next-highest priority.

If two routers in a VRRP group come up at the same time and have the same priority value, the interface’s physical IP addresses are used as tie-breakers to decide which is MASTER. The router with the higher IP address will become MASTER.

Configure the VRRP Group’s priority with the following command in the VRRP mode:

Task	Command Syntax	Command Mode
Configure the priority for the VRRP group.	INTERFACE -VRID	**priority** *priority* Range: 1-255 Default: 100

Figure 58-7. Command Example: priority in Interface VRRP mode

```
FTOS(conf-if-gi-1/2)#vrrp-group 111
FTOS(conf-if-gi-1/2-vrid-111)#priority 125
```

Figure 58-8. Command Example Display: show vrrp

```
FTOS#show vrrp
------------------
GigabitEthernet 1/1, VRID: 111, Net: 10.10.10.1
State: Master, Priority: 255, Master: 10.10.10.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 2343, Gratuitous ARP sent: 5
Virtual MAC address:
 00:00:5e:00:01:6f
Virtual IP address:
 10.10.10.1 10.10.10.2 10.10.10.3 10.10.10.10
Authentication: (none)
------------------
GigabitEthernet 1/2, VRID: 111, Net: 10.10.2.1
State: Master, Priority: 125, Master: 10.10.2.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 601, Gratuitous ARP sent: 2
Virtual MAC address:
 00:00:5e:00:01:6f
Virtual IP address:
 10.10.2.2 10.10.2.3
Authentication: (none)
FTOS(conf)#
```

Configure VRRP Authentication

Note: Authentication is not available for IPv6 VRRP.

Simple authentication of VRRP packets ensures that only trusted routers participate in VRRP processes. When authentication is enabled, FTOS includes the password in its VRRP transmission, and the receiving router uses that password to verify the transmission.

Note: All virtual routers in the VRRP group must be configured the same: authentication must be enabled with the same password or authentication is disabled.

Configure simple authentication with the following command in VRRP configuration mode:

Task	Command Syntax	Command Mode
Configure a simple text password.	**authentication-type simple** [*encryption-type*] *password* *encryption-type:* 0 indicates an unencrypted password in the configuration; 7 indicates an encrypted password in the configuration. *password: plain text*	INTERFACE-VRID

Note: As shown in Figure 58-9, the VRRP authentication password that you configure is displayed in encrypted form in **show running-config** (EXEC Privilege) and **show config** (INTERFACE) command output. To display the VRRP authentication password (as well as all other FTOS passwords) in clear text in **show** command output, you must enter the **no service password-encryption** (CONFIGURATION) command. To remove the currently configured VRRP authentication password, enter the **no authentication-type simple** [*encryption-type*] *password* command.

Figure 58-9. Command Example: authentication-type simple

```
FTOS(conf-if-gi-1/1-vrid-111)#authentication-type simple 0 force10
FTOS(conf-if-gi-1/1-vrid-111)#show config
!
 vrrp-group 111
  authentication-type simple 7 387a7f2df5969da4
  priority 255
  virtual-address 10.10.10.1
  virtual-address 10.10.10.2
  virtual-address 10.10.10.3
  virtual-address 10.10.10.10
FTOS(conf-if-gi-1/1-vrid-111)#no authentication-type simple 0 force10
FTOS(conf-if-gi-1/1-vrid-111)#show config
!
 vrrp-group 111
  priority 255
  virtual-address 10.10.10.1
  virtual-address 10.10.10.2
  virtual-address 10.10.10.3
  virtual-address 10.10.10.10
```

Disable Preempt

The **preempt** command is enabled by default, and it forces the system to change the MASTER router if another router with a higher priority comes online.

Prevent the BACKUP router with the higher priority from becoming the MASTER router by disabling **preempt**.

Note: All virtual routers in the VRRP group must be configured the same: all configured with preempt enabled or configured with preempt disabled.

Since preempt is enabled by default, disable the preempt function with the following command in the VRRP mode. Re-enable preempt by entering the **preempt** command. When preempt is enabled, it does not display in the show commands, because it is a default setting.,

Task	Command Syntax	Command Mode
Prevent any BACKUP router with a higher priority from becoming the MASTER router.	**no preempt**	INTERFACE-VRID

Figure 58-10. Command Example: no preempt

```
FTOS(conf-if-gi-1/1)#vrrp-group 111
FTOS(conf-if-gi-1/1-vrid-111)#no preempt
FTOS(conf-if-gi-1/1-vrid-111)#show conf
```

Figure 58-11. Command Example Display: show config in VRID mode

```
FTOS(conf-if-gi-1/1-vrid-111)#show conf
!
 vrrp-group 111
  authentication-type simple 7 387a7f2df5969da4
  no preempt
  priority 255
  virtual-address 10.10.10.1
  virtual-address 10.10.10.2
  virtual-address 10.10.10.3
  virtual-address 10.10.10.10
FTOS(conf-if-gi-1/1-vrid-111)#
```

Change the Advertisement interval

By default, the MASTER router transmits a VRRP advertisement to all members of the VRRP group every 1 second, indicating it is operational and is the MASTER router. If the VRRP group misses 3 consecutive advertisements, then the election process begins and the BACKUP virtual router with the highest priority transitions to MASTER.

Note: Dell Force10 recommends you to increase the VRRP advertisement interval to a value higher than the default value of 1 second to avoid throttling VRRP advertisement packets. If you do change the time interval between VRRP advertisements on one router, you must change it on all participating routers.

Change that advertisement interval with the following command in the VRRP mode:

Task	Command Syntax	Command Mode
Change the advertisement interval setting.	**advertise-interval** *seconds* Range: 1-255 seconds IPv4 Default: 1 second IPv6 Default: 100 centiseconds	INTERFACE-VRID

Figure 58-12. Command Example: advertise-interval

```
FTOS(conf-if-gi-1/1)#vrrp-group 111
FTOS(conf-if-gi-1/1-vrid-111)#advertise-interval 10
FTOS(conf-if-gi-1/1-vrid-111)#
```

Figure 58-13. Command Example Display: advertise-interval in VRID mode

```
FTOS(conf-if-gi-1/1-vrid-111)#show conf
!
 vrrp-group 111
  advertise-interval 10
  authentication-type simple 7 387a7f2df5969da4
  no preempt
  priority 255
  virtual-address 10.10.10.1
  virtual-address 10.10.10.2
  virtual-address 10.10.10.3
  virtual-address 10.10.10.10
FTOS(conf-if-gi-1/1-vrid-111)#
```

Track an Interface or Object

In previous releases, you could set FTOS to track the state of an interface for a specified virtual group. Starting in release 8.4.1.0, you can track additional objects for a virtual group, such as Layer 3 interfaces (IPv4 and IPv6), IPv4/IPv6 route reachability, and thresholds of IPv4/IPv6 route metrics. For information on how to track supported objects, refer to Chapter 31, “Object Tracking,” on page 677.

Each VRRP group can track changes in the status of up to 12 interfaces and up to 20 additional objects, which may affect the priority of the VRRP group. If a tracked interface or object goes down, the VRRP group’s priority is decreased by a default value of 10 (also known as cost). If the state of a tracked interface or object goes up, the VRRP group’s priority is increased by 10.

The lowered priority of a VRRP group may trigger an election. Because Master/Backup VRRP routers are selected based on the VRRP group’s priority, tracking interfaces and/or objects ensures that the best VRRP router is the Master for a group. In object and interface tracking, the following conditions apply:

- The sum of the costs of all tracked interfaces and objects cannot equal or exceed the priority of the VRRP group.
- If the VRRP group is configured as the Owner router (priority 255), tracking for the group is disabled, irrespective of the state of the tracked interfaces and objects. The priority of the owner group always remains as 255 and does not change.

For a virtual group, you can track the line-protocol state or the routing status of any of the following interfaces with the **interface** *interface* parameter:

- 1-Gigabit Ethernet: Enter **gigabitethernet** *slot/port* in the **track** *interface* command (see Step 1 below).
- 10-Gigabit Ethernet: Enter **tengigabitethernet** *slot/port*.
- Port channel: Enter **port-channel** *number*, where valid port-channel numbers are:
  - For the C-Series and S-Series, 1 to 128
  - For the E-Series: 1 to 32 for EtherScale, 1 to 255 for TeraScale, and 1 to 512 for ExaScale
- SONET: Enter **sonet** *slot/port*.
- VLAN: Enter **vlan** *vlan-id*, where valid VLAN IDs are from 1 to 4094.

For a virtual group, you can also track the status of a configured object (**track** *object-id* command) by entering its object number. See Object Tracking Configuration on page 681 for more information.

Note that you can configure a tracked object for a VRRP group (using the **track** *object-id* command in INTERFACE-VRID mode) before you actually create the tracked object (using a **track** *object-id* command in CONFIGURATION mode). However, no changes in the VRRP group’s priority will occur until the tracked object is defined and determined to be down.

In addition, if you configure a VRRP group on an interface that belongs to a VRF instance and later configure object tracking on an interface for the VRRP group, the tracked interface must belong to the VRF instance.

To track an interface or configured object for a virtual group, use the **track** command in the VRRP mode:

Task	Command Syntax	Command Mode
Monitor an interface or a configured object and, optionally, reconfigure the cost value to be subtracted from the VRRP group priority if the status of the tracked object goes DOWN.	**track** {*interface* | *object-id*} [**priority-cost** *cost*] Valid object IDs are from 1 to 65535. Cost range: 1-254. Default: 10	INTERFACE-VRID
(Optional) Display the configuration and UP or DOWN state of tracked objects, including the client (VRRP group) that is tracking an object's state.	**show track**	EXEC EXEC Privilege
(Optional) Display the configuration and UP or DOWN state of tracked interfaces and objects in VRRP groups, including the time since the last change in an object's state.	**show vrrp**	EXEC EXEC Privilege
(Optional) Display the configuration of tracked objects in VRRP groups on a specified interface.	**show running-config interface** *interface*	EXEC EXEC Privilege

Note: The sum of all the costs for all tracked interfaces and objects must be less than or equal to the configured priority of the VRRP group.

Figure 58-14. Command Example: track *interface*

```
FTOS(conf-if-gi-1/1)#vrrp-group 111
FTOS(conf-if-gi-1/1-vrid-111)#track gigabitethernet 1/2
FTOS(conf-if-gi-1/1-vrid-111)#
```

Figure 58-15. Command Example: show configuration in VRID mode

```
FTOS(conf-if-gi-1/1-vrid-111)#show configuration
!
 vrrp-group 111
  advertise-interval 10
  authentication-type simple 7 387a7f2df5969da4
  no preempt
  priority 255
  track GigabitEthernet 1/2
  virtual-address 10.10.10.1
  virtual-address 10.10.10.2
  virtual-address 10.10.10.3
  virtual-address 10.10.10.10
FTOS(conf-if-gi-1/1-vrid-111)#
```

Figure 58-16. Command Example: show track

```
FTOS#show track

Track 2
  IPv6 route 2040::/64 metric threshold
  Metric threshold is Up (STATIC/0/0)
   5 changes, last change 00:02:16
  Metric threshold down 255 up 254
  First-hop interface is GigabitEthernet 13/2
  Tracked by:
    VRRP GigabitEthernet 7/30 IPv6 VRID 1

Track 3
  IPv6 route 2050::/64 reachability
  Reachability is Up (STATIC)
   5 changes, last change 00:02:16
  First-hop interface is GigabitEthernet 13/2
  Tracked by:
    VRRP GigabitEthernet 7/30 IPv6 VRID 1
```

Figure 58-17. Command Example: show vrrp

```
FTOS#show vrrp
------------------

GigabitEthernet 7/30, IPv6 VRID: 1, Version: 3, Net: fe80::201:e8ff:fe01:95cc
VRF: 0 default-vrf
State: Master, Priority: 100, Master: fe80::201:e8ff:fe01:95cc (local)
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 310
Virtual MAC address:
 00:00:5e:00:02:01
Virtual IP address:
 2007::1 fe80::1
Tracking states for 2 resource Ids:
2 - Up IPv6 route, 2040::/64, priority-cost 20, 00:02:11
3 - Up IPv6 route, 2050::/64, priority-cost 30, 00:02:11
```

Figure 58-18. Command Example: show running-config interface

```
FTOS#show running-config interface gigabitethernet 7/30

interface GigabitEthernet 7/30
 no ip address
 ipv6 address 2007::30/64

 vrrp-ipv6-group 1
  track 2 priority-cost 20
  track 3 priority-cost 30
  virtual-address 2007::1
  virtual-address fe80::1
 no shutdown
```

VRRP on a VRF Interface

VRRP is supported with Virtual Routing and Forwarding (VRF) only on platform: E

Starting in release 8.4.1.0, you can configure the VRRP feature on interfaces that belong to a non-default Virtual Routing and Forwarding (VRF) instance on E-Series routers. In previous releases, the VRRP feature was not supported on interfaces that were configured for VRF. For a sample VRRP configuration on a VRF interface, see VRRP in VRF Configuration on page 1149.

The VRF feature allows a physical router to partition itself into multiple virtual routers (VRs) so that multiple instances of a routing table can co-exist within the same router at the same time. The control and data plane are isolated in each VR so that traffic does not flow across VRs. For more information, refer to Chapter 57, “Virtual Routing and Forwarding (VRF),” on page 1109.

In a virtualized network that consists of multiple VRFs, various overlay networks can exist on a shared physical infrastructure:

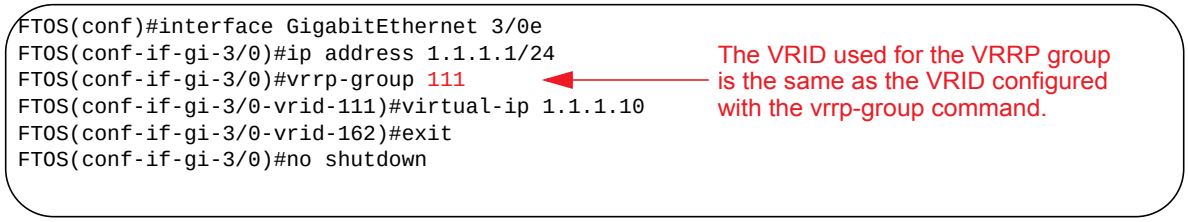
- The same IP addresses or overlapping IP subnets can exist in different VRFs. (If two interfaces are assigned to the same VRF, you cannot configure overlapping IP subnets or the same IP address on them.)
- The same VRRP virtual address can exist in different VRFs.

Nodes (hosts and servers) that are part of the VRFs can be configured with IP static routes for reaching specific destinations through a given gateway in a VRF. VRRP can provide high availability and protection for next-hop static routes by eliminating a single point of failure in the default static routed network.

Note: On E-Series routers, the VRID used by the VRRP protocol changes according to whether VRF microcode is loaded or not:

• When VRF microcode is not loaded in CAM, the VRID for a VRRP group is the same as the VRID number configured with the **vrrp-group** or **vrrp-ipv6-group** command:

Figure 58-19. VRID used when VRF microcode is not loaded

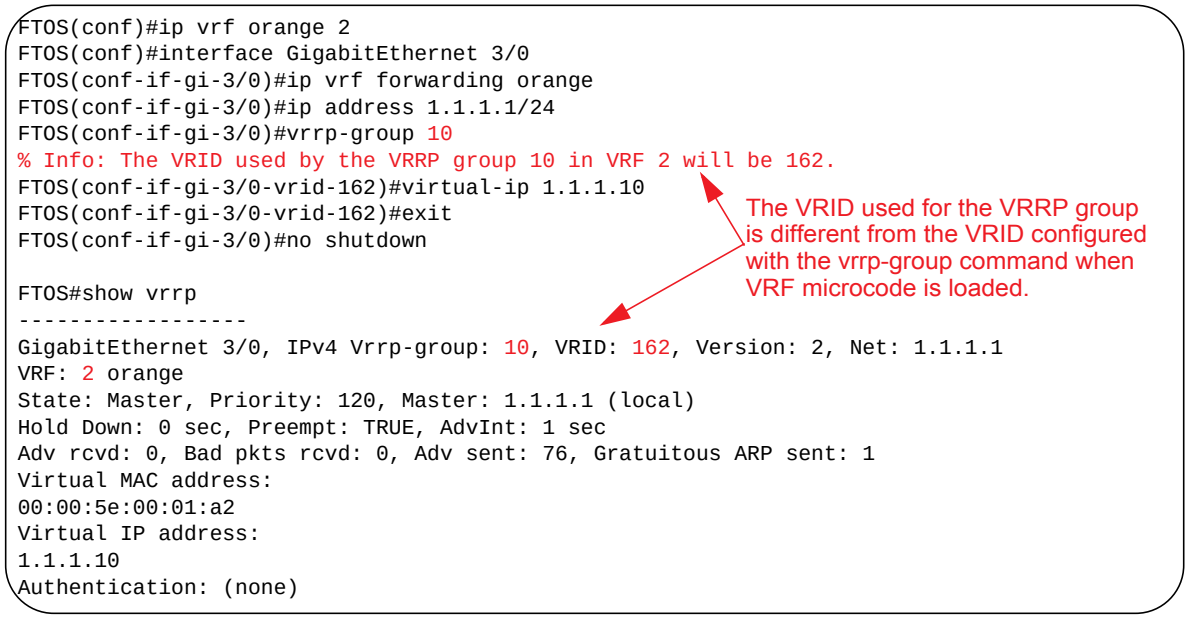


• When VRF microcode is loaded in CAM, the VRID for a VRRP group is equal to 16 times the **vrrp-group vrid** or **vrrp-ipv6-group vrid** number plus the **ip vrf vrf-id** number.

For example, if VRF microcode is loaded and VRRP group 10 is configured in VRF 2, the VRID used for the VRRP group is (16 x 10) + 2, or 162. This VRID value is used in the lowest byte of the virtual MAC address of the VRRP group and is also used for VRF routing.

Note that the actual VRID used by a VRRP group is displayed below the command line when you enter the **vrrp-group** or **vrrp-ipv6-group** command in VRRP-group configuration mode, and in **show vrrp** command output:

Figure 58-20. VRID used when VRF microcode is loaded



Important: You must configure the same VRID on neighboring routers (Dell Force10 or non-Force10) in the same VRRP group in order for all routers to interoperate.

Sample Configurations

VRRP for IPv4 Configuration

The configuration in Figure 58-21 shows how to enable IPv4 VRRP. This example does not contain comprehensive directions and is intended to provide guidance for only a typical VRRP configuration. You can copy and paste from the example to your CLI. Be sure you make the necessary changes to support your own IP addresses, interfaces, names, etc. Figure 58-21 shows the VRRP topology created with the CLI configuration in Figure 58-22.

Figure 58-21. VRRP for IPv4 Topology

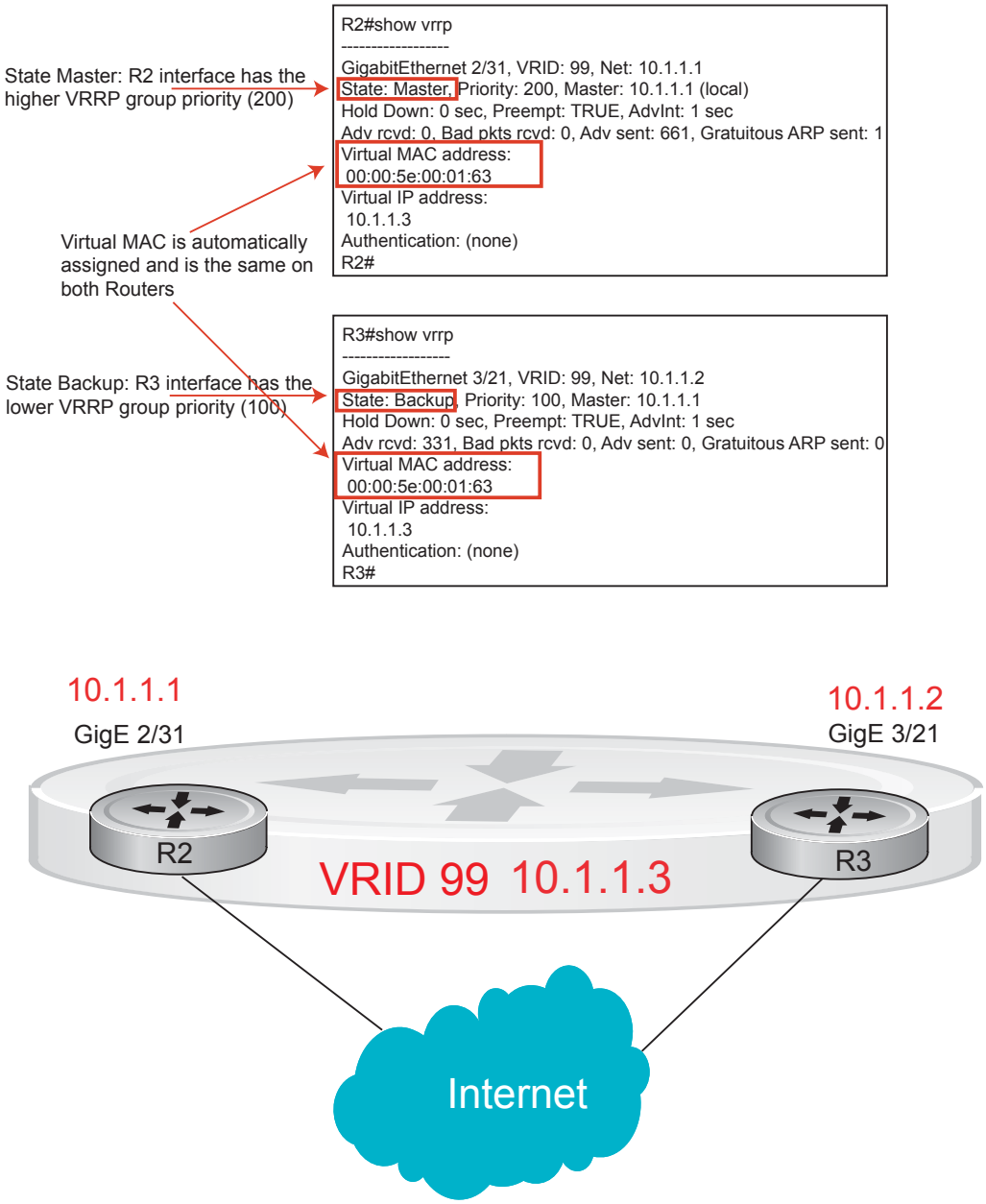


Figure 58-22. Configure VRRP for IPv4

Router 2

```
R2(conf)#int gi 2/31
R2(conf-if-gi-2/31)#ip address 10.1.1.1/24
R2(conf-if-gi-2/31)#vrrp-group 99
R2(conf-if-gi-2/31-vrid-99)#priority 200
R2(conf-if-gi-2/31-vrid-99)#virtual 10.1.1.3
R2(conf-if-gi-2/31-vrid-99)#no shut
R2(conf-if-gi-2/31)#show conf
!
interface GigabitEthernet 2/31
 ip address 10.1.1.1/24
!
 vrrp-group 99
  priority 200
  virtual-address 10.1.1.3
 no shutdown
R2(conf-if-gi-2/31)#end

R2#show vrrp
------------------
GigabitEthernet 2/31, VRID: 99, Net: 10.1.1.1
State: Master, Priority: 200, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 817, Gratuitous ARP sent: 1
Virtual MAC address:
 00:00:5e:00:01:63
Virtual IP address:
 10.1.1.3
Authentication: (none)
R2#
```

Router 3

```
R3(conf)#int gi 3/21
R3(conf-if-gi-3/21)#ip address 10.1.1.2/24
R3(conf-if-gi-3/21)#vrrp-group 99
R3(conf-if-gi-3/21-vrid-99)#virtual 10.1.1.3
R3(conf-if-gi-3/21-vrid-99)#no shut
R3(conf-if-gi-3/21)#show conf
!
interface GigabitEthernet 3/21
 ip address 10.1.1.1/24
!
 vrrp-group 99
  virtual-address 10.1.1.3
 no shutdown
R3(conf-if-gi-3/21)#end
R3#show vrrp
------------------
GigabitEthernet 3/21, VRID: 99, Net: 10.1.1.2
State: Backup, Priority: 100, Master: 10.1.1.1
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 698, Bad pkts rcvd: 0, Adv sent: 0, Gratuitous ARP sent: 0
Virtual MAC address:
 00:00:5e:00:01:63
Virtual IP address:
 10.1.1.3
Authentication: (none)
```

VRRP for IPv6 Configuration

Figure 58-22 shows an example of a VRRP for IPv6 configuration in which the IPv6 VRRP group consists of two routers. This example does not contain comprehensive directions and is intended to provide guidance for only a typical VRRP configuration. You can copy and paste from the example to your CLI. Be sure you make the necessary changes to support your own IP addresses, interfaces, names, etc. Figure 58-21 shows the VRRP for IPv6 topology with the CLI configuration in Figure 58-22.

Figure 58-23. VRRP for IPv6 Topology

Master State: Although both R2 and R3 have the same priority (100), R2 is the master in the VRRP group because the R2 interface has a higher IPv6 address.

Virtual MAC is automatically assigned and is the same on both Routers

You must configure both a virtual IPv6 address and a virtual link local (fe80) address for an IPv6 VRRP group

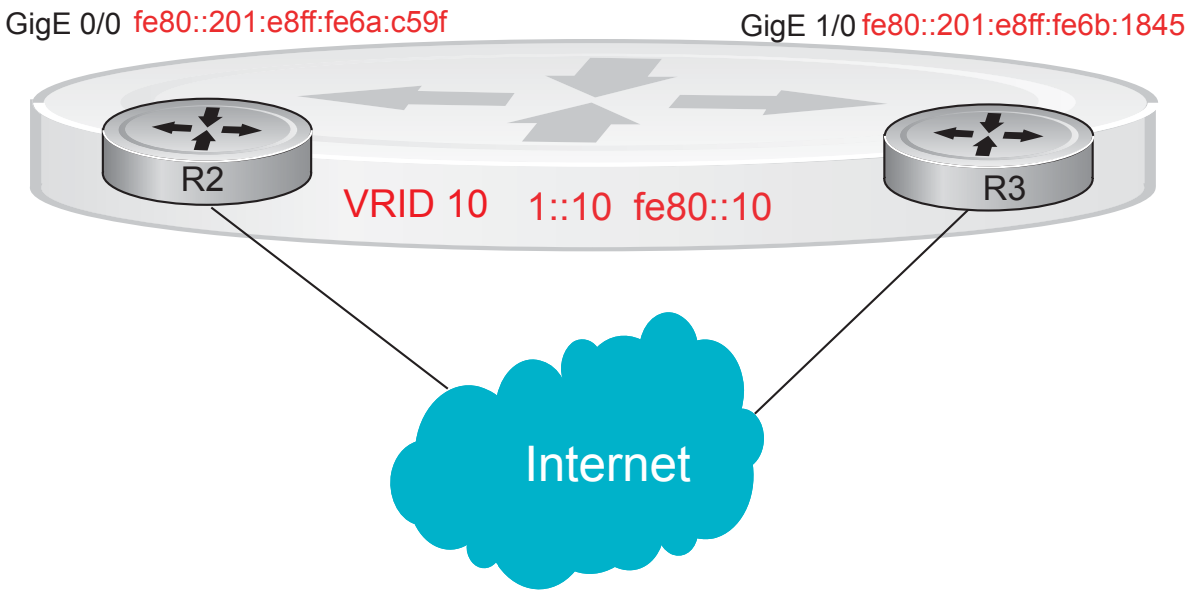
```
R2#show vrrp
------------------
GigabitEthernet 0/0, IPv6 VRID: 10, Net: fe80::201:e8ff:fe6a:c59f
VRF: 0 default-vrf
State: Master, Priority: 100, Master: fe80::201:e8ff:fe6a:c59f (local)
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 135
Virtual MAC address:
 00:00:5e:00:02:0a
Virtual IP address:
1::10  fe80::10
R2#
```

Backup State: R3 is the backup in the VRRP group because the R3 interface has a lower IPv6 address.

Virtual MAC is automatically assigned and is the same on both Routers

You must configure both a virtual IPv6 address and a virtual link local (fe80) address for an IPv6 VRRP group

```
R3#show vrrp
------------------
GigabitEthernet 1/0, IPv6 VRID: 10, Net: fe80::201:e8ff:fe6b:1845
VRF: 0 default-vrf
State: Backup Priority: 100, Master: fe80::201:e8ff:fe6a:c59f
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 135
Virtual MAC address:
 00:00:5e:00:02:0a
Virtual IP address:
1::10  fe80::10
R2#
```



Note: In a VRRP or VRRPv3 group, if two routers come up with the same priority and another router already has MASTER status, the router with master status continues to be master even if one of two routers has a higher IP or IPv6 address.

Figure 58-24. Configure VRRP for IPv6

Router 2

```
R2(conf)#interface gigabitethernet 0/0
R2(conf-if-gi-0/0)#no ip address
R2(conf-if-gi-0/0)#ipv6 address 1::1/64
R2(conf-if-gi-0/0)#vrrp-group 10
R2(conf-if-gi-0/0-vrid-10)#virtual-address fe80::10
R2(conf-if-gi-0/0-vrid-10)#virtual-address 1::10
R2(conf-if-gi-0/0-vrid-10)#no shutdown
R2(conf-if-gi-0/0)#show config
interface GigabitEthernet 0/0
 ipv6 address 1::1/64
 vrrp-group 10
  priority 100
  virtual-address fe80::10
  virtual-address 1::10
 no shutdown
R2(conf-if-gi-0/0)#end

R2#show vrrp
------------------
GigabitEthernet 0/0, IPv6 VRID: 10, Version: 3, Net:fe80::201:e8ff:fe6a:c59f
State: Master, Priority: 100, Master: fe80::201:e8ff:fe6a:c59f (local)
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 135
Virtual MAC address:
00:00:5e:00:02:0a
Virtual IP address:
1::10 fe80::10
```

You must configure a virtual link local (fe80) address for each VRRPv3 group created for an interface. The VRRPv3 group becomes active as soon as you configure the link local address. Afterwards, you can configure the group's virtual IPv6 address.

The virtual IPv6 address you configure should be the same as the IPv6 subnet to which the interface belongs.

Router 3

```
R3(conf)#interface gigabitethernet 1/0
R3(conf-if-gi-1/0)#no ipv6 address
R3(conf-if-gi-1/0)#ipv6 address 1::2/64
R3(conf-if-gi-1/0)#vrrp-group 10
R2(conf-if-gi-1/0-vrid-10)#virtual-address fe80::10
R2(conf-if-gi-1/0-vrid-10)#virtual-address 1::10
R3(conf-if-gi-1/0-vrid-10)#no shutdown
R3(conf-if-gi-1/0)#show config
interface GigabitEthernet 1/0
 ipv6 address 1::2/64
 vrrp-group 10
  priority 100
  virtual-address fe80::10
  virtual-address 1::10
 no shutdown
R3(conf-if-gi-1/0)#end

R3#show vrrp
------------------
GigabitEthernet 1/0, IPv6 VRID: 10, Version: 3, Net:
fe80::201:e8ff:fe60:1845
State: Backup, Priority: 100, Master: fe80::201:e8ff:fe6a:c59f
Hold Down: 0 centisec, Preempt: TRUE, AdvInt: 100 centisec
Accept Mode: FALSE, Master AdvInt: 100 centisec
Adv rcvd: 11, Bad pkts rcvd: 0, Adv sent: 0
Virtual MAC address:
00:00:5e:00:02:0a
Virtual IP address:
1::10 fe80::10
```

Although R2 and R3 have the same default priority (100), R2 is elected master in the VRRPv3 group because the GigE 0/0 interface has a higher IPv6 address than the GigE 1/0 interface on R3.

VRRP in VRF Configuration

The example in this section shows how to enable VRRP operation in a VRF virtualized network for the following scenarios:

- Multiple VRFs on physical interfaces running VRRP
- Multiple VRFs on VLAN interfaces running VRRP

To view a VRRP in VRF configuration, use the **show** commands described in Displaying a VRRP in VRF Configuration on page 1154.

Non-VLAN Scenario

Figure 58-25. VRRP in VRF: Non-VLAN Example

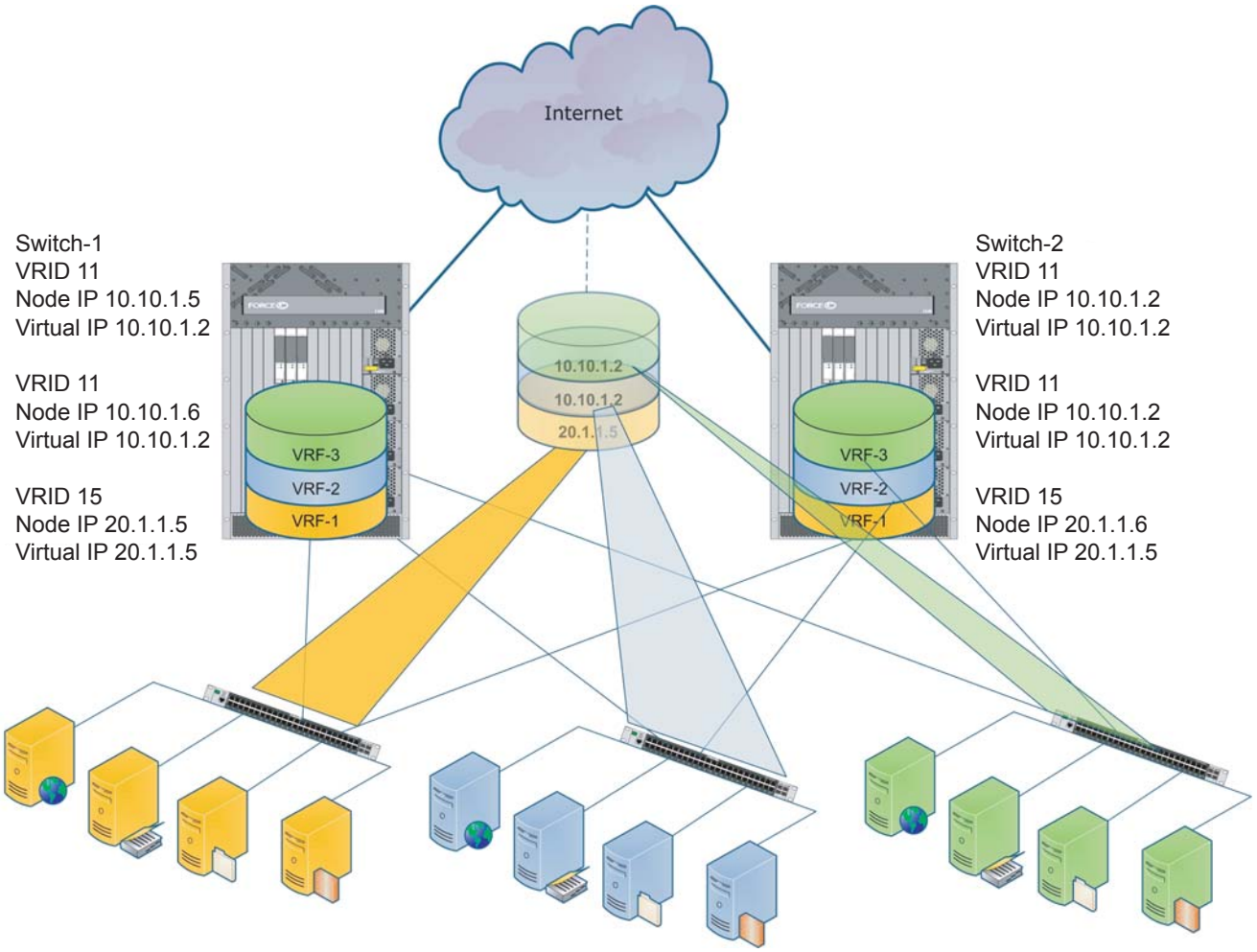


Figure 58-25 shows a typical use case in which three virtualized overlay networks are created by configuring three VRFs in two E-Series switches. The default gateway to reach the internet in each VRF is a static route with the next hop being the virtual IP address configured in VRRP. In this scenario, a single VLAN is associated with each VRF.

Both Switch-1 and Switch-2 have three VRF instances defined: VRF-1, VRF-2, and VRF-3. Each VRF has a separate physical interface to a LAN switch and an upstream VPN interface to connect to the Internet. Both Switch-1 and Switch-2 use VRRP groups on each VRF instance in order that there is one master and one backup router for each VRF. In VRF-1 and VRF-2, Switch-2 serves as owner-master of the VRRP group and Switch-1 serves as the backup. On VRF-3, Switch-1 is the owner-master and Switch-2 is the backup.

Note that in VRF-1 and VRF-2 on Switch-2, the virtual IP and node IP address, subnet, and VRRP group are the same. On Switch-1, the virtual IP address, subnet, and VRRP group are the same in VRF-1 and VRF-2, but the IP address of the node interface is unique. There is no requirement for the virtual IP and node IP addresses to be the same in VRF-1 and VRF-2; similarly, there is no requirement for the IP addresses to be different. In VRF-3, the node IP addresses and subnet are unique.

Figure 58-26. VRRP in VRF: Switch-1 Non-VLAN Configuration

Switch-1

```
S1(conf)#ip vrf default-vrf 0
!
S1(conf)#ip vrf VRF-1 1
!
S1(conf)#ip vrf VRF-2 2
!
S1(conf)#ip vrf VRF-3 3
!
S1(conf)#interface GigabitEthernet 12/1
S1(conf-if-gi-12/1)#ip vrf forwarding VRF-1
S1(conf-if-gi-12/1)#ip address 10.10.1.5/24
S1(conf-if-gi-12/1)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 1 will be 177.
S1(conf-if-gi-12/1-vrid-101)#priority 100
S1(conf-if-gi-12/1-vrid-101)#virtual-address 10.10.1.2
S1(conf-if-gi-12/1)#no shutdown
!
S1(conf)#interface GigabitEthernet 12/2
S1(conf-if-gi-12/2)#ip vrf forwarding VRF-2
S1(conf-if-gi-12/2)#ip address 10.10.1.6/24
S1(conf-if-gi-12/2)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 2 will be 178.
S1(conf-if-gi-12/2-vrid-101)#priority 100
S1(conf-if-gi-12/2-vrid-101)#virtual-address 10.10.1.2
S1(conf-if-gi-12/2)#no shutdown
!
S1(conf)#interface GigabitEthernet 12/3
S1(conf-if-gi-12/3)#ip vrf forwarding VRF-3
S1(conf-if-gi-12/3)#ip address 20.1.1.5/24
S1(conf-if-gi-12/3)#vrrp-group 15
% Info: The VRID used by the VRRP group 15 in VRF 3 will be 243.
S1(conf-if-gi-12/3-vrid-105)#priority 255
S1(conf-if-gi-12/3-vrid-105)#virtual-address 20.1.1.5
S1(conf-if-gi-12/3)#no shutdown
```

Figure 58-27. VRRP in VRF: Switch-2 Non-VLAN Configuration

Switch-2

```
S2(conf)#ip vrf default-vrf 0
!
S2(conf)#ip vrf VRF-1 1
!
S2(conf)#ip vrf VRF-2 2
!
S2(conf)#ip vrf VRF-3 3
!
S2(conf)#interface GigabitEthernet 12/1
S2(conf-if-gi-12/1)#ip vrf forwarding VRF-1
S2(conf-if-gi-12/1)#ip address 10.10.1.2/24
S2(conf-if-gi-12/1)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 1 will be 177.
S2(conf-if-gi-12/1-vrid-101)#priority 255
S2(conf-if-gi-12/1-vrid-101)#virtual-address 10.10.1.2
S2(conf-if-gi-12/1)#no shutdown
!
S2(conf)#interface GigabitEthernet 12/2
S2(conf-if-gi-12/2)#ip vrf forwarding VRF-2
S2(conf-if-gi-12/2)#ip address 10.10.1.2/24
S2(conf-if-gi-12/2)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 2 will be 178.
S2(conf-if-gi-12/2-vrid-101)#priority 255
S2(conf-if-gi-12/2-vrid-101)#virtual-address 10.10.1.2
S2(conf-if-gi-12/2)#no shutdown
!
S2(conf)#interface GigabitEthernet 12/3
S2(conf-if-gi-12/3)#ip vrf forwarding VRF-3
S2(conf-if-gi-12/3)#ip address 20.1.1.6/24
S2(conf-if-gi-12/3)#vrrp-group 15
% Info: The VRID used by the VRRP group 15 in VRF 3 will be 243.
S2(conf-if-gi-12/3-vrid-105)#priority 100
S2(conf-if-gi-12/3-vrid-105)#virtual-address 20.1.1.5
S2(conf-if-gi-12/3)#no shutdown
```

VLAN Scenario

In another scenario, VRF-1, VRF-2, and VRF-3 use a single physical interface with multiple tagged VLANS (instead of separate physical interfaces) to connect to the LAN. In this case, three VLANs are configured: VLAN-100, VLAN-200, and VLAN-300. Each VLAN is a member of one VRF. A physical interface (gigabitethernet 0/1) attaches to the LAN and is configured as a tagged interface in VLAN-100, VLAN-200, and VLAN-300. The rest of this user case is the same as the non-VLAN scenario.

This VLAN scenario often occurs in a service-provider network in which VLAN tags are configured for traffic from multiple customers on customer-premises equipment (CPE), and separate VRF instances associated with each VLAN are configured on the provider edge (PE) router in the point-of-presence (POP).

Figure 58-28. VRRP in VRF: Switch-1 VLAN Configuration

Switch-1

```
S1(conf)#ip vrf VRF-1 1
!
S1(conf)#ip vrf VRF-2 2
!
S1(conf)#ip vrf VRF-3 3
!
S1(conf)#interface GigabitEthernet 12/4
S1(conf-if-gi-12/4)#no ip address
S1(conf-if-gi-12/4)#switchport
S1(conf-if-gi-12/4)#no shutdown
!
S1(conf-if-gi-12/4)#interface vlan 100
S1(conf-if-vl-100)#ip vrf forwarding VRF-1
S1(conf-if-vl-100)#ip address 10.10.1.5/24
S1(conf-if-vl-100)#tagged gigabitethernet 12/4
S1(conf-if-vl-100)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 1 will be 177.
S1(conf-if-vl-100-vrid-101)#priority 100
S1(conf-if-vl-100-vrid-101)#virtual-address 10.10.1.2
S1(conf-if-vl-100)#no shutdown
!
S1(conf-if-gi-12/4)#interface vlan 200
S1(conf-if-vl-200)#ip vrf forwarding VRF-2
S1(conf-if-vl-200)#ip address 10.10.1.6/24
S1(conf-if-vl-200)#tagged gigabitethernet 12/4
S1(conf-if-vl-200)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 2 will be 178.
S1(conf-if-vl-200-vrid-101)#priority 100
S1(conf-if-vl-200-vrid-101)#virtual-address 10.10.1.2
S1(conf-if-vl-200)#no shutdown
!
S1(conf-if-gi-12/4)#interface vlan 300
S1(conf-if-vl-300)#ip vrf forwarding VRF-3
S1(conf-if-vl-300)#ip address 20.1.1.5/24
S1(conf-if-vl-300)#tagged gigabitethernet 12/4
S1(conf-if-vl-300)#vrrp-group 15
% Info: The VRID used by the VRRP group 15 in VRF 3 will be 243.
S1(conf-if-vl-300-vrid-101)#priority 255
S1(conf-if-vl-300-vrid-101)#virtual-address 20.1.1.5
S1(conf-if-vl-300)#no shutdown
```

Figure 58-29. VRRP in VRF: Switch-2 VLAN Configuration

Switch-2

```
S2(conf)#ip vrf VRF-1 1
!
S2(conf)#ip vrf VRF-2 2
!
S2(conf)#ip vrf VRF-3 3
!
S2(conf)#interface GigabitEthernet 12/4
S2(conf-if-gi-12/4)#no ip address
S2(conf-if-gi-12/4)#switchport
S2(conf-if-gi-12/4)#no shutdown
!
S2(conf-if-gi-12/4)#interface vlan 100
S2(conf-if-vl-100)#ip vrf forwarding VRF-1
S2(conf-if-vl-100)#ip address 10.10.1.2/24
S2(conf-if-vl-100)#tagged gigabitethernet 12/4
S2(conf-if-vl-100)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 1 will be 177.
S2(conf-if-vl-100-vrid-101)#priority 255
S2(conf-if-vl-100-vrid-101)#virtual-address 10.10.1.2
S2(conf-if-vl-100)#no shutdown
!
S2(conf-if-gi-12/4)#interface vlan 200
S2(conf-if-vl-200)#ip vrf forwarding VRF-2
S2(conf-if-vl-200)#ip address 10.10.1.2/24
S2(conf-if-vl-200)#tagged gigabitethernet 12/4
S2(conf-if-vl-200)#vrrp-group 11
% Info: The VRID used by the VRRP group 11 in VRF 2 will be 178.
S2(conf-if-vl-200-vrid-101)#priority 255
S2(conf-if-vl-200-vrid-101)#virtual-address 10.10.1.2
S2(conf-if-vl-200)#no shutdown
!
S2(conf-if-gi-12/4)#interface vlan 300
S2(conf-if-vl-300)#ip vrf forwarding VRF-3
S2(conf-if-vl-300)#ip address 20.1.1.6/24
S2(conf-if-vl-300)#tagged gigabitethernet 12/4
S2(conf-if-vl-300)#vrrp-group 15
% Info: The VRID used by the VRRP group 15 in VRF 3 will be 243.
S2(conf-if-vl-300-vrid-101)#priority 100
S2(conf-if-vl-300-vrid-101)#virtual-address 20.1.1.5
S2(conf-if-vl-300)#no shutdown
```

Displaying a VRRP in VRF Configuration

To display information on a VRRP group that is configured on an interface that belongs to a VRF instance, enter the **show running-config track** [**interface** *interface*] command:

Figure 58-30. Command Example: show running-config track interface

```
FTOS#show running-config interface gigabitethernet 13/4

interface GigabitEthernet 13/4
 ip vrf forwarding red
 ip address 192.168.0.1/24

 vrrp-group 4
  virtual-address 192.168.0.254
 no shutdown
```

To display information on the VRRP groups configured on interfaces that belong to a VRF instance, enter the **show vrrp vrf** [**vrf** *instance*] command:

Figure 58-31. Command Example: show vrrp vrf

```
FTOS#show vrrp vrf red
------------------
GigabitEthernet 13/4, IPv4 Vrrp-group: 4, VRID: 65, Version: 2, Net: 192.168.0.1
VRF: 1 red
State: Master, Priority: 100, Master: 192.168.0.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 0, Bad pkts rcvd: 0, Adv sent: 9, Gratuitous ARP sent: 1
Virtual MAC address:
 00:00:5e:00:01:41
Virtual IP address:
 192.168.0.254
Authentication: (none)
```

59

FTOS XML Feature

FTOS XML Feature is supported on platforms: C E

This chapter describes the FTOS XML Feature in the following major sections:

XML Functionality

Through SSH/Telnet client sessions, FTOS XML provides a way of interfacing with the system by entering XML-formatted requests and retrieving XML output. See The Form of XML Requests and Responses on page 1156.

FTOS XML supports the following functionality:

- Configure both physical and logical interfaces
- Layer 2 and Layer 3 Standard ACLs
- Layer 2 and Layer 3 Extended ACLs
- Supported show commands and their output. Some show command options supported by FTOS are not supported in XML, so each option that is supported in XML is listed separately here for clarity:
  - Protocol commands:
    - **show ip bgp neighbors** (no parameters accepted)
    - **show qos statistics**
    - **show qos statistics wred-profile**
  - System commands:
    - **show chassis**
    - **show rpm** *slot ID*
    - **show rpm all**

- **show linecard** *slot ID*
- **show linecard all**
- **show sfm** *slot ID*
- **show logging** *1-65535*
- **show logging reverse**
- **show sfm**
- **show sfm all**
- **show version**
- **show running-config**—Only the full report is supported, no options.
- **show interfaces**—All the options are supported except **rate**:

The Form of XML Requests and Responses

To send an XML-formatted command through a Telnet or SSH client session, you first use the **terminal xml** command to inform FTOS that you wish to switch to XML mode. See Run an FTOS XML session on page 1159.

Note: FTOS accepts well-formed XML requests, except that it does not currently support XML Namespaces.

Request Format

You can then enter XML-formatted requests that conform to the following schema. Every XML request begins with an XML declaration, followed by a “Method” type tag, followed by an “Operation” type tag, as shown in this shell schema:

```
<?xml version="1.0" encoding="UTF-8"?>
<Request MajorVersion="1" MinorVersion="0">
  <Method>
    <Operation>
      <command>
    :: ! The number of allowed <command> tag sets depends on the type of request. !
    ::
      </command>
    </Operation>
  </Method>
</Request>
```

Currently, for “Method”, you must enter “cli”. In place of “Operation”, you enter either “configuration” or “action”, depending on the CLI mode that you want to invoke:

Namespace	Description
<configuration>	This tag tells the CLI to invoke the CONFIGURATION mode. These requests encapsulate configuration modification commands.

<action>	This tag tells the CLI to invoke the EXEC PRIVILEGE mode. These requests encapsulate “show” commands.

Response Format

Similarly, every response from FTOS begins with the XML declaration, followed by a “Response” tag:

```
<?xml version="1.0" encoding="UTF-8"?>
<Response MajorVersion="1" MinorVersion="0">
    ::
    ::
 </Response>
```

What goes between the Response tags depends on the type of response, as discussed next.

The Configuration Request and Response

To create a configuration request, you know from the introduction above that you put “<cli>” in place of the “<Method>” tag in the schema, and you put “<configuration>” in place of “<Operation>”. The number of configuration commands in one request is not restricted.

Just as you enter commands in the CLI, you have the option of entering abbreviated commands in XML messages. For example, instead of using the full **show running-config** statement, you can enter **show run**. Also, spaces before or after the command are allowed, as shown in the following example.

The following sequence of XML tags shows the structure of a configuration request containing several commands:

```
<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<cli>
<configuration>
<command>ip access standard test2 </command>
<command> seq 10 deny any</command>
<command> seq 20 permit host 10.1.1.1 count </command>
<command>seq 30 deny 10.2.0.0 /16</command>
</configuration>
</cli>
</request>
```

The response from FTOS, if the command executes successfully, is as follows:

```
<?xml version="1.0" encoding="UTF-8"?>
<response MajorVersion="1" MinorVersion="0">
<responseType>NO_ERROR</responseType>
<responseSeverity>SEVERITY_INFO</responseSeverity>
<responseMsg>Xml request successfully processed.</responseMsg>
</response>
```

For details on responses to error conditions, see XML Error Conditions and Reporting on page 1162.

The “Show” Request and Response

To generate an XML request that encapsulates a “show” command (to request a report), you use the <action> tag instead of the <configuration> tag as the Operation type. The schema of a show request allows only one <command>, as shown here for the **show linecard** command. (Note that “<command>**show line all**</command>” demonstrates that you can use both an abbreviated form of the command and options, just as in the standard CLI):

```
<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<cli>
<action>
<command>show line all</command>
</action>
</cli>
</request>
```

The response from FTOS, if the command executes successfully, presents all of the content that you would get in the equivalent CLI report. Note that the data are encapsulated in self-explanatory XML tags. The following is an example of a **show linecard** report embedded in XML tags:

```
<?xml version="1.0" encoding="UTF-8" ?>
<response MajorVersion="1" MinorVersion="0">
<action>
<linecard>
<slotId>3</slotId>
<status>online</status>
<nextBoot>online</nextBoot>
<reqType>E48TF3</reqType>
<numPorts>0</numPorts>
<swVer>6.5.1.1</swVer>
</linecard>
</action>
</response>
```

Configuration Task List

In addition to supporting show commands, FTOS XML currently also supports ACL configuration:

- Run an FTOS XML session on page 1159
- Configure a standard ACL on page 1161
- Configure an extended ACL on page 1161
- Apply an IP ACL on page 1161
- Create an egress ACL and apply rules to the ACL on page 1162

Run an FTOS XML session

Use the following procedure to start, run, and close an FTOS XML session:

Step	Command Syntax	Command Mode	Purpose
1	**terminal xml**	EXEC Privilege	Invoke XML interface in Telnet and SSH client sessions.
2	[Construct input to the CLI by following the XML request schema, as described in The Form of XML Requests and Responses on page 1156.]	FTOS XML	Cut and paste your XML request from a text editor or other type of XML presentation tool, or type your XML request line by line.
3	Press **Ctrl-Y** (or press **Enter** twice, creating an empty line).	FTOS XML	Execute the request. Alternatively, to cancel the request (only possible before sending) and get a fresh XML prompt, press **Ctrl-C**.
4	Press **Ctrl-Z** (or enter **terminal no xml** as the <command> string in the XML request <action> schema).	FTOS XML	Exit from FTOS XML mode.

Figure 59-1, below, illustrates entering FTOS XML mode. Figure 59-1 on page 1159, below, illustrates the full sequence of invoking an XML session, entering a command, receiving a success response, and leaving the session with the **terminal no xml** command in XML:

Figure 59-1. Example of Entering FTOS XML mode from the CL

```
FTOS# terminal xml
FTOS(xml)#
Enter XML request with CTRL-Y or empty line
Clear XML request with CTRL-C
Exit XML mode with CTRL-Z:
```

Figure 59-2. Example of a Successful XML Session

```
FTOS# terminal xml
FTOS(xml)#
Enter XML request with CTRL-Y or empty line
Clear XML request with CTRL-C
Exit XML mode with CTRL-Z:

<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<cli>
<configuration>
<command>ip access standard test1</command>
</configuration>
</cli>
</request>

<?xml version="1.0" encoding="UTF-8"?>
<response MajorVersion="1" MinorVersion="0">
<responseType>NO_ERROR</responseType>
<responseSeverity>SEVERITY_INFO</responseSeverity>
<responseMsg>Xml request successfully processed.</responseMsg>
</response>

FTOS(xml)#
Enter XML request with CTRL-Y or empty line
Clear XML request with CTRL-C
Exit XML mode with CTRL-Z:
<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<cli>
<action>
<command>terminal no xml</command>
</action>
</cli>
</request>

FTOS#
```

Configure a standard ACL

To configure a standard ACL with XML, first enter FTOS XML mode, and then construct a configuration request, as described above. An example of a complete standard ACL configuration request message is:

```
<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<cli>
<configuration>
<command> ip access list standard ToOspf</command>
<command> seq 5 deny any</command>
<command> seq 10 deny 10.2.0.0 /16</command>
<command> seq 15 deny 10.3.0.0 /16</command>
<command> seq 20 deny 10.4.0.0 /16</command>
<command> seq 25 deny 10.5.0.0 /16</command>
<command> seq 30 deny 10.6.0.0 /16</command>
<command> seq 35 deny 10.7.0.0 /16</command>
<command> seq 40 deny 10.8.0.0 /16</command>
<command> seq 45 deny 10.9.0.0 /16</command>
<command> seq 50 deny 10.10.0.0 /16</command>
</configuration>
</cli>
</request>
```

Configure an extended ACL

To configure an extended ACL through XML, enter FTOS XML mode and construct an XML configuration request (see Run an FTOS XML session on page 1159). An example of a complete request message is:

```
<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<cli>
<configuration>
<command> interface GigabitEthernet 0/0</command>
<command> ip address 10.2.1.100 255.255.255.0 </command>
<command> ip access-group nimule in no shutdown</command>
</configuration>
</cli>
</request>
```

Apply an IP ACL

To apply the IP ACL (standard or extended) that you created, above, to a physical or port channel interface, construct an XML configuration request (see Run an FTOS XML session on page 1159) that encapsulates the appropriate CLI commands, as exemplified here:

```
<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<cli>
<configuration>
<command> interface GigabitEthernet 0/0</command>
<command> ip address 10.2.1.100 255.255.255.0 </command>
<command> ip access-group nimule in no shutdown</command>
</configuration>
</cli>
</request>
```

Create an egress ACL and apply rules to the ACL

To create an egress ACL and apply rules to the ACL in one single XML request, first enter FTOS XML mode, and then construct the configuration request (see Run an FTOS XML session on page 1159). The following example shows a configuration request message that accomplishes this task:

```
<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<cli>
<configuration>
<command> interface GigabitEthernet 0/0</command>
<command> ip access-group abcd out</command>
<command> ip access-list extended abcd</command>
<command> seq 5 permit tcp any any</command>
<command> seq 10 deny icmp any any</command>
<command> permit 1.1.1.2</command>
</configuration>
</cli>
</request>
```

XML Error Conditions and Reporting

This section contains examples of various error conditions that might occur in an XML transaction, and the associated responses that the XML generates. Note also, as shown below by the “NO_ERROR” message, that the same response message format is used for a successful configuration request.

The general form of the response is as follows:

```
<?xml version="1.0" encoding="UTF-8"?>
<response MajorVersion="1" MinorVersion="0">
<responseType></responseType>
<responseSeverity></responseSeverity>
<responseMsg></responseMsg>
</response>
```

Summary of XML Limitations

- The XML response to a **show running-configuration** request is encoded in one single XML tag, instead of the standard XML-encoded response.
- A **show** command, in an XML request, requires <action> for the operation tag; the request is not supported if <configuration> is used for the operation tag.
- Only allowed one show command is supported within a single XML request.
- XML namespace is not supported.

Error Messages

The following strings can appear after the <responseType> tag:

- XML_PARSE_ERROR
- CLI_PARSE_ERROR—This error is caused by:
    - — Malformed XML or mismatched XML tags

    - Invalid CLI commands or keywords
    - Invalid range of data specified in the CLI command
- XML_SCHEMA_ERROR—This error is caused by:
    - Invalid XML method or operation tags
    - Invalid object hierarchy or value out of range
- APPLICATION_ERROR—This error is caused by a failure to process the request, or a problem on the FTOS task.
- NO_ERROR—The XML request processed successfully.

The following strings can appear after the <responseSeverity> tag:

- SEVERITY_INFO—This string indicates no error, and is paired with NO_ERROR after the <responseType> tag.
- SEVERITY_ERROR—This string is paired with one of the other four possible <responseType> strings besides NO_ERROR.

The following strings can appear after the <responseMsg> tag:

- “Xml request successfully processed” (paired with NO_ERROR)
- “% Error: Parsing error is detected in the XML request” (paired with XML_PARSE_ERROR)
- “% Error: Schema error is detected in the XML request” (paired with XML_SCHEMA_ERROR)
- “% Error: CLI Parsing error is detected in the XML request” (paired with CLI_PARSE_ERROR)
- “% Error: [content varies, depending on the error]” (paired with APPLICATION_ERROR, indicating an application error from a backend task)

Examples of Error Conditions

XML parsing error

The following XML request is missing the XML declaration (the first line in the schema):

```
<request MajorVersion="1" MinorVersion="0">
<cli>
<configuration>
<command>ip access standard test2</command>
</configuration>
</cli>
</request>
```

The XML response to that malformed request is:

```
<?xml version="1.0" encoding="UTF-8"?>
<response MajorVersion="1" MinorVersion="0">
<responseType>XML_PARSE_ERROR</responseType>
<responseSeverity>SEVERITY_ERROR</responseSeverity>
<responseMsg>% Error: Parsing error detected in the XML request.</responseMsg>
</response>
```

XML schema error

This following XML request has transposed the <configuration> and <cli> tag sets:

```
<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<configuration>
<cli>
<command>ip access standard test2</command>
</cli>
</configuration>
</request>
```

The XML response to that malformed request is:

```
<?xml version="1.0" encoding="UTF-8"?>
<response MajorVersion="1" MinorVersion="0">
<responseType>XML_SCHEMA_ERROR</responseType>
<responseSeverity>SEVERITY_ERROR</responseSeverity>
<responseMsg>% Error: Schema error detected in the XML request.</responseMsg>
</response>
```

XML command error

The following XML request contains an invalid CLI command:

```
<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<cli>
<configuration>
<command>ip access test test1</command>
</configuration>
</cli>
</request>
```

The XML response to that invalid request is:

```
<?xml version="1.0" encoding="UTF-8"?>
<response MajorVersion="1" MinorVersion="0">
<responseType>CLI_PARSE_ERROR</responseType>
<responseSeverity>SEVERITY_ERROR</responseSeverity>
<responseMsg><command>ip access test test1</command></responseMsg>
</response>
```

XML application error

The command in this XML request makes an invalid request:

```
<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<cli>
<configuration>
<command>ip access standard test1</command>
<command>seq 10 permit host 1.2.3.4 log count bytes</command>
</configuration>
</cli>
</request>
```

The error response contains a <responseSeverity> of “APPLICATION_ERROR”, <responseSeverity>SEVERITY_ERROR, and a <responseMsg> of “% Error: Seq number does not exist.”

The second command in this XML request also makes an invalid request:

```
<?xml version="1.0" encoding="UTF-8"?>
<request MajorVersion="1" MinorVersion="0">
<cli>
<configuration>
<command>ip access standard test1</command>
<command>no permit host 2.2.3.4 log count bytes</command>
</configuration>
</cli>
</request>
```

The error response contains a <responseSeverity> of “APPLICATION_ERROR”, <responseSeverity> of “APPLICATION_ERROR” and a <responseMsg> of “% Error: Access-list entry does not exist.”

Using display xml as a Pipe Option

Also, at a CLI prompt in EXEC privilege mode (“enable mode”), you can retrieve XML-formatted responses to the show commands supported by XML (see the list of supported show commands in the section XML Functionality on page 1155). The following table describes how to format a show command with a pipe option that will request that the show command report be presented with XML formatting.

Command Syntax	Command Mode	Purpose
show *keyword* | **display xml**	EXEC privilege	FTOS treats “ | **display xml**” as a request to format the show command report in XML format.

As shown in the following Figure 59-3, FTOS formats the response with the XML tags from the same response schema used by the XML response, discussed in The “Show” Request and Response on page 1158. For more on pipe options, see Filtering show Command Outputs on page 43.

Figure 59-3. Example: show linecard 0 | display xml

```
FTOS>#show linecard 0 | display xml
<?xml version="1.0" encoding="UTF-8" ?>
<response MajorVersion="1" MinorVersion="0">
<action>
<linecard>
<slotId>0</slotId>
<status>online</status>
<nextBoot>online</nextBoot>
<reqType>EXW2PF3 - 2-port 10GE LAN/WAN PHY line card with XFP optics (EF3)</reqType>
<curType>EXW2PF3 - 2-port 10GE LAN/WAN PHY line card with XFP optics (EF3)</curType>
<hwRevBase>1.1</hwRevBase>
<hwRevPortPipe0>1.1</hwRevPortPipe0>
<hwRevPortPipe1>n/a</hwRevPortPipe1>
<numPorts>2</numPorts>
<upTime>1 hr, 32 min</upTime>
<swVer>4.4.3.243</swVer>
<lcJumboCapable>yes</lcJumboCapable>
<lcBootFlashA>2.3.0.6 [booted]</lcBootFlashA>
<lcBootFlashB>2.3.0.6  </lcBootFlashB>
<totMemSize>268435456</totMemSize>
<lcTemperature>37</lcTemperature>
<powerStatus>AC</powerStatus>
<voltage>ok</voltage>
<serialNum>0039034</serialNum>
<partNum>7520017400</partNum>
<productRev>08</productRev>
<vendorId>04</vendorId>
<dateCode>01332005</dateCode>
<countryCode>01</countryCode>
</linecard>
</action>
</response>
FTOS>
```

60

C-Series Debugging and Diagnostics

In addition to standard manageability features such as LEDs, SNMP alarms and traps, and Syslogging, the C-Series supports several diagnostic and debugging features that are crucial to isolating and resolving support issues during the operations and maintenance phase.

Switch Fabric overview

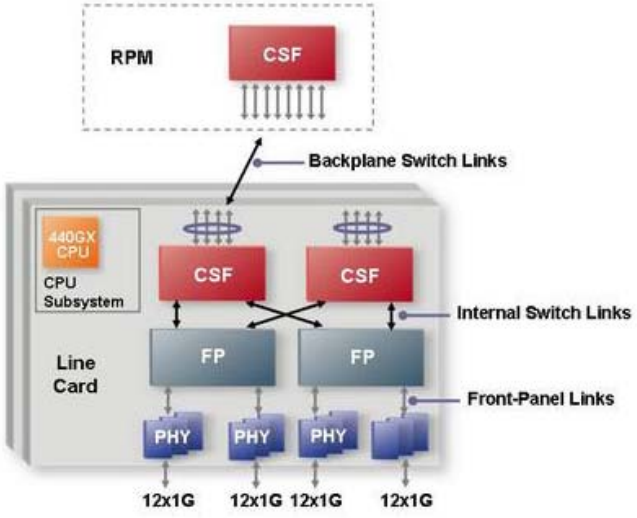
The switch fabric is formed through the installed RPMs and line cards via C-Series Switch Fabric (CSF) ASICs.

Each RPM includes four CSFs, each of which provides eight Backplane Data (BDP) links, one link for each line card slot. In total, an RPM provides 32 BDP links of forwarding capacity.

Each line card includes two CSFs. Six of the eight links on the CSFs are used as follows:

- Up to four ports—ports 1 to 4—connect to the Forwarding Processors (FP). These ports are referred to as the Internal Dataplane (IDP) link.
- Ports 5-8 connect to the RPMs. These ports are referred to as the BDP links.

Figure 60-1. Architecture Diagram of the 1x48GE Line Card



The number of FPs varies with the line card type, as shown in Table 60-1.

Table 60-1. FPs, CSFs, and IDP Links by Line Card Type

Line Card Type	# of FPs	# of CSFs	IDP Links Used on CSF
48x1GE	2	2	• Ports 1 and 2 connect to the FPs. • Ports 3 and 4 are unused.
96x1GE	4	2	Ports 1-4 connect to the FPs.
4x10GE	2	2	Ports 1-4 connect to the FPs.

Switch Fabric link monitoring

FTOS Switch Manager (SWMGR) task monitors the BDP links on the RPM. This task also monitors the overall state of the switch fabric and reports any changes via Syslog messages.

FTOS Switch Agent (SWAGT) monitors the IDP and BDP links on the line cards.

FTOS Link Monitoring task continually polls the status of the IDP and BDP links. If it finds an open link, the system brings down the link and reports the condition via a message similar to the one shown in Message 1.

Message 1 FTOS Link Monitoring Syslog Message Example

```
Mar 12 21:01:18: %RPM1-P:CP %SWMGR-1-BDP_LINK_DETECT: Backplane datapath link status for RSM0 Switch
fabric unit# 0 port# 0 => DOWN
!- Describes only the state of the port.
Mar 10 16:58:28: %RPM1-P:CP %SWMGR-1-IDP_LINK_DETECT: Internal datapath link status for Linecard#5
Switch unit# 1 port# 24 and Switch fabric unit# 3 port# 1=> DOWN
!- Describes the state of the link.
```

Note: These messages are not reported when a line card is reset by a user command.

If a backplane link on a line card goes down, the RPM side of the link stays up to avoid duplicate reporting.

Bringing down an IDP or BDP link causes the card to be powered-off and placed into a "card problem - port pipe problem" state. Use the **show linecard** command to view the status of a line card.

If a single BDP link to the active RPM is down, the line card will be placed in an error state. Use the **show switch links** command to view the status of the dataplane links, as shown in Figure 60-2.

Figure 60-2. show switch links backplane Command Example

```
FTOS#show switch links backplane
Switch fabric backplane link status:
SFM0 Links Status              SFM1 Links Status
LC SlotID   Port0 | Port1 | Port2 | Port3 | Port4 | Port5 | Port6 | Port7
   0        not present
   1        not present
   2        not present
   3        not present
   4        not present
   5         up      up      up      up    up/down up/down up/down up/down
   6        not present
   7        not present
up - Both ends of the link are up
down - Both ends of the link are down
up / down - SFM side up and LC side down
down / up - SFM side down and LC side up
```

To monitor the status of a virtual SFM, use the **show sfm** command shown in Figure 60-3. The system reports an “active” status if all CSF ASICs on the RPM initialize successfully, whether or not any line cards are installed and the BDP links are up.

Figure 60-3. show sfm Command Example

```
FTOS#show sfm
Switch Fabric State:  up
-- SFM 0 --
Status        : active
Module Type   : SFM - Switch Fabric Module
Up Time       : 1 day, 6 hr, 0 min
-- SFM 1 --
Status        : not present
```

Use the FTOS Syslogging feature to monitor the overall status of the switch fabric. Changes in switch fabric status are reported via messages similar those in Message 2.

Message 2 Switch Fabric Status Change Syslog Message Example

```
00:00:13: %RPM1-P:CP %TSM-6-SFM_FULL_PARTIAL_STATE: SW_FAB_UP_1 SFM in the system
00:00:13: %RPM1-P:CP %TSM-6-SFM_SWITCHFAB_STATE: Switch Fabric: UP
```

Runtime hardware status monitoring

The FTOS Poll Manager (POLLMGR) process reads the key status registers on hardware sub-components to pro-actively identify and report a hardware fault. An example Syslog message is shown in Message 3.

Message 3 Poll Manager Syslog Message Example

```
%RPM1-P:CP %POLLMGR-2-BPL_IRC_ERR: Back Plane Link Error
```

The Poll Manager runs automatically in the background and cannot be disabled. The possible Poll Manager Syslog messages are given in Table 60-2.

Table 60-2. Poll Manager Syslog Message Description

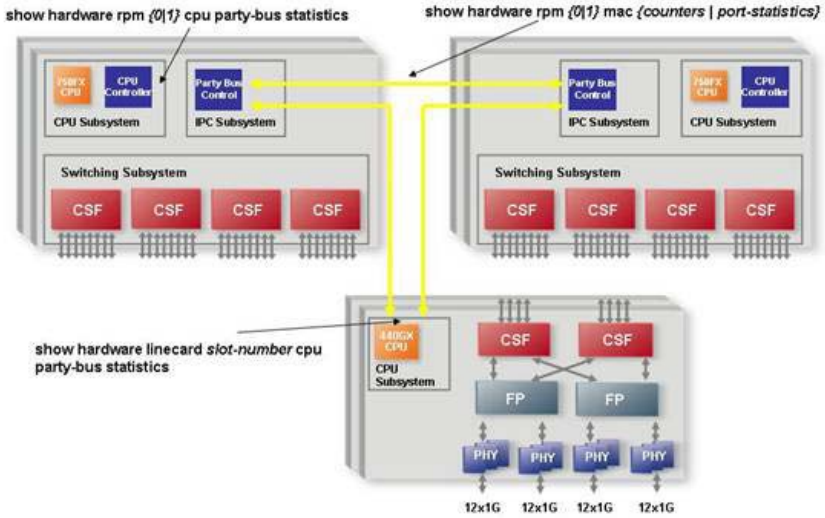
Message	Description
POLLMGR-2-ALT_RPM_STATE	Reports that either the standby RPM is not present or has been detected.
POLLMGR-2-USER_FLASH_DETECT	Reports the status of the flash disks. If reported during boot up, this message indicates either: **a** External flash disk missing in slot0: **b** Internal flash disk missing in flash: If reported during runtime, this message indicates either: **a** External flash disk removed from slot0: **b** Internal flash disk removed from flash:

Table 60-2. Poll Manager Syslog Message Description

Message	Description
POLLMGR-2-POLLMGR_RPM_ECC_ERR_DETECT	Indicates that the system detected a single-bit ECC memory error in the RPM CPU memory (SDRAM). The system tracks the number of multi-bit errors and resets the system after a certain number of such errors are recorded. Upon reset, the system writes a failure trace file to the TRACE_LOG directory for analysis by Dell Force10.
POLLMGR-2-POLLMGR_BPL_IRC_ERR	Indicates that the system detected an error on the internal IPC switch subsystem connection between the two RPMs. This connection is referred to as Inter-RPM Communication (IRC). When a number of consecutive IRC heartbeat messages are lost, the system will declare an IRC timeout via a Syslog message and reset the system. This message suggests that a hardware fault on the RPM may have caused the IRC timeout. To troubleshoot this issue: • Verify that the RPMs are fully inserted. • Try a swap test of the RPMs. • Capture the output of the following show hardware commands: **•show hardware rpm** *number* **cpu party-bus statistics** **•show hardware rpm** *number* **mac counters** **•show hardware rpm** *number* **mac port-statistics rpm** *number* (of alternate RPM)
POLLMGR-2-POLLMGR_PTYBUS_LINK_SCAN	Indicates the internal IPC party bus connection to a line card has changed to down. IPC, or inter-process communication, is the protocol used among the RPM and line card CPUs to exchange information. The underlying IPC subsystem uses internal Ethernet links. To troubleshoot this condition: • Capture the output of the **show hardware linecard cpu party-bus statistics** command, and forward it to Dell Force10.

Figure 60-4 illustrates the IPC subsystem, including the IRC links between the RPMs, and the relevant troubleshooting commands.

Figure 60-4. IPC Sub-system



Inter-CPU timeouts

The CP monitors the health status of the other processors using heartbeat messaging exchange.

FTOS automatically saves critical information about the IPC failure to NVRAM. Such information includes:

- Status counters on the internal Ethernet interface
- Traffic profile of the inter-CPU bus
- Kernel drops
- High CPU exception conditions

Upon the next boot, this information is uploaded to a file in the CRASH_LOG directory. Use the following command sequence beginning in EXEC mode to capture this file for analysis by the Dell Force10 TAC.

Step	Task	Command	Mode
1	Display the directories in flash memory. The output should include: **1 drwx 2048 Jan 01 1980 00:00:06 CRASH_LOG_DIR**	**dir flash:**	EXEC
2	Change to the CRASH_LOG directory.	**cd CRASH_LOG_DIR**	EXEC
3	View any saved files in the CRASH_LOG directory. The naming convention is: *sysinfo_RPMIDProcessorID _ timestamp* For example: **sysinfo_RPM1CP_20060616_013125**	**dir**	EXEC Privilege
4	View the contents of the file.	**show file flash://CRASH_LOG_DIR/[file_name]**	EXEC Privilege

In a dual RPM system, the two RPMs send synchronization messages via inter-RPM communication (IRC). As described in the High Availability chapter, an RPM failover can be triggered by loss of the heartbeat (similar to a keepalive message) between the two RPMs. FTOS reports this condition via syslog messages, as follows:

Message 4 RPM Failover Syslog Messages

```
20:29:07: %RPM1-S:CP %IRC-4-IRC_WARNLINKDN: Keepalive packet 7 to peer RPM is lost
20:29:07: %RPM1-S:CP %IRC-4-IRC_COMMDOWN: Link to peer RPM is down
%RPM1-S:CP %RAM-4-MISSING_HB: Heartbeat lost with peer RPM. Auto failover on heart beat lost.
%RPM1-S:CP %RAM-6-ELECTION_ROLE: RPM1 is transitioning to Primary RPM.
```

FTOS automatically saves critical information, about the IRC failure, to NVRAM. Use the same three-step procedure to capture this file for analysis by Dell Force10.

FTOS actually saves up to three persistent files depending upon the type of failure. When reporting an RPM failover triggered by a loss of the IPC or IRC heartbeats, look for failure records in the following directories:

- Application or kernel core dump RP in the CORE_DUMP_DIR
- CP trace log file (look for a filename with the phrase “failure_trace”) in the TRACE_LOG_DIR
- RP and/or CP sysinfo file in the CRASH_LOG_DIR, as explained above

Bootup diagnostics

During bootup and reset of a card, diagnostics check the status of key hardware sub-components and verify that all ASICs initialize successfully.

Recognizing bootup failure

Any detected failures or errors during bootup are reported via Syslog. The messages in Message 5 and Message 6 might be reported for line card failures and RPMs, respectively.

Message 5 Line Card Boot Up Failure Syslog Messages

```
%CHAGT-5-LINK_STATUS_DOWN: Link status bad for port pipe [number] on line card [number]
%CHAGT-5-PORT PIPE DOWN: Port pipe [number] down or errored on line card [number]
```

Message 6 RPM Boot Up Failure Syslog Messages

```
%CHMGR-2-RPM_ISOLATED: Active RPM is unable to talk to line cards
%CHMGR-3-RAM_STANDBY_RPM_FAULT: Secondary RPM fault detected
%CHMGR-3-RPM_POST_PORTPIPE_FAIL: RPM port pipe fails on boot up test
%CHMGR-3-RPM_DRIVER_OPEN_FAIL: RPM driver open fails on boot up
%CHMGR-3-RPM_SWITCH_OPEN_FAIL: RPM switch driver fails on boot up
%CHMGR-3-RPM_POST_RTC_FAIL: RPM RTC fails on boot up test
```

Troubleshoot bootup failure

If these messages are seen, collect the output of the **show console lp** and **show tech** commands and contact the Dell Force10 Technical Assistance Center.

Environmental monitoring

All C-Series components use environmental monitoring hardware to detect overtemperature, undervoltage, and overvoltage conditions. Use the **show environment** command to monitor the components for any major or minor alarm conditions. The output in Figure 60-5 displays the environment status of the RPM.

Figure 60-5. show environment rpm Command Example

```
FTOS#show environment rpm

--  RPM Environment Status  --
Slot  Status        Temp   Voltage
-------------------------------------
  0   active        33C    ok
  1   not present
```

Recognize an overtemperature condition

An overtemperature condition occurs, for one of two reasons:

- The card genuinely is too hot.
- A sensor has malfunctioned.

Inspect cards adjacent to the one reporting the condition to discover the cause.

- If directly adjacent cards are not normal temperature, suspect a genuine overheating condition.
- If directly adjacent cards are normal temperature, suspect a faulty sensor.

When the system detects a genuine over-temperature condition, it powers off the card. To recognize this condition, look for the system messages in Message 7.

Message 7 Over Temperature Condition System Messages

```
CHMGR-2-MAJOR_TEMP: Major alarm: chassis temperature high (temperature reaches or exceeds threshold of [value]C)
CHMGR-2-TEMP_SHUTDOWN_WARN: WARNING! temperature is [value]C; approaching shutdown threshold of [value]C
```

To view the programmed alarm thresholds levels, including the shutdown value, execute the **show alarms threshold** command shown in Figure 60-6.

Figure 60-6. show alarms threshold Command Example

```
FTOS#show alarms threshold

--  Temperature Limits (deg C)  --
-------------------------------------------------------------
          Minor     Minor Off   Major    Major Off   Shutdown
Linecard   75         70         80         77         85
RPM        65         60         75         70         80
FTOS#
```

Troubleshoot an overtemperature condition

To troubleshoot an over-temperature condition:

1. Use the **show environment** commands to monitor the temperature levels.

2. Check air flow through the system. On the C-Series, air flows sideways from right to left. Ensure the air ducts are clean and that all fans are working correctly.
3. Once the software has determined that the temperature levels are within normal limits, the card can be re-powered safely. Use the **power-on** command in EXEC mode to bring the line card back online.

In addition, Dell Force10 requires that you install blanks in all slots without a line card to control airflow for adequate system cooling.

Note: Exercise care when removing a card; if it has exceeded the major or shutdown thresholds, the card could be hot to the touch!

Recognize an under-voltage condition

If the system detects an under-voltage condition and declares an alarm. To recognize this condition, look for the system messages in Message 8.

Message 8 Under-voltage Condition System Messages

```
%CHMGR-1-CARD_SHUTDOWN: Major alarm: Line card 2 down - auto-shutdown due to under voltage
```

This message in Message 8 indicates that the specified card is not receiving enough power. In response, the system first shuts down Power over Ethernet (PoE). If the under-voltage condition persists, line cards are shut down, then RPMs.

Troubleshoot an under-voltage condition

To troubleshoot an under-voltage condition, check that the correct number of power supplies are installed and their Status LEDs are lit.

Trace logs

In addition to the syslog buffer, FTOS buffers trace messages which are continuously written by various FTOS software tasks to report hardware and software events and status information. Each trace message provides the date, time, and name of the FTOS process. All messages are stored in a ring buffer and can be saved to a file either manually or automatically upon failover.

Some trace files are automatically saved and stored in the **flash:/TRACE_LOG_DIR** directory for SW and HW Traces of the CP and for all Linecards. This directory contains the **TRACE_CURR_BOOT** directory which in turn contains the saved trace buffer files.

The **TRACE_LOG_DIR/TRACE_CURR_BOOT** directory can be reached by FTP or by using the **show file** command from the **flash://TRACE_LOG_DIR** directory.

Note: At reload this directory is renamed to **flash:/TRACE_LOG_DIR/TRACE_LAST_BOOT** and a new empty **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT** directory is created.

Automatic trace log updates

The system automatically saves trace files to the internal flash. The files are saved to the TRACE_LOG_DIR directory on the flash, and are named so that they can be viewed in a logical order. The first automatic CP software trace file is labeled **sw_trace_RPM0CP.0**, and the first automatic CP hardware trace file is labeled **hw_trace_RPM0CP.0**. Up to five trace logs are saved before the system begins overwriting them (sw_trace_RPM0CP.1, and so on until sw_trace_RPM0CP.4, hw_trace_RPM0CP.1, and so on until hw_trace_RPM0CP.4).

These files are saved in **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT**. At reload this directory is renamed to **flash:/TRACE_LOG_DIR/TRACE_LAST_BOOT** and an empty **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT** directory is created.

Trace file *hw_trace_RPM0CP.0* is not overwritten so that chassis bootup message are preserved.

The CP and LP trace file names are:

- **CP [SW trace]** : sw_trace_RPM0CP.0, sw_trace_RPM0CP.1, sw_trace_RPM0CP.2, sw_trace_RPM0CP.3 and sw_trace_RPM0CP.4
- **CP [HW trace]** : hw_trace_RPM0CP.0, hw_trace_RPM0CP.1, hw_trace_RPM0CP.2, hw_trace_RPM0CP.3 and hw_trace_RPM0CP.4
- **LP [SW trace]** : sw_trace_LP[0-7].0, sw_trace_LP[0-7].1, sw_trace_LP[0-7].2, sw_trace_LP[0-7].3 and sw_trace_LP[0-7].4
- **LP [HW trace]** : hw_trace_LP[0-7].0, hw_trace_LP[0-7].1, hw_trace_LP[0-7].2, hw_trace_LP[0-7].3 and hw_trace_LP[0-7].4

Trace files are saved in the directory **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT**. Upon a system reload this directory is renamed **flash:/TRACE_LOG_DIR/TRACE_LAST_BOOT**, and an empty **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT** directory is created.

Save a hardware log to a file on the flash

The RPM and line card trace logs are enabled by default. The trace logs are saved automatically but you can save the contents of a buffer manually with the CLI. The files are named the same whether they are saved automatically or manually.

To manually write the contents of an RPM log to the internal flash:

Task	Command Syntax	Command Mode
Write the RPM trace log to flash.	**upload trace-log cp** [cmd-history | hw-trace | sw-trace]	EXEC Privilege

To manually write the contents of a line card log to the internal flash:

Task	Command Syntax	Command Mode
Write the line card trace log to flash.	**upload trace-log linecard** *[0-7]* **| [hw-trace | sw-traceupload trace-log cp** | [cmd-history | hw-trace | sw-trace]	EXEC Privilege

Figure 60-7. TRACE_CURR_BOOT Directory example

```
FTOS#cd /flash/TRACE_LOG_DIR/TRACE_CURR_BOOT
FTOS#dir
hw_traceLP0.0       hw_traceLP2.1       hw_traceLP2.4       hw_traceRPM0_CP.2
hw_traceLP0.1       hw_traceLP2.2       hw_traceRPM0_CP.0   hw_traceRPM0_CP.3
hw_traceLP2.0       hw_traceLP2.3       hw_traceRPM0_CP.1   hw_traceRPM0_CP.4
```

Disable writing the contents of a hardware log to the internal the flash:

Task	Command Syntax	Command Mode
Stop the writing the hardware log to flash.	**trace disable**	EXEC Privilege

View the hardware log contents or clear them with the following commands.:

Task	Command Syntax	Command Mode
View the content of the hardware log.	**show trace hardware** *slot-number*	EXEC Privilege
Clear the hardware log buffer	**clear trace hardware**	EXEC Privilege

Manual reload messages

When the chassis is reloaded manually (through the CLI), trace messages in all of the buffers (software and hardware) in CP and linecards are saved to the flash as *reload_traceRPM0_CP* and *reload_traceLP1* in **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT**. After reload, you can see these files in **flash:/TRACE_LOG_DIR/TRACE_LAST_BOOT**.

When the trace messages are being saved on reload, Message 9 is displayed.

Message 9 Saving Trace Messages

```
Starting to save trace messages… Done.
```

The CP and LP trace file names at chassis reload are:

- **CP:** reload_traceRPM0_CP
- **LP:** reload_traceLP[0-7]

Figure 60-8. TRACE_LAST_BOOT Directory example

```
FTOS#cd /flash/TRACE_LOG_DIR/TRACE_LAST_BOOT
FTOS#dir
hw_traceLP0.0       hw_traceLP2.3       hw_traceRPM0_CP.3   reload_traceLP6
hw_traceLP0.1       hw_traceLP2.4       hw_traceRPM0_CP.4   reload_traceRPM0_CP
hw_traceLP2.0       hw_traceRPM0_CP.0   reload_traceLP0
hw_traceLP2.1       hw_traceRPM0_CP.1   reload_traceLP2
hw_traceLP2.2       hw_traceRPM0_CP.2   reload_traceLP3
```

CP software exceptions

When a RPM resets due to a software exception, the linecard trace files are saved to **flash:/TRACE_LOG_DIR** directory.

The CP and LP trace file names in the case of a software exception are:

- **CP:** failure_trace_RPM1_CP
- **LP:** failure_trace_RPM1_LP[0-7]

For systems with a single RPM, the linecard traces are saved on the failed RPM itself.

For systems with dual RPM, linecard trace logs are saved when the CP, RP1, or RP2 crashes. The linecard trace logs are saved on the new Primary RPM. The linecard trace file name identifies the failed RPM. For example, if RPM0 fails the trace files saved in RPM1 with filename as **failure_trace_RPM0_LP1**.

Command history

The command-history trace feature captures all commands entered by all users of the system with a time stamp and writes these messages to a dedicated trace log buffer. When the **show command-history** command is entered, the system displays a trace message for each executed command. No password information is saved to the file. The trace messages are not saved to a file or directory on the internal system flash.

To view the command-history trace, use the **show command-history** command, as shown in Figure 60-9.

Figure 60-9. Command Example: show command-history

```
FTOS#show command-history
[6/16 16:22:3]: CMD-(CLI):[enable]by default from console
[6/16 16:22:6]: CMD-(CLI):[show cam-profile]by default from console
[6/16 16:38:9]: CMD-(TEL0):[enable]by admin from vty0 (10.11.48.30)
[6/16 16:38:10]: CMD-(CLI):[show qos statistics]by default from console
[6/16 16:38:21]: CMD-(TEL0):[show command-history]by admin from vty0 (10.11.48.30)
```

Clearing the command history

Clear the command history buffer using the command **clear command-history** from EXEC Privilege mode, as shown in Figure 60-10.

Figure 60-10. Clearing the Command History

```
FTOS#show command-history 10
[12/3 15:40:17]: CMD-(CLI):[show config]by default from console
[12/3 15:40:22]: CMD-(CLI):[ping 10.11.80.201]by default from console
[12/3 15:40:46]: CMD-(CLI):[show interfaces managementethernet 0/0]by default from console
[12/3 15:40:49]: CMD-(CLI):[shutdown]by default from console
[12/3 15:40:59]: CMD-(CLI):[no shutdown]by default from console
[12/3 15:41:1]: CMD-(CLI):[interface managementethernet 0/0]by default from console
[12/3 15:41:2]: CMD-(CLI):[shutdown]by default from console
[12/3 15:41:7]: CMD-(CLI):[ping 10.11.80.201]by default from console
[12/3 21:45:46]: CMD-(CLI):[enable]by default from console
[12/3 21:47:18]: CMD-(CLI):[show command-history 10]by default from console
FTOS#clear command-history
FTOS#show command-history 10
[12/3 21:47:43]: CMD-(CLI):[show command-history 10]by default from console
FTOS#
```

Advanced debugging commands

The C-Series supports advanced debugging commands for isolating suspected software and hardware support issues.

debug commands

The debug command tree provides packet- and event-level debugging for major protocols, as shown in Figure 60-11.

Figure 60-11. debug Command Tree

```
FTOS#debug ?
aaa                    AAA debug information
arp                    IP ARP debug information
cpu-traffic-stats      Start collecting CPU traffic statistics
dot1x                  Dot1x debug information
ftpserver              FTP Server debug information
ifm                    IFM Debug information
ip                     IP debug information
ntp                    NTP debug information
ppp                    PPP debug commands
radius                 RADIUS debug information
spanning-tree          Spanning tree debug information
tacacs+                TACACS+ debug information
```

show hardware commands

The **show hardware** command tree consists of EXEC Privilege commands that have been created or changed specially for use with the C-Series. These commands display information from a hardware sub-component, such as an FP or CSF ASIC, and from hardware-based feature tables.

Table 60-3 lists the **show hardware** commands available as of the latest FTOS version.

Note: show hardware commands should only be used under the guidance of Dell Force10 Technical Assistance Center.

Table 60-3. show hardware Commands

Command	Description
show hardware interface phy	View link status information, including the transmitted and received auto-negotiation control words. Use the **registers** keyword to capture a dump of key PHY registers for your technical support representative.
show hardware drops	View internal packet-drop counters on a line card or RPM
show hardware rpm mac counters **clear hardware rpm mac counters**	Enter the keyword **counters** keyword to view or clear the receive and transmit frame counters for the party bus switch in the IPC subsystem on the RPM.
show hardware rpm mac port-statistics **clear hardware rpm mac port-statistics**	Enter the keyword **port-statistics** to view or clear detailed Ethernet statistics for the specified port on the party bus switch.
show hardware rpm cpu management	View internal interface status information for the RPM CPU port which connects to the external management interface.
show hardware cpu party-bus **clear hardware cpu party-bus**	View or clear statistics for the party-bus port on the CPU of the specified line card or RPM.

Table 60-3. show hardware Commands

Command	Description
show hardware cpu data-plane	View driver-level statistics for the data-plane port on the CPU for the specified line card or RPM.
show hardware unit	Views advanced counters, statistics, and register information for the FP and CSF ASICs.

Recognizing a High CPU Condition

A high CPU condition exist when any of the messages in Message 10 appear.

Message 10 High CPU Condition

```
Feb 13 13:56:16: %RPM1-S:CP %CHMGR-5-TASK_CPU_THRESHOLD: Cpu usage above threshold for task
"sysAdmTsk"(100.00%) in CP.

Feb 13 13:56:20: %RPM1-S:CP %CHMGR-5-CPU_THRESHOLD: Overall cp cpu usage above threshold. Cpu5SecUsage
(100)

Feb 13 13:56:20: %RPM1-S:CP %CHMGR-5-TASK_CPU_THRESHOLD_CLR: Cpu usage drops below threshold for task
"sysAdmTsk"(0.00%) in CP.
```

Troubleshoot a high CPU condition

If FTOS indicates a high CPU condition or you suspect one:

Step	Task	Command Syntax	Command Mode
1	Enable **debug cpu-traffic-stats**, and monitor the output with the **show cpu-traffic-stats** command. These commands indicate the physical interface through which the questionable traffic is arriving.	**debug cpu-traffic-stats** **show cpu-traffic-stats**	CONFIGURATION EXEC Privilege
2	Review the **show ip traffic** command output. This command displays the types of IP traffic destined to the CPU.	**show ip traffic**	EXEC Privilege

Monitoring hardware components with SNMP

The SNMP traps and OIDs in Table 60-4 provide information on C-Series hardware components.

Table 60-4. SNMP Traps and OIDs

OID String	OID Name	Description
RPM		
.1.3.6.1.4.1.6027.3.1.1.3.8	chRPMMajorAlarmStatus	Fault status of the major alarm LED on the RPM
.1.3.6.1.4.1.6027.3.1.1.3.9	chRPMMinorAlarmStatus	Fault status of the minor alarm LED on the RPM
.1.3.6.1.4.1.6027.3.1.1.4.0.11	chAlarmRpmUp	Trap generated when the status of primary or secondary RPM changes to up and running
.1.3.6.1.4.1.6027.3.1.1.4.0.12	chAlarmRpmDown	Trap generated when the status of primary or secondary RPM changes to down, either by software reset or by being physically removed from the chassis
Line Card		
.1.3.6.1.4.1.6027.3.1.1.2.3.1.15	chSysCardOperStatus	Operational status of the card. • If the chSysCardAdminStatus is up, the valid state is ready—the card is present and ready and the chSysCardOperStatus status is up. • If the chSysCardAdminStatus is down the service states can be: • **offline**: the card is not used. • **cardNotmatch**: the card does not match what is configured • **cardProblem**: a hardware problem has been detected on the card. • **diagMode**: the card is in the diagnostic mode. **Note:** chSysCardFaultStatus is supported only the C-Series.
.1.3.6.1.4.1.6027.3.1.1.4.0.1	chAlarmCardDown	Trap reported when a card operational status changes to down
.1.3.6.1.4.1.6027.3.1.1.4.0.2	chAlarmCardUp	Trap reported when a card operational status changes to up
.1.3.6.1.4.1.6027.3.1.1.4.0.3	chAlarmCardReset	Trap reported when a card is reset
.1.3.6.1.4.1.6027.3.1.1.4.0.7	chAlarmCardProblem	Trap reported when a card operational status changes to card problem
.1.3.6.1.4.1.6027.3.1.1.1.4.0.5	chAlarmCardMismatch	Trap generated when the configured card does not match the installed card
.1.3.6.1.4.1.6027.3.1.1.1.4.0.6	chAlarmCardRemove	Trap generated when a card is removed
Power Supply Unit		

Table 60-4. SNMP Traps and OIDs

OID String	OID Name	Description
.1.3.6.1.4.1.6027.3.1.1.2.1.1.2	chSysPowerSupplyOperStatus	Each entry in the chSysPowerSupplyTable includes a set of objects which describe the status of a particular power supply.
.1.3.6.1.4.1.6027.3.1.1.4.0.13	chAlarmPowerSupplyDown	Trap generated when the power supply status changes to non-operational
.1.3.6.1.4.1.6027.3.1.1.4.0.17	chAlarmPowerSupplyClear	Trap generated when the power supply status changes to operational.
.1.3.6.1.4.1.6027.3.1.1.4.0.32	chAlarmMajorPS	Trap generated when a power supply major alarm is issued
.1.3.6.1.4.1.6027.3.1.1.4.0.33	chAlarmMajorPSClr	Trap generated when a power supply major alarm is cleared
.1.3.6.1.4.1.6027.3.1.1.4.0.34	chAlarmMinorPS	Trap generated when a power supply minor alarm is issued
.1.3.6.1.4.1.6027.3.1.1.4.0.35	chAlarmMinorPSClr	Trap generated when a power supply minor alarm is cleared
Fan Tray		
.1.3.6.1.4.1.6027.3.1.1.2.2.1.2	chSysFanTrayOperStatus	Each entry in the chSysFanTrayTable includes a set of objects that describe the status of a particular fan tray, as identified by the chSysFanTrayIndex
.1.3.6.1.4.1.6027.3.1.1.4.0.36	chAlarmMinorFanBad	Trap generated when the status of one or more fans changes to down, and generates a minor alarm
.1.3.6.1.4.1.6027.3.1.1.4.0.21	chAlarmMinorFanBadClear	Trap generated when the minor alarm on the one or more fans is cleared
.1.3.6.1.4.1.6027.3.1.1.4.0.16	chAlarmFanTrayDown	Trap generated when all fans are down and/or when the fan tray status changes to missing or down
.1.3.6.1.4.1.6027.3.1.1.4.0.20	chAlarmFanTrayClear	Trap generated when all fans and/or the fan tray status changes to operational

Hardware watchdog timer

The hardware watchdog command automatically reboots an FTOS switch/router with a single RPM that is unresponsive. This is a last resort mechanism intended to prevent a manual power cycle.

Command	Description
hardware watchdog	Enable the hardware watchdog mechanism.

Offline diagnostics

Note: As the SFM on the C-Series is a logical concept only, the FORCE10-CHASSIS-MIB SFM-related SNMP alarms and traps are not used.

The offline diagnostics test suite is useful for isolating faults and debugging hardware.

Diagnostics are invoked from the FTOS CLI. While diagnostics are running, the status can be monitored via the CLI. The tests results are written to a file in flash memory and can be displayed on screen. Detailed statistics for all tests are collected. These statistics include:

- last execution time
- first and last test pass time
- first and last test failure time
- total run count
- total failure count
- consecutive failure count
- error code.

The diagnostics tests are grouped into three levels:

- **Level 0** checks the device inventory and verifies the existence of the devices (e.g., device ID test).
- **Level 1** verifies that the devices are accessible via designated paths (e.g., line integrity tests) and tests the internal parts (e.g., registers) of the devices.
- **Level 2** performs on-board loopback tests on various data paths (e.g., data port pipe and Ethernet).

Configuration task list

Note: This procedure assumes you have already loaded an FTOS image. These instructions illustrates the process of running offline diagnostics using line cards, but the process is the same for RPMs. Only the command keyword **linecard** must change to **rpm**. See the Command Line Reference Guide for details.

1. Take the line card offline. page 1185.
2. Run offline diagnostics. page 1185.
3. View offline diagnostic test results. page 1185.
4. Bring the line card back online. page 1188.

Important points to remember

- Offline diagnostics can only be run on offline line cards and on the standby route processor module (RPM). The primary RPM cannot be not tested.
- Diagnostics test only connectivity, not the entire data path.

- The complete diagnostics test suite normally runs for 4 to 6 minutes; the 48-port 1-Gigabit line card takes slightly longer than the 4-port 10-Gigabit line card.

Take the line card offline

Place the line card in an offline state using the **offline linecard** command, as shown in Figure 60-12.

Figure 60-12. offline linecard Command Example

```
FTOS#offline linecard 5
00:50:05: %RPM0-P:CP %CHMGR-2-CARD_DOWN: Line card 5 down - card offline
00:50:05: %RPM0-P:CP %IFMGR-1-DEL_PORT: Removed port: Te 5/0-3
```

Use the **show linecard all** command to confirm offline status, as shown in Figure 60-13.

Figure 60-13. show linecard all Command Example

```
FTOS#show linecard all
--  Line cards  --
Slot  Status        NxtBoot    ReqTyp   CurTyp   Version     Ports
--------------------------------------------------------------------------
  0   not present
  1   online        online     E48TB    E48TB    2.2.1.1     48
  2   not present
  3   not present
  4   not present
  5   not present
  6   offline       online     E48TB    E48TB    2.2.1.1     48
FTOS#
```

Run offline diagnostics

Start diagnostics on the line card using the **diag linecard** command, as shown in Figure 60-14.

Figure 60-14. diag linecard Command Example

```
FTOS#diag linecard 5
FTOS#00:50:44: %EX4PB:5 %DIAGAGT-6-DA_DIAG_STARTED: Starting diags on slot 5
00:50:44 : Approximate time to complete these Diags ... 5 Min
FTOS#
```

View offline diagnostic test results

Use the **show diag** command to view a brief report of the test results, as shown in Figure 60-15.

Figure 60-15. show diag linecard Command Example

```
FTOS#show diag linecard 5
Diag status of Linecard slot 5:
------------------------------------------------------------
  Card is currently offline.
  Card alllevels diag issued at THU FEB 08, 2018 04:10:06 PM.
  Current diag status:            Card diags are in progress.
------------------------------------------------------------
00:54:19 : Diagnostic test results are stored on file: flash:/TestReport
-LC-5.txt
00:54:19: %EX4PB:5 %DIAGAGT-6-DA_DIAG_DONE: Diags finished on slot 5
FTOS#
```

Use the **show file flash:***/filename* view the detailed test results in the test report saved to flash memory on the RPM. Use the command. Figure 60-16 shows the filename of the test results, and Figure 60-16 shows the contents of the file.

Note: Report any test failures to your Dell Force10 technical support engineer.

Figure 60-16. Viewing Offline Diagnostics Test Results

```
FTOS#show diag linecard 5
Diag status of Linecard slot 5:
----------------------------------------------------------------
    Card is currently offline.
    Card alllevels diag issued at THU FEB 08, 2018 04:10:05 PM.
    Current diag status:             Card diags are done.
    Duration of execution:           3 min 35 sec.
    Diagonostic test results located:        flash:/TestReport-LC-5.txt
----------------------------------------------------------------
*********************************C-Series Diagnostics*******************
LCM Board serial Number : 0060384
CPU Version : Line Processor: AMCC 440GX (rev D)
FPGA firmware Version : 1.20
Diag image based on build : CS-1-1-509
LCM Board Voltage levels - 3.260000 V, 2.480000 V, 1.770000 V, 1.500000 V, 1.200
000 V
LCM Board temperature : 29 Degree C
LCM present on Slot : 5.
*********************LCM EEPROM INFO*************************
********MFG INFO******************
Data in Lp Eeprom is listed...........
Vendor Id: 00
Country Code: 01
Date Code: 01012007
Serial Number: 0060384
Part Number: 1234
Product Revision: 1
Product Order Number: LC-CB-10GE-4P
Card Id: 402
*********SW INFO******************
Data in Lp Eeprom is listed...........
Chassis Type: 6
Chassis Mode: 4
Backplane version: 1
******************** Starting iteration 1. ********************
******************* LEVEL 0 DIAGNOSTICS *********************
Test 1 - NVRAM Access test ..................................... PASS
Test 3 - FPGA Access Test ...................................... PASS
Test 4 - Probing Test for volt/Temp sensor ...................... PASS
Test 6 - Probing for POE device 1 ...............................    NOT APPL
Test 7 - Probing for POE device 2 ...............................    NOT APPL
Test 8 - Probing for POE device 3 ...............................    NOT APPL
Test 9 - Probing for POE device 4 ...............................    NOT APPL
Test 10 - EEPROM access test ................................... PASS
Test 11 - PCI CPU BRG 0 Level0 Test ............................ PASS
Test 12 - PCI F10 DEV 0 Level0 Test ............................ PASS
Test 13 - PCI F10 DEV 1 Level0 Test ............................ PASS
Test 14 - PCI F10 DEV 2 Level0 Test ............................    NOT APPL
Test 15 - PCI F10 DEV 3 Level0 Test ............................    NOT APPL
Test 16 - PCI F10 DEV 4 Level0 Test ............................ PASS
Test 17 - PCI F10 DEV 5 Level0 Test ............................ PASS
Test 21 - PHY MGT DEV 0 Level0 Test ............................ PASS
Test 22 - PHY IPC DEV 0 Level0 Test ............................ PASS
Test 23 - PHY IPC DEV 1 Level0 Test ............................ PASS
Test 26 - FPGA Flash Primary Test ..............................     PASS
Test 27 - FPGA Firmware Compare Test ...........................     PASS
*************** LEVEL 1 DIAGNOSTICS*****************************
Test 100 - Performing SDRAM ECC test ........................... PASS
Test 101 - SDRAM Pseudo Random test ............................ PASS
```

Figure 60-17. Viewing Offline Diagnostics Test Results (continued)

```
Test 107 - NVRAM Address Line test ............................... PASS
Test 108 - NVRAM Data Line Test .................................. PASS
Test 110 - NVRAM Read Write test ................................. PASS
.Test 111 - FLASH Write Read test ................................ PASS
Test 112 - FPGA Registers Verification Test ...................... PASS
Test 113 - FPGA Level1 Test ...................................... PASS
Test 114 - FPGA Data bus walking 0 and 1's test ................... PASS
Test 115 - Temp/volt monitor write read test ...................... PASS
Test 116 - Reg verification test POE manager 1 .....    NOT APPL
Test 117 - Reg verification test POE manager 2 .....    NOT APPL
Test 118 - Reg verification test POE manager 3 .....    NOT APPL
Test 119 - Reg verification test POE manager 4 .....    NOT APPL
Test 120 - EEPROM write read test ................................ PASS
Test 122 - Blinking Status LEDs Test ............................. PASS
Test 123 - PHY MGT DEV 0 Level1 Test ............................. PASS
Test 124 - PHY IPC DEV 2 level1 Test ............................. PASS
Test 125 - PHY IPC DEV 3 level1 Test ............................. PASS
Test 126 - Local Eeprom MFG block checksum test ................... PASS
Test 127 - Local Eeprom SW block checksum test .................... PASS
Test 128 - 3.3 V Brick Load test ................................. PASS
Test 129 - 2.5 V Brick Load test ................................. PASS
Test 130 - 1.5 V Brick Load test ................................. PASS
Test 131 - 1.25 V Brick Load test ................................ PASS
Test 132 - 1.8 V Brick Load test ................................. PASS
Test 133 - XFP Verification Test 0 ............................... PASS
Test 134 - XFP Verification Test 1 ............................... PASS
Test 135 - XFP Verification Test 2 ............................... PASS
Test 136 - XFP Verification Test 3 ............................... PASS
Test 137 - XFP Verification Test 4 .....    NOT APPL
Test 138 - XFP Verification Test 5 .....    NOT APPL
Test 139 - XFP Verification Test 6 .....    NOT APPL
Test 140 - XFP Verification Test 7 .....    NOT APPL
************** LEVEL 2 DIAGNOSTICS*********************************
Test 200 - MAC MGT DEV 0 Level2 Test ............................. PASS
Test 201 - MAC IPC DEV 0 Level2 Test ............................. PASS
Test 202 - MAC IPC DEV 1 Level2 Test ............................. PASS
Test 203 - PHY MGT DEV 0 Level2 Test ............................. PASS
Test 204 - PHY IPC DEV 0 Level2 Test ............................. PASS
Test 205 - PHY IPC DEV 1 Level2 Test ............................. PASS
Test 216 - F10-SFM Port wise Traffic Test with PHY Loopback ....... PASS
Test 218 - LCM SNAKE Test using CPU Traffic with MAC Loopback ..... PASS
Test 219 - LCM SNAKE Test using CPU Traffic with PHY Loopback ..... PASS
Test 220 - LCM Port wise Traffic Test using CPU traffic - MAC Loopbac PASS
Test 221 - LCM Port wise Traffic Test using CPU traffic - PHY Loopbac PASS
Test 222 - POE I2C Interface stress test on Unit - 0 .....    NOT APPL
Test 223 - POE I2C Interface stress test on Unit - 1 .....    NOT APPL
Test 224 - POE I2C Interface stress test on Unit - 2 .....    NOT APPL
Test 225 - POE I2C Interface stress test on Unit - 3 .....    NOT APPL
*********** FORCE10 C series Diagnostics END*****************************
Number of Diagnostics performed 69
Number of Diagnostics failed 0
 End of Diags
 Duration of execution:        3 min 18 sec
```

Bring the line card online

Bring the card back online using the **online linecard** command. The card will be reset. Use the **show linecard all** command to verify the online status of the line card.

Buffer tuning

Buffer Tuning allows you to modify the way your switch allocates buffers from its available memory, and helps prevent packet drops during a temporary burst of traffic.

The C-Series and S-Series ASICs implement the key functions of queuing, feature lookups, and forwarding lookups in hardware.

- Forwarding Processor (FP) ASICs provide Ethernet MAC functions, queueing and buffering, as well as store feature and forwarding tables for hardware-based lookup and forwarding decisions. 1G and 10G interfaces use different FPs.
- Switch Fabric (CSF) ASICs are on the C-Series only. They provide some queuing while also providing the physical pathway through which frames are switched between ports when the source and destination ports are attached to different FP ASICs.

Table 60-5 describes the type and number of ASICs per platform.

Table 60-5. ASICS by Platform

Hardware	FP	CSF
48-port LC on C-Series	2	2

You can tune buffers at three locations, as shown in Figure 60-18.

1. CSF – Output queues going from the CSF.
2. FP Uplink—Output queues going from the FP to the CSF IDP links.
3. Front-End Link—Output queues going from the FP to the front-end PHY.

All ports support eight queues, 4 for data traffic and 4 for control traffic. All 8 queues are tunable.

Physical memory is organized into cells of 128 bytes. The cells are organized into two buffer pools—dedicated buffer and dynamic buffer.

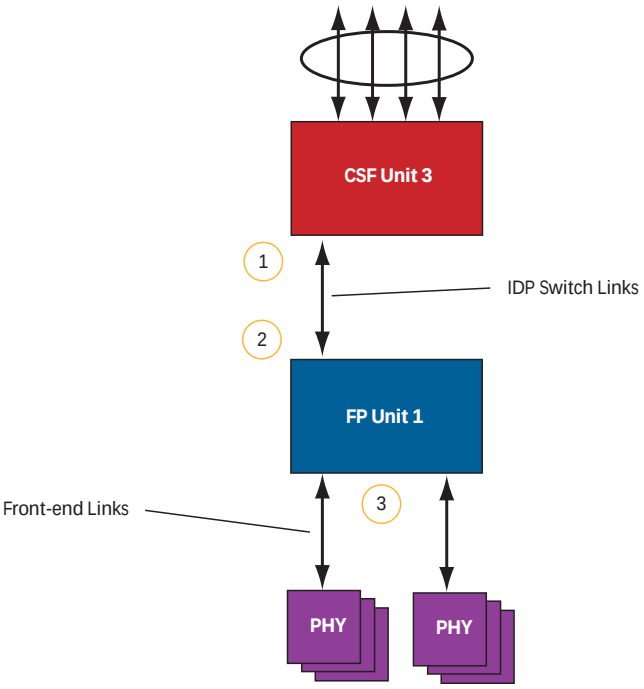
- **Dedicated buffer** is reserved memory that cannot be used by other interfaces on the same ASIC or by other queues on the same interface. This buffer is always allocated, and no dynamic recarving takes place based on changes in interface status. Dedicated buffers introduce a trade-off. They provide each interface with a guaranteed minimum buffer to prevent an overused and congested interface from starving all other interfaces. However, this minimum guarantee means the buffer manager does not reallocate the buffer to an adjacent congested interface, which means that in some cases, memory is underused.
- **Dynamic buffer** is shared memory that is allocated as needed, up to a configured limit. Using dynamic buffers provides the benefit of statistical buffer sharing. An interface requests dynamic buffers when its dedicated buffer pool is exhausted. The buffer manager grants the request based on three conditions:
  - the number of used and available dynamic buffers
  - the maximum number of cells that an interface can occupy

- Available packet pointers (2k per interface). Each packet is managed in the buffer using a unique packet pointer. Thus, each interface can manage up to 2k packets.

You can configure dynamic buffers per port on both 1G and 10G FPs and per queue on CSFs. By default, the FP dynamic buffer allocation is 10 times oversubscribed. For 48-port 1G card:

- Dynamic Pool= Total Available Pool(16384 cells) – Total Dedicated Pool =5904 cells
- Oversubscription ratio = 10
- Dynamic Cell Limit Per port= 59040/29 =2036 cells

Figure 60-18. Buffer Tuning Points



When to tune buffers

Dell Force10 recommends exercising caution when configuring any non-default buffer settings, as tuning can significantly affect system performance. The default values work for most cases.

As a guideline, consider tuning buffers if traffic is very bursty (and coming from several interfaces). In this case:

- Reduce the dedicated buffer on all queues/interfaces,
- Increase the dynamic buffer on all interfaces
- Increase the cell pointers on a queue that you are expecting will receive the largest number of packets.

Buffer tuning commands

Task	Command	Command Mode
Define a buffer profile for the FP queues.	**buffer-profile fp fsqueue**	CONFIGURATION
Define a buffer profile for the CSF queues.	**buffer-profile csf csqueue**	CONFIGURATION
Change the dedicated buffers on a physical 1G interface.	**buffer dedicated**	BUFFER PROFILE
Change the maximum amount of dynamic buffers an interface can request.	**buffer dynamic**	BUFFER PROFILE
Change the number of packet-pointers per queue.	**buffer packet-pointers**	BUFFER PROFILE
Apply the buffer profile to a line card.	**buffer fp-uplink linecard**	CONFIGURATION
Apply the buffer profile to a CSF to FP link.	**buffer csf linecard**	CONFIGURATION

FTOS Behavior: If you attempt to apply a buffer profile to a non-existent port-pipe, FTOS displays the following message. However, the configuration still appears in the running-config.

```
%DIFFSERV-2-DSA_BUFF_CARVING_INVALID_PORT_SET: Invalid FP port-set 2 for linecard 2. Valid
range of port-set is <0-1>
```

Configuration changes take effect immediately and appear in the running configuration. Since under normal conditions all ports do not require the maximum possible allocation, the configured dynamic allocations can exceed the actual amount of available memory; this is called oversubscription. If you choose to oversubscribe the dynamic allocation, a burst of traffic on one interface might prevent other interfaces from receiving the configured dynamic allocation, which causes packet loss.

You cannot allocate more than the available memory for the dedicated buffers. If the system determines that the sum of the configured dedicated buffers allocated to the queues is more than the total available memory, the configuration is rejected, returning a syslog message similar to the following.

Message 11 Buffer Allocation Error on C-Series

```
Mar 26 01:54:16: %E48VB:0 %DIFFSERV-2-DSA_DEVICE_BUFFER_UNAVAILABLE: Unable to allocate dedicated
buffers for linecard 0, port pipe 0, egress port 24 due to unavailability of cells
```

FTOS Behavior: When you move to a different chassis a line card that has a buffer profile applied at interface level on the fp-uplink, the line card retains the buffer profile. To return the line card to the default buffer profile, remove the current profile using the command **no buffer-profile fp-uplink linecard** from INTERFACE mode, and then reload the chassis.

FTOS Behavior: When you remove a buffer-profile using the command **no buffer-profile** [**fp** | **csf**] from CONFIGURATION mode, the buffer-profile name still appears in the output of **show buffer-profile** [**detail** | **summary**]. After a line card reset, the buffer profile correctly returns to the default values, but the profile name remains. Remove it from the **show buffer-profile** [**detail** | **summary**] command output by entering **no buffer** [**fp-uplink** |**csf**] **linecard port-set buffer-policy** from CONFIGURATION mode and **no buffer-policy** from INTERFACE mode.

Display the allocations for any buffer profile using the show commands in Figure 60-20. Display the default buffer profile using the command **show buffer-profile** {**summary** | **detail**} from EXEC Privilege mode, as shown in Figure 60-19.

Figure 60-19. Display the Default Buffer Profile

```
FTOS#show buffer-profile detail interface gigabitethernet 0/1
Interface Gi 0/1
Buffer-profile -
Dynamic buffer 194.88 (Kilobytes)
Queue#               Dedicated Buffer     Buffer Packets
                     (Kilobytes)
0                    2.50                 256
1                    2.50                 256
2                    2.50                 256
3                    2.50                 256
4                    9.38                 256
5                    9.38                 256
6                    9.38                 256
7                    9.38                 256
```

Figure 60-20. Displaying Buffer Profile Allocations

```
FTOS#show running-config interface tengigabitethernet 2/0 !
interface TenGigabitEthernet 2/0
no ip address
mtu 9252
switchport
no shutdown
buffer-policy myfsbufferprofile

FTOS#sho buffer-profile detail int gi 0/10
Interface Gi 0/10
Buffer-profile fsqueue-fp
Dynamic buffer 1256.00 (Kilobytes)
Queue#               Dedicated Buffer     Buffer Packets
                     (Kilobytes)
0                    3.00                 256
1                    3.00                 256
2                    3.00                 256
3                    3.00                 256
4                    3.00                 256
5                    3.00                 256
6                    3.00                 256
7                    3.00                 256

FTOS#sho buffer-profile detail fp-uplink stack-unit 0 port-set 0
Linecard 0 Port-set 0
Buffer-profile fsqueue-hig
Dynamic Buffer 1256.00 (Kilobytes)
Queue#               Dedicated Buffer     Buffer Packets
                     (Kilobytes)
0                    3.00                 256
1                    3.00                 256
2                    3.00                 256
3                    3.00                 256
4                    3.00                 256
5                    3.00                 256
6                    3.00                 256
7                    3.00                 256

FTOS#show buffer-profile detail csf linecard 2 port-set 0
Buffer-profile mybufferprofile
Queue#               Dedicated Buffer     Buffer Packets
                     (Bytes)
0                    80                   100
1                    160                  20
2                    240                  30
3                    400                  40
4                    7680                 255
5                    10240                255
6                    10240                255
7                    10240                255
```

Use a pre-defined buffer profile

FTOS provides two pre-defined buffer profiles, one for single queue (i.e non-QoS) applications, and one for four queue (i.e QoS) applications.

Task	Command Syntax	Command Mode
Apply one of two pre-defined buffer profiles for all port-pipes in the system.	**buffer-profile global** [**1Q**|**4Q**]	CONFIGURATION

You must reload the system for the global buffer-profile to take effect (Message 12).

Message 12 Reload After Applying Global Buffer Profile

```
% Info: For the global pre-defined buffer profile to take effect, please save the config and reload the
system.
```

FTOS Behavior: After you configure **buffer-profile global 1Q**, Message 12 is displayed during every bootup. Only one reboot is required for the configuration to take effect; afterwards this bootup message may be ignored.

FTOS Behavior: The buffer profile does not returned to the default, **4Q**, if you configure **1Q**, save the running-config to the startup-config, and then delete the startup-config and reload the chassis. The only way to return to the default buffer profile is to explicitly configure **4Q**, and then reload the chassis.

The **buffer-profile global** command fails if you have already applied a custom buffer-profile on an interface.

Message 13 Global Buffer Profile Error

```
% Error: User-defined buffer profile already applied. Failed to apply global pre-defined buffer profile.
Please remove all user-defined buffer profiles.
```

Similarly, when buffer-profile global is configured, you cannot not apply buffer-profile on any interface.

Message 14 Global Buffer Profile Error

```
% Error: Global pre-defined buffer profile already applied. Failed to apply user-defined buffer profile
on interface Gi 0/1. Please remove global pre-defined buffer profile.
```

If the default buffer-profile (4Q) is active, FTOS displays an error message instructing you to remove the default configuration using the command no buffer-profile global.

Sample configuration

The two general types of network environments are sustained data transfers and voice/data. Dell Force10 recommends a single-queue approach for data transfers. Figure 60-21 is a sample configuration for a C-Series 48-port line card that uses the default packet pointer values.

Figure 60-21. Single Queue Application With Default Packet Pointers

```
!
buffer-profile fp fsqueue-fp
buffer dedicated queue0 3 queue1 3 queue2 3 queue3 3 queue4 3 queue5 3 queue6 3 queue7 3
buffer dynamic 1256
!
 buffer-profile fp fsqueue-hig
 buffer dedicated queue0 3 queue1 3 queue2 3 queue3 3 queue4 3 queue5 3 queue6 3 queue7 3
 buffer dynamic 1256

!
buffer fp-uplink linecard 0 port-set 0 buffer-policy fsqueue-hig
buffer fp-uplink linecard 0 port-set 1 buffer-policy fsqueue-hig
!
Interface range gi 0/1 - 48
buffer-policy fsqueue-fp

FTOS#sho run int gi 0/10
!
interface GigabitEthernet 0/10
 no ip address
 switchport
 no shutdown
 buffer-policy fsqueue-fp
FTOS#
```

61

E-Series TeraScale Debugging and Diagnostics

This chapter addresses E-Series TeraScale Debugging and Diagnostics TeraScale platforms. Refer to Chapter 63, E-Series ExaScale Debugging and Diagnostics for information relating to that platform.

In addition to the FTOS high availability features, E-Series and FTOS support several diagnostics and debug features that are integral components to delivering maximum uptime. These features consist of the following:

Note: These diagnostics and debugability features are available on TeraScale systems only, unless specifically noted.

Overview

The FTOS diagnostics and debugging features are a proactive approach to maximizing system uptime and reducing meantime to resolution (MTTR) when a problem occurs. This feature set includes a combination of proactive and reactive components designed to alert the user to network events, automatically collect information on the event, and allow the user to collect diagnostic information from the system.

- Proactive component
  - The system health check detects, reports, and takes action on an error in real time.
  - When an automatic corrective action is not appropriate, the system health check reports the detected anomaly, in real time, via a syslog message and/or SNMP.
- Reactive component
  - When an error condition is asserted, appropriate show and debug commands are available to assist in identifying the condition as well as rapid fault isolation.

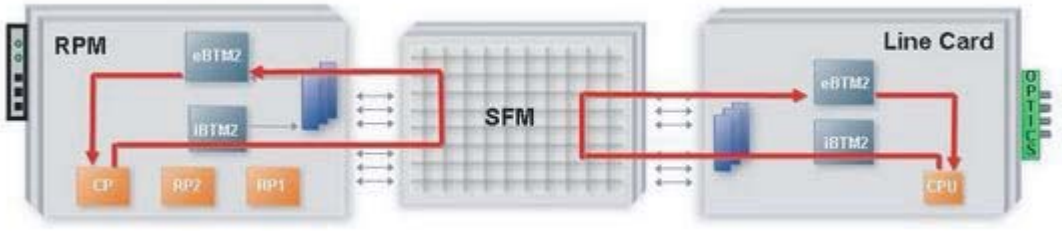
System health checks

An automatic runtime loopback test monitors the overall health status of the dataplane. This loopback test runs while the system's switch fabric is up; detecting potential blockages in the system's usual data transfer path.

Runtime dataplane loopback check

This is a dataplane loopback health check. Periodically, the primary RPM and each line card, in an online start, sends a packet through the dataplane channels, verifying the packet is returned, and then verifying the dataplane is functioning as expected (see Figure 61-1). Both portpipes on the line cards are tested.

Figure 61-1. Dataplane Loopback



If three consecutive packets are lost, an error message is logged and then one of the following happens:

- The RPM-SFM runtime loopback test failure initiates an SFM *walk* whenever it is enabled, feasible and necessary. The system automatically places each SFM (in sequential order) in an offline state, runs the loopback test, and then places the SFM back in an active state. This continues until the system determines a working SFM combination. If no working combination is found, the system restores to the pre-walking SFM state and the switch fabric state remains up. No more SFM walks are conducted as long as the SFM settings remain unchanged (setting changes include SFM reset, power off/on, and hotswap). However, the runtime loopback tests will continue with failure messages being logged every five minutes.

Note: SFM walking assumes a chassis with the maximum number of SFMs in an active state.

The loopback runtime test results reflect the overall health status of the dataplane. SFM walking can help to identify a single faulty SFM which is persistently dropping all traffic. For any partial packet loss, the loopback test results can only indicate that there is partial packet loss on the dataplane.

When an automatic SFM walk is conducted, events are logged to indicate the start and completion of the SFM walk and the results. A complete system message set is shown below.

FTOS Behavior: In very rare circumstances, FTOS is not able to recover from a SFM looback failure. You must recover manually from the loopback failure, by power-cycling the SFM. See Power the SFM on/off on page 1202.

Message 1 SFM walk message example

```
%TSM-2-RPM_LOOPBACK_FAIL: RPM-SFM dataplane loopback test failed

%TSM-2-SFM_WALK_START: Automatic SFM walk-through started

%TSM-6-RPM_LOOPBACK_PASS: RPM-SFM dataplane loopback test succeeded

%TSM-2-BAD_SFM_DISABLED: Bad SFM in slot 0 detected and disabled

%TSM-2-SFM_WALK_SUCCEED: Automatic SFM walk-through succeeded
```

- An SFM walk will not be able to identify multiple faulty SFMs, faulty linecards, or faulty RPM. In this case, the following event is logged.

Message 2 SFM walk Event Log

```
%TSM-2-RPM_LOOPBACK_FAIL: RPM-SFM dataplane loopback test failed

%TSM-2-SFM_WALK_START: Automatic SFM walk-through started

%TSM-2-SFM_WALK_FAIL: Automatic SFM walk-through failed to identify single faulty SFM
```

- If a line card runtime loopback test fails, the system does *not* launch an SFM walk. A message is logged indicating the failure.

Message 3 Loopback test failure

```
%TSM-2-RPM_LOOPBACK_FAIL: Linecard-SFM dataplane loopback test failed on linecard 6
```

The runtime dataplane loopback test is enabled by default. To disable this feature, use the following command.

Task	Command	Mode
Disable the runtime loopback test on the primary RPM and line cards. To re-enable, use the **no dataplane-diag disable loopback** command	**dataplane-diag disable loopback**	CONFIGURATION

Note: Disabling the runtime loopback test prevents the **sfm-walk** command and **sfm-bringdown** commands from taking effect.

Disable RPM-SFM walk

If a full set of SFMs are online during the runtime loopback test and an RPM-SFM runtime loopback test failure occurs, an automatic SFM walk is launched in an attempt to determine if the failure is due to a faulty SFM. If confirmed, the single faulty SFM is identified and disabled by default.

To disable the automatic SFM walk that is launched after an RPM-SFM runtime loopback test failure, use the following command in CONFIGURATION mode.

Task	Command	Mode
Disable the automatic SFM walk that is launched after an RPM-SFM runtime loopback test failure. To re-enable the automatic SFM walk, use the **no dataplane-diag disable sfm-walk** command.	**dataplane-diag disable sfm-walk**	CONFIGURATION

Note: Disabling the **sfm-walk** command prevents the **sfm-bringdown** command from taking effect.

RPM-SFM bring down

If a full set of SFMs are online during the runtime loopback test and a RPM-SFM runtime loopback test failure occurs, an automatic SFM walk is launched in an attempt to determine if the failure is due to a faulty SFM. If confirmed, the single SFM is identified and disabled (bringdown) by default.

To disable the automatic bring-down of an SFM that is identified by the SFM walk during the RPM-SFM runtime loopback test, use the following CONFIGURATION mode command.

Task	Command	Mode
Disable the automatic bring down of the single faulty SFM identified by the SFM walk during the RPM-SFM runtime loopback test. To re-enable the automatic bring down of an SFM, use the **no dataplane-diag disable sfm-bringdown** command.	**dataplane-diag disable sfm-bringdown**	CONFIGURATION

Manual loopback test

This manual dataplane loopback test is a supplemental test to the automatic runtime loopback test and can be initiated regardless if the runtime loopback test is enabled or disabled. Use this test to verify that the dataplane is actually functional *even* when a switch fabric status is down but there are at least (max-1) SFMs in active or diag failure state.

Task	Command	Mode
Execute a manual dataplane loopback test: • all-loopback – Both the RPM and the line card dataplane loopback test is done. • rpm-loopback – Only the RPM dataplane loopback test is done. This test can be run when the switch fabric is in either an operational or a non-operational state.	**diag sfm** [**all-loopback | rpm-loopback**]	EXEC

If the RPM-SFM or line card-SFM loopback test detects an SFM failure, an attempt is made to isolate a single faulty SFM by automatically *walking* the SFMs. For this failure case, error messages similar to the runtime loopback test error are generated.

Note: The dataplane runtime loopback configuration does not apply to this manual loopback test.

In the example in Figure 61-2, the manual loopback tests is successful, and no SFM failure is detected.

Figure 61-2. diag sfm all-loopback command Example

```
FTOS#diag sfm all-loopback
Proceed with dataplane loopback test [confirm yes/no]:yes
SFM loopback test completed successfully.

FTOS#
```

If the test passes when the switch fabric is down and there are at least (max-1) SFMs in the chassis, then the system will bring the switch fabric back up automatically. Like the runtime loopback test, the manual loopback test failure will not bring the switch fabric down.

Note: Line card-SFM loopback test failure, during the manual test, will trigger an SFM walk.

Power the SFM on/off

If you suspect that an SFM is faulty and would like to manually disable it to determine whether any packet loss or forwarding issues are resolved, execute the following command.

Task	Command	Mode
Power on or off a specific SFM.	**power-**{**off | on**} **sfm** *slot-number*	EXEC

Note: Execute this command only during an offline diagnostics; this command may bring down the switch fabric.

When there are a full set of SFMs online, powering down one SFM will reduce the total bandwidth supported by the chassis, and may affect data flow. A warning message is issued at the command line that requires user confirmation to proceed with the command (Figure 61-3).

Figure 61-3. power-off sfm command with data traffic warning message

```
FTOS#power-off sfm 0
SFM0 is active. Powering it off it might impact the data traffic.
Proceed with power-off [confirm yes/no]:yes
Feb 15 23:52:53: %RPM1-P:CP %CHMGR-2-MINOR_SFM: Minor alarm: only eight working SFM
FTOS#
```

Since this command is for diagnostic purposes, you can power off more than one SFM which may cause a switch fabric module to go down. A warning message is issued at the command line and requires user confirmation to proceed with the command (Figure 61-4).

Figure 61-4. power-off sfm command with switch fabric down warning message

```
FTOS#power-off sfm 1
WARNING!! SFM1 is active. Powering it off it will cause Switch Fabric to go down!!
Proceed with power-off [confirm yes/no]:yes
Feb 16 00:03:19: %RPM1-P:CP %TSM-6-SFM_SWITCHFAB_STATE: Switch Fabric: DOWN
Feb 16 00:03:20: %RPM1-P:CP %CHMGR-0-MAJOR_SFM: Major alarm: Switch fabric down
FTOS#
```

Once the SFM is powered off, the SFM status indicates that the SFM has been powered off by the user. Use the show sfm all command to display the status (Figure 61-5).

Figure 61-5. show sfm all command Example

```
FTOS#show sfm all
Switch Fabric State:  down   (Not enough working SFMs)
Switch Mode: SFM

--  Switch Fabric Modules  --
Slot  Status
------------------------------------------------------------------------------
  0   power off            (SFM powered off by user)
  1   power off            (SFM powered off by user)
  2   power off            (SFM powered off by user)
  3   active
  4   active
  5   active

FTOS#
```

Reset the SFM

When the SFM is taken offline due to an error condition, you can execute the **reset sfm** command and initiate a manual recovery.

Task	Command	Mode
Reset a specific SFM module (power-off and then power-on).	**reset sfm** *slot-number*	EXEC

When an error is detected on an SFM module, this command is a manual recovery mechanism. Since this command can be used with *live* traffic running, the switch fabric will not go down if the switch fabric is in an UP state. When there is a full set of SFMs online in the chassis, resetting one SFM will reduce the total bandwidth supported by the chassis and may effect data flow. A warning message is issued at the command line and requires user confirmation to proceed. (Figure 61-6)

Figure 61-6. reset sfm command example

```
FTOS#reset sfm 0
SFM0 is active. Resetting it might temporarily impact data traffic.
Proceed with reset [confirm yes/no]:yes
Feb 16 00:39:30: %RPM1-P:CP %TSM-5-SFM_DISCOVERY: Found SFM 0
FTOS#
```

This command does not permit resetting any SFM when the system has (max-1) SFM and switch fabric is up (Figure 61-7).

Figure 61-7. reset sfm error message

```
FTOS#FTOS#reset sfm 1
% Error: SFM1 is active. Resetting it will impact data traffic.
FTOS#
```

Note: Resetting an SFM in a power-off state is not permitted. Use the command **power-on sfm** to bring the SFM back to a power-on state.

SFM channel monitoring

In addition to monitoring the datapath, the SFM channels can be monitored using the Per-Channel Deskew FIFO Overflow (PCDFO) polling feature on all line cards and RPMs in both EtherScale and TeraScale E1200, E600, and E300 chassis. Like the datapath loopback feature, the PCDFO polling feature is enabled by default.

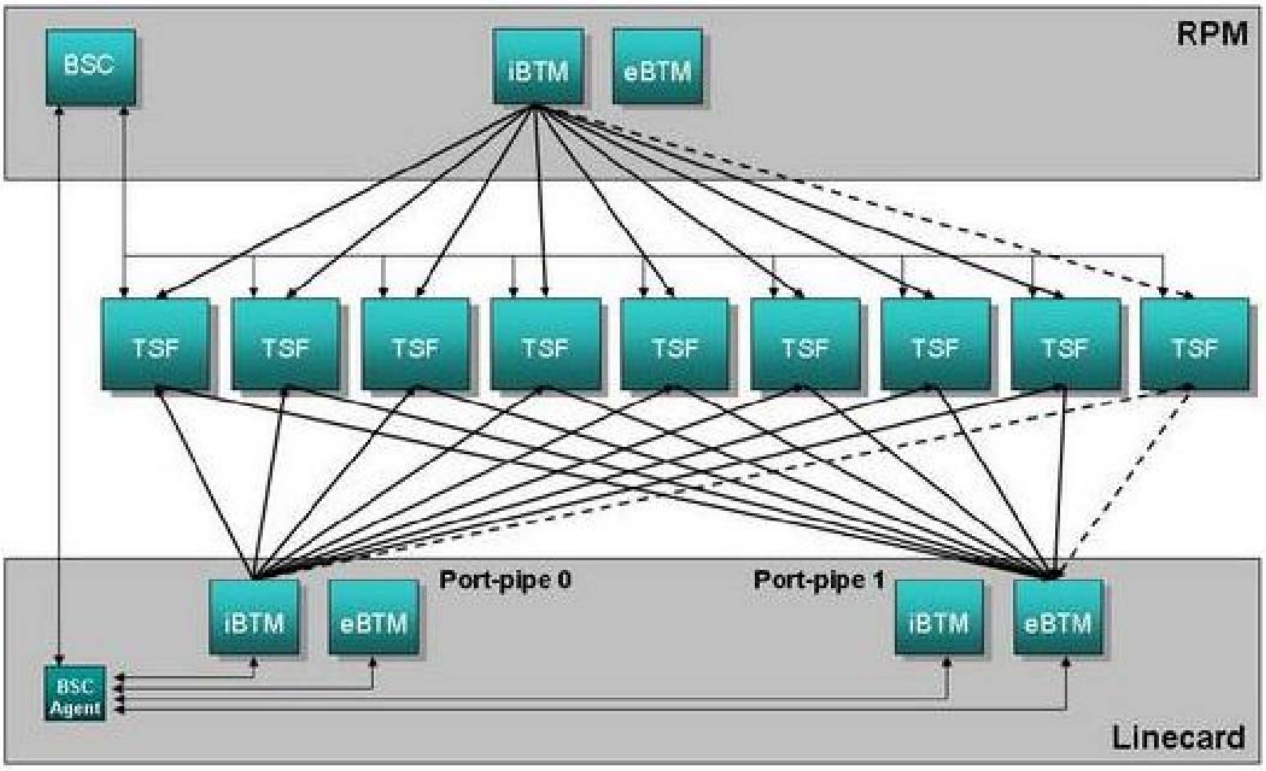
Note: This feature is not supported on the E600i chassis.

The PCDFO polling feature monitors data received over the switch fabric. When a DFO error is detected, no automatic action is initiated by the system. The message issued is similar to:

Message 4 PCDFO error message

```
%RPM1-P:CP %CHMGR-2-SFM_PCDFO: PCDFO error detected for SFM4
```

The following graphic illustrates the E600 and E1200 switch fabric architecture. Each ingress and egress Buffer and Traffic Management (BTM) ASIC maintains nine channel connections to the TeraScale Switch Fabric (TSF) ASIC.



Respond to PCDFO events

Troubleshooting PCDFO events requires applying some human intelligence to differentiate between transient and systematic failures. PCDFO events can be caused by several factors, including:

- Backplane noise
- Data corruption
- Bad epoch timing
- Mis-configuration of backplane

There are two PCDFO error types: Transient and Systematic. Transient error are non-persistent events that occur as one-events during normal operation. Systematic errors are repeatable events. For example, some hardware device or component is malfunctioning in such a way that it persistently exhibits incorrect behavior.

For the transient case, PCDFO errors are not reported to the log. The hardware system automatically recovers from the error state, and the dataplane continues to function properly. In persistent case, PCDFO errors will appear in the log, and the error state is likely to remain if not handled.

With PCDFO error data alone, it is impossible to arrive at a conclusion which will pinpoint the cause for PCDFO error or reason for packets drop. For example, it is quite possible to have multiple line cards/RPM show different channels with PCDFO error. Nonetheless, PCDFO status is a very useful data point as an indication of the health of the dataplane, particularly when an error is persistent.

To disable the PCDFO polling feature, use the following command in CONFIGURATION mode.

Task	Command	Mode
Disable the PCDFO polling feature. To re-enable, use the **no dataplane-diag disable dfo-reporting** command.	dataplane-diag disable dfo-reporting	CONFIGURATION

Detection of a PCDFO event causes the system to generate a message similar to the following.

Message 5 PCDFO error detection

```
%RPM1-P:CP %CHMGR-2-SFM_PCDFO: PCDFO error detected for SFM #
```

Events are logged when PCDFO error first occurs on any SFM and when PCDFO error pattern changes.

No automatic action is taken by the system when a DFO error is detected. If such an error is reported, note the SFM slot number identified in the message and contact Dell Force10 technical support. In addition, to confirm that the identified SFM needs to be replaced, use the diag sfm all-loopback to execute a manual dataplane loopback test.

Inter-CPU timeouts

Each RPM consists of three CPUs:

- Control Processor (CP)
- Routing Processor 1 (RP1)
- Routing Processor 2 (RP2)

The three CPUs use Fast Ethernet connections to communicate to each other and to the line card CPUs using Inter-Processor Communication (IPC). The CP monitors the health status of the other processors using heartbeat messaging exchange.

Message 6 CP monitor

```
%RPM1-P:CP %IPC-2-STATUS: target rp2 not responding
%RPM0-S:CP %RAM-6-FAILOVER_REQ: RPM failover request from active peer: Auto failover on failure
%RPM0-S:CP %RAM-6-ELECTION_ROLE: RPM0 is transitioning to Primary RPM.
%RPM0-P:CP %TSM-6-SFM_SWITCHFAB_STATE: Switch Fabric: UP
```

FTOS automatically saves critical information about the IPC failure to NVRAM. Such information includes:

- Status counters on the internal Ethernet interface
- Traffic profile of the inter-CPU bus
- Kernel drops
- High CPU exception conditions

Upon the next boot, this information is uploaded to a file in the CRASH_LOG directory. Use the following command sequence beginning in EXEC mode to capture this file for analysis by the Dell Force10 TAC.

Step	Task	Command	Mode
1	Display the directories in flash memory. The output should include: **1 drwx 2048 Jan 01 1980 00:00:06 CRASH_LOG_DIR**	dir flash:	EXEC
2	Change to the CRASH_LOG directory.	cd CRASH_LOG_DIR	EXEC
3	View any saved files in the CRASH_LOG directory. The naming convention is: *sysinfo_RPMIDProcessorID _ timestamp* For example: **sysinfo_RPM1CP_20060616_013125** **sysinfo_RPM1RP1_20060616_013248** **sysinfo_RPM1RP2_20060616_013249**	dir	EXEC Privilege
4	View the contents of the file.	show file flash://CRASH_LOG_DIR/**[file_name]**	EXEC Privilege

In a dual RPM system, the two RPMs send synchronization messages via inter-RPM communication (IRC). As described in the High Availability chapter, an RPM failover can be triggered by loss of the heartbeat (similar to a keepalive message) between the two RPMs. FTOS reports this condition via syslog messages, as follows:

Message 7 RPM heartbeat report

```
20:29:07: %RPM1-S:CP %IRC-4-IRC_WARNLINKDN: Keepalive packet 7 to peer RPM is lost
20:29:07: %RPM1-S:CP %IRC-4-IRC_COMMDOWN: Link to peer RPM is down
%RPM1-S:CP %RAM-4-MISSING_HB: Heartbeat lost with peer RPM. Auto failover on heart beat lost.
%RPM1-S:CP %RAM-6-ELECTION_ROLE: RPM1 is transitioning to Primary RPM.
```

FTOS automatically saves critical information, about the IRC failure, to NVRAM. Use the same three-step procedure to capture this file for analysis by Dell Force10.

FTOS actually saves up to three persistent files depending upon the type of failure. When reporting an RPM failover triggered by a loss of the IPC or IRC heartbeats, look for failure records in the following directories:

- Application or kernel core dump RP in the CORE_DUMP_DIR
- CP trace log file (look for a filename with the phrase “failure_trace”) in the TRACE_LOG_DIR
- RP and/or CP sysinfo file in the CRASH_LOG_DIR, as explained above

Debug commands

FTOS supports an extensive suite of debug commands for troubleshooting specific problems while working with Dell Force10 technical support staff. All debug commands are entered in privileged EXEC mode. See the *FTOS Command Reference* for details.

Hardware watchdog timer

The hardware watchdog command automatically reboots an FTOS switch/router with a single RPM that is unresponsive. This is a last resort mechanism intended to prevent a manual power cycle.

Command	Description
hardware watchdog	Enable the hardware watchdog mechanism.

Show hardware commands

The show hardware command tree consists of privileged EXEC commands created or changed specially for use with the E-Series. These commands display information from a hardware sub-component, such as the Buffer and Traffic Management (BTM) ASIC and the Forwarding and Packet Classification (FPC) ASIC. They should be used only under the guidance of Dell Force10 technical support staff.

The following table lists the show hardware commands. For detailed information on these and other commands, see the *FTOS Command Line Interface Reference* document.

Command	Description
show hardware rpm *slot-number* **mac counters** [**port** *port-number*] **clear hardware rpm** *slot-number* **mac counters**	View or clear the receive- and transmit- counters for the party-bus control switch on the IPC subsystem of the RPM.
show hardware rpm *slot-number* **cp {data-plane | management-port}** | **party-bus}** {**counters | statistics**} **show hardware rpm** *slot-number* {**rp1** | **rp2**} **{data-plane | party-bus}** {**counters | statistics**}	Display advanced debugging information for the RPM processors.
show hardware linecard *number* **port-set** *pipe-number* **fpc forward {counters | drops | spi {err-counters | spichannel# counters} | status}**	Display receive and transmit counters, error counters and status registers for the forwarding functional area of the FPC (flexible packet classification engine).
show hardware linecard *number* **port-set** *pipe-number* **fpc lookup detail**	Display diagnostic and debug information related to the lookup functional area of the Flexible Packet Classification (FPC).
show running-config hardware-monitor	Display hardware-monitor action-on-error settings.
show cpu-interface-stats	Provides an immediate snapshot of internal RPM and line card CPU health counters.

Offline diagnostics

These diagnostics can be useful for isolating faults and debugging TeraScale hardware installed in a chassis.

Diagnostics are invoked from the FTOS CLI. While diagnostics are running, the status can be monitored via the CLI. The tests results are written to a file in flash memory and can be displayed on screen. Detailed statistics for all tests are collected and include:

- last execution time
- first test pass time and last test pass time
- first test failure time and last test failure time

- total run count
- total failure count
- consecutive failure count
- error code

The diagnostics tests are grouped into three levels:

Level 0—Check the inventory of devices. Verify the existence of devices (e.g., device ID test).

Level 1—Verify the devices are accessible via designated paths (e.g., line integrity tests). Test the internal parts (e.g., registers) of devices.

Level 2—Perform on-board loopback tests on various data paths (e.g., data port-pipe and Ethernet).

Important points to remember

- Offline diagnostics can be run only on an offline line card and on a standby route processor module (RPM). The primary RPM is not tested.
- Diagnostics test only connectivity and not the entire data path.
- A line card must be put into an offline state before diagnostics are run.
- Complete diagnostics test suite normally runs for 5 to 7 minutes on a single port-pipe line card and 12 to 15 minutes on a dual port-pipe line card. Running diagnostics on LC-EF-GE-90M cards may take slightly longer.

Offline configuration task list

Use the following steps to run offline diagnostics on the E-Series. This procedure assumes the FTOS image is installed.

1. Place the line card in an offline state with the **offline linecard** command. Use the **show linecard** command to confirm the new status.

```
Foce10#offline line 4
Mar 27 05:18:26: %RPM0-P:CP %CHMGR-2-CARD_DOWN: Line card 4 down - card offline
Mar 27 05:18:26: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 4/3
```

2. Start diagnostics on the line card with the **diag** command. The system will confirm that diagnostics tests are running by displaying the syslog message shown below.

```
FTOS#diag linecard 4 ?
alllevels               Execute level 0-2 diags (default)
level0                  Execute level 0 diags
level1                  Execute level 1 diags
level2                  Execute level 2 diags
terminate               Stops the running test

FTOS#diag linecard 4
Mar 27 01:54:00: %E12PD3:2 %DIAGAGT-6-DA_DIAG_STARTED: Starting diags on slot 4
Mar 27 02:05:47: %E12PD3:2 %DIAGAGT-6-DA_DIAG_DONE: Diags finished on slot 4
```

3. Execute the **show diag** command to view a report of the test results.

```
FTOS#show diag linecard 4
Diag status of Linecard slot 4:
------------------------------------------------------------------
    Card is currently offline.
    Card level0 diag issued at TUE Mar 27, 2007 05:19:35 AM.
    Current diag status:             Card diags are done (FAIL).
    Duration of execution:           0 min 0 sec.
    Number of diags performed:       39
    Number of diags passed:          36
    Number of diags failed:          3
    Number of notification received: 80
    Last notification received at:   TUE Mar 27, 2007 05:19:35 AM
```

4. Report any test failures to your Dell Force10 technical support engineer.
5. Bring the card back online with the **online linecard** {*slot#*} command. The card will be reset.

Parity error detection and correction

There are two types of parity errors: transient and real.

- **Transient Parity Error**— implies that a read value was corrupted in transit but that the actual memory may not be corrupt. Transient errors are further categorized as a recoverable and phantom.
  - **Recoverable Transient Parity Error**—a transient parity error indicated by SRAM that FTOS was able to correct (rewrite).
  - **Phantom Transient Parity Error**—a recoverable parity error for which the software was unable to determine a problem location.
- **Non-recoverable Error**—implies a persistent corrupted memory location for which the only means of recovery is to reboot the line card.

FTOS has the ability determine the error type when a parity error occurs and correct recoverable transient parity errors using the Parity Error Correction feature. During the SRAM scanning function, anytime the system detects a parity error, it must determine which type it is. To distinguish between the two types of parity errors (transient and non-recoverable), the system maintains a copy of all SRAM writes. If a location does not match the SRAM copy or causes another parity error indication in the status register, the system rewrites the location. After rewriting the location, the system again reads the location and checks the status register for parity error indication. If the location fails either of these two tests after a rewrite, then the parity error is non-recoverable, the location is marked as corrupt, and FTOS generates log messages. If the location passes these tests, then the parity error is transient (recoverable), the location is marked as having had a transient parity error, and FTOS continues the scan to the end of the SRAM bank.

Enable parity error correction

Parity Error Correction is disabled by default and consumes 25 megabytes of line card memory when you you enable it.

To enable Parity Error Correction:

Step	Task	Command	Command Mode
1	Verify that the line card has sufficient memory to enable this feature, as shown in Figure 61-8.	**show processes memory lp**	EXEC Privilege
2	Enable Parity Error Correction	**hardware monitor linecard asic fpc parity-correction**	CONFIGURATION
3	Reload the linecard.	**reset linecard**	EXEC Privilege

FTOS displays Message 8 on the console, when you enable Parity Error Correction. FTOS also records the messages in Message 9 in the trace log when you enable and disable Parity Error Correction, or insert a line card with Parity Error Correction enabled.

Message 8 Parity Error Correction Enabled

```
%RPM0-P: CP %CHMGR-5-PARITY_CORRECTION: FPC parity correction feature will be on next reload.
```

Message 9 Parity Error Correction Enabled

```
[1/1 0:0:25] LCMGR-(lcMgr):lcMgrSetParityCorrection(): Enable parity correction
[5/15 16:13:14] LCMGR-(lcMgr):lcMgrSetParityCorrection(): Disable parity correction
```

Figure 61-8. Viewing the Available Memory on a Line Card

```
FTOS#show processes memory lp 9
Total:   197534888, MaxUsed:   112999116, CurrentUsed:   87782876, CurrentFree:
109752012
    TaskName  TotalAllocated       TotalFreed       MaxHeld    CurrentHolding

 ppdT2lSramP        25169040       25165840      25169040            3200
```

Recognize a transient parity error

FTOS generates a message similar to Message 10 and Message 11 when it encounters a phantom or recoverable transient parity error.

Message 10 Console Phantom Transient Parity Error Message

```
Apr 29 18:01:24: %EXW16PG:6 %POLLMGR-5-FPC_NOTIFY: Line card detected FPC 6 parity - Transient phantom
```

Message 11 Console Recoverable Transient Parity Error Message

```
Apr 29 18:01:05: %EXW16PG:6 %POLLMGR-5-FPC_NOTIFY: Line card detected FPC 6 parity - Transient
recoverable
```

The line card status does not reflect transient errors until FTOS encounters five recoverable or 50 phantom transient errors on a card within an hour, as shown in Figure 61-9. The text “Last Event” indicates the last type of parity error (transient or real) that occurred.

Use SNMP to poll the number of transient errors using the objects chSysCardParityPhantomError and chSysCardParityRecoverableError. If it has been more than one hour since the last occurrence of the same type of error, the counter associated with that type of error is reset to 1.

Figure 61-9. Viewing the Parity Status of a Line Card

```
FTOS#show linecard 6

--  Line card 6 --
Status        : online
Next Boot     : online
Required Type : E48TF - 48-port 10/100/1000Base-T line card with RJ-45 interfaces (EF)
Current Type  : E48TF - 48-port 10/100/1000Base-T line card with RJ-45 interfaces (EF)
Hardware Rev  : Base  - 1.1  PP0 - 1.0  PP1 - 1.0
Num Ports     : 48
Up Time       : 3 min, 44 sec
FTOS Version  : 6.5.4.1
Jumbo Capable : yes
Boot Flash    : A: 2.3.1.3     B: 2.3.1.3 [booted]
Memory Size   : 268435456 bytes
Temperature   : 42C
Power Status  : AC
Voltage       : ok
Serial Number : 0045149
Part Number   : 7520016602 Rev 06
Vendor Id     : 04
Date Code     : 01442005
Country Code  : 01
Parity Status : Last Event - FPC DDR Bank [191:128], Transient Failure, Running Count 5
FTOS#show linecard 7 | find “Parity Status”
Parity Status : Last Event - FPC DDR Bank [191:128], Real Failure, Address 0x11ffe00
```

Recognize a non-recoverable parity error

FTOS generates a message similar to Message 12 when it encounters a real transient parity error.

Message 12 Console Real Parity Error Message

```
Apr 29 18:52:50: %RPM0-P:CP %CHMGR-2-CARD_PARITY_ERR: Linecard 6 pp 0 FPC SRAM Hard parity error:
Address 0x85000004 Index 0x80000
```

FTOS also generates an SNMP trap, TR_CHM_SRAM_PARITY_NONRECOVERABLE, at the same time it generates a console message for non-recoverable parity errors. Use SNMP to poll the number of non-recoverable errors using the objects chSysCardParityNonrecovrableError.

Trace logs

In addition to the syslog buffer, FTOS buffers trace messages which are continuously written by various FTOS software tasks to report hardware and software events and status information. Each trace message provides the date, time, and name of the FTOS process. All messages are stored in a ring buffer and can be saved to a file either manually or automatically upon failover.

Some trace files are automatically saved and stored in the **flash:/TRACE_LOG_DIR** directory for SW and HW Traces of the CP and for all Linecards. This directory contains the **TRACE_CURR_BOOT** file which in turn contains the saved trace buffer files.

The **TRACE_LOG_DIR/TRACE_CURR_BOOT** files can be reached by FTP or by using the **show file** command from the **flash://TRACE_LOG_DIR** directory.

Note: At reload this directory is renamed to **flash:/TRACE_LOG_DIR/TRACE_LAST_BOOT** and a new empty **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT** directory is created.

Buffer full condition

When the Trace Ring Buffer fills up, trace logs are saved into the flash so the buffer can be clear for further trace activity. The saved file is named *hw_trace_RPM0CP.0*, for example. If the buffer fills a second time, a second file is created as *hw_trace_RPM0CP.1* and saved to the flash. Following the fifth file created (*hw_trace_RPM0CP.4*), the saved files are overwritten starting with the .1 version (*hw_trace_RPM0CP.1*).

These files will be saved in **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT**. At reload this directory is renamed as **flash:/TRACE_LOG_DIR/TRACE_LAST_BOOT** and an empty **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT** directory is created.

Trace file *hw_trace_RPM0CP.0* is not overwritten so that chassis bootup message are preserved.

The CP and LP trace file names are:

Note: Line card 1 is taken as the example for the following filenames.

- **CP [SW trace]** : sw_trace_RPM0CP.0, sw_trace_RPM0CP.1, sw_trace_RPM0CP.2, sw_trace_RPM0CP.3 and sw_trace_RPM0CP.4
- **CP [HW trace]** : hw_trace_RPM0CP.0, hw_trace_RPM0CP.1, hw_trace_RPM0CP.2, hw_trace_RPM0CP.3 and hw_trace_RPM0CP.4
- **LP [SW trace]** : sw_trace_LPX.0, sw_trace_LP1.1, sw_trace_LP1.2, sw_trace_LP1.3 and sw_trace_LP1.4
- **LP [HW trace]** : hw_trace_LPX.0, hw_trace_LP1.1, hw_trace_LP1.2, hw_trace_LP1.3 and hw_trace_LP1.4

Trace files are saved in the directory **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT**. Upon a system reload this directory is renamed **flash:/TRACE_LOG_DIR/TRACE_LAST_BOOT**, and an empty **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT** directory is created.

Manual reload condition

When the chassis is reloaded manually (through the CLI), trace messages in all of the buffers (software and hardware) in CP and linecards are saved to the flash as *reload_traceRPM0_CP* and *reload_traceLP1* in **flash:/TRACE_LOG_DIR/TRACE_CURR_BOOT**. After reload, you can see these files in **flash:/TRACE_LOG_DIR/TRACE_LAST_BOO***T*.

When the trace messages are being saved on reload, Message 13 is displayed.

Message 13 Saving Trace Messages

```
Starting to save trace messages… Done.
```

The CP and LP trace file names at chassis reload are:

- **CP:** reload_traceRPM0_CP
- **LP:** reload_traceLP1

CP software exceptions

When a RPM resets due to a software exception, the linecard trace files are saved to **flash:/TRACE_LOG_DIR** directory.

The CP and LP trace file names in the case of a software exception are:

- **CP:** failure_trace_RPM1_CP
- **LP:** failure_trace_RPM1_LP1

For systems with a single RPM, the linecard traces are saved on the failed RPM itself.

For systems with dual RPM, linecard trace logs are saved when the CP, RP1, or RP2 crashes. The linecard trace logs are saved on the new Primary RPM. The linecard trace file name identifies the failed RPM. For example, if RPM0 fails the trace files saved in RPM1 with filename as **failure_trace_RPM0_LP1**.

View trace buffer content

The command-history trace feature captures all commands entered by all users of the system with a time stamp and writes these messages to a dedicated trace log buffer. The system generates a trace message for each executed command. No password information is saved to the file.

To view the command-history trace, use the **show command-history** command, as shown in Figure 61-10.

Figure 61-10. show command-history Command Example

```
FTOS#show command-history
[12/5 10:57:8]: CMD-(CLI):service password-encryption
[12/5 10:57:12]: CMD-(CLI):hostname Force10
[12/5 10:57:12]: CMD-(CLI):ip telnet server enable
[12/5 10:57:12]: CMD-(CLI):line console 0
[12/5 10:57:12]: CMD-(CLI):line vty 0 9
[12/5 10:57:13]: CMD-(CLI):boot system rpm0 primary flash://FTOS-CB-1.1.1.2E2.bin
```

Write the contents of the trace buffer

The trace logs are saved to automatically but you can save the contents of a buffer manually via the CLI.

To manually write the contents of a trace buffer on CP to a file on the flash:

Step	Task	Command Syntax	Command Mode
1	Write the buffered trace log to flash.	**upload trace-log cp** [**cmd-history | hw-trace | sw-trace**]	EXEC Privilege

To manually write the contents of a trace buffer on LP to a file on the flash:

Step	Task	Command Syntax	Command Mode
1	Write the buffered trace log to flash.	**upload trace-log** [**rp1 | rp2 | linecard**] *number* [**hw-trace** | **sw-trace**]	EXEC Privilege

Clear the trace buffer

Clear the command history buffer using the command clear command-history from EXEC Privilege mode, as shown in Figure 61-11.

Figure 61-11. Clearing the Command History

```
FTOS#show command-history 10
[12/3 15:40:17]: CMD-(CLI):[show config]by default from console
[12/3 15:40:22]: CMD-(CLI):[ping 10.11.80.201]by default from console
[12/3 15:40:46]: CMD-(CLI):[show interfaces managementethernet 0/0]by default from console
[12/3 15:40:49]: CMD-(CLI):[shutdown]by default from console
[12/3 15:40:59]: CMD-(CLI):[no shutdown]by default from console
[12/3 15:41:1]: CMD-(CLI):[interface managementethernet 0/0]by default from console
[12/3 15:41:2]: CMD-(CLI):[shutdown]by default from console
[12/3 15:41:7]: CMD-(CLI):[ping 10.11.80.201]by default from console
[12/3 21:45:46]: CMD-(CLI):[enable]by default from console
[12/3 21:47:18]: CMD-(CLI):[show command-history 10]by default from console
FTOS#clear command-history
FTOS#show command-history 10
[12/3 21:47:43]: CMD-(CLI):[show command-history 10]by default from console
FTOS#
```

Recognize a high CPU condition

A high CPU condition exists when any of the messages in Message 14 appear.

Message 14 High CPU Condition

```
Feb 13 13:56:16: %RPM1-S:CP %CHMGR-5-TASK_CPU_THRESHOLD: Cpu usage above threshold for task
"sysAdmTsk"(100.00%) in CP.


Feb 13 13:56:20: %RPM1-S:CP %CHMGR-5-CPU_THRESHOLD: Overall cp cpu usage above threshold.
Cpu5SecUsage (100)


Feb 13 13:56:20: %RPM1-S:CP %CHMGR-5-TASK_CPU_THRESHOLD_CLR: Cpu usage drops below threshold
for task "sysAdmTsk"(0.00%) in CP.
```

Configure an action upon a hardware error

You can configure FTOS to take an action if it encounters an BTM, FPC, or MAC hardware error.

Buffer traffic manager hardware errors

FTOS displays Message 15, Message 16, Message 17, or Message 18 depending on the type of BTM error. In this case, configure an action using the command **hardware monitor linecard asic btm action-on-error**. You may place the line card in a problem state, reset the card, or shutdown all ports on the card.

Message 15 Queue Memory Error

```
%RPM0-P:CP %CHMGR-1-QMERR_RDBE: (0x30004) Double bit error detected in SRAM pointer memory on
Ingress BTM port-pipe 0 Line card slot 9.

%RPM0-P:CP %CHMGR-1-QMERR_RDBE: (0x30004) Double bit error detected in SRAM pointer memory on
Egress  BTM port-pipe 0 Line card slot 9.
```

Message 16 Buffer Memory Error

```
RPM0-P:CP %CHMGR-3-BM_STATUS_DBE: Double-bit error detected when reading internal header from
buffer memory on Ingress BTM port-pipe 0 Line card slot 9. Check for temporary or sustained
packet loss with show hardware commands.

%RPM0-P:CP %CHMGR-3-BM_STATUS_DBE: Double-bit error detected when reading internal header
from buffer memory on Egress  BTM port-pipe 0 Line card slot 9. Check for temporary or
sustained packet loss with show hardware commands.
```

Message 17 Low Free Memory Error

```
%RPM0-P:CP %CHMGR-1-LF_MEM_ERR: Low free memory error detected on Ingress BTM port-pipe 0
Line card slot 9

%RPM0-P:CP %CHMGR-1-LF_MEM_ERR: Low free memory error detected on Egress BTM port-pipe 1 Line
card slot 9
```

Message 18 Start-of-Packet/End-of-Packet Memory Error

```
%RPM0-P:CP %CHMGR-1-SOP_EOP_ERR: SOP/EOP error detected on Ingress BTM port-pipe 1 Line card
slot 9

%RPM0-P:CP %CHMGR-1-SOP_EOP_ERR: SOP/EOP error detected on Unknown slot 9
```

Flexible packet classifier hardware errors

FTOS displays Message 19 in case of a parity error on an FPC. Configure an action using the command **hardware monitor linecard asic fpc action-on-error**. You may place the line card in a problem state, reset the card, or shutdown all ports on the card.

Message 19 Parity Error

```
%RPM0-P:CP %CHMGR-2-CARD_PARITY_ERR: Linecard 9 pp 0 FPC SRAM Hard parity error:  Address
0x85000 Index 0x80000 Check the Hardware Log

%E48TF:9 %POLLMGR-5-FPC_NOTIFY: Line card detected FPC 9 parity - Transient phantom

%E48TF:9 %POLLMGR-5-FPC_NOTIFY: Line card detected FPC 9 parity - Transient recoverable
```

Line card MAC hardware errors

FTOS displays Message 20 in case of a port hang error. Configure an action using the command **hardware monitor mac action-on-error port-shutdown**. You may only shutdown all ports on the card with this command.

Message 20 Port Hang Error

```
%E48TF:9 %IFAGT-5-PORT_HUNG: Port hang detected on slot 9 port 2

%E48TF:9 %IFAGT-5-PORT_HUNG: Port hang detected on slot 9 port 26
```

Core dumps

RPM core dumps

The RPM supports two types of core dumps for RPMs—application and kernel.

Kernel core dump—The E-Series supports kernel core dumps for CP and for RP1/RP2 using a naming convention of *f10{cp|rp{1|2}}.kcore.gz*.

RP kernel core dumps are enabled by default. New files are written in flash until space is exhausted, in which case the write is aborted.

CP kernel core dumps are disabled by default. Enable them using the command **logging coredump cp** from CONFIGURATION mode. If you use the keyword **cp** with this command, the system creates a file, named f10cp.kcore.gz, that preserves space on the internal flash so that there is always enough space for a core dump. Undoing this command using the **no logging coredump cp** removes this file. The CP kernel core dumps are overwritten every time there is a new core dump. You should manually upload kernel core dumps periodically if you want to save this information.

You may choose to write the core dump directly to an FTP server using the keyword **server**. However, the server option supports only RP coredumps; it does not support CP coredumps. By default the kernel core dump is sent to the root directory of the internal flash CP and the CORE_DUMP_DIR directory for RP.

Application core dump—On the E-Series, the application core dump has the file name format *f10{cp|rp{1|2}}<yymmddhhmmss>.acore.gz*, where <*yymmddhhmmss*> is a time stamp, and FTOS writes it to the internal flash.

Note: The kernel core dump can be large and may take up to 5 to 30 minutes to upload. FTOS does not overwrite application core dumps, so you should delete them as necessary to conserve space on the flash; if the flash is out of memory, the core dump is aborted. On the S-Series, if the FTP server is not reachable, the application core dump is aborted.

The FTOS High Availability module is aware of the core dump upload, and it does not reboot the crashed RPM until the core dump has completed or is aborted.

Line card core dumps

Line card core dumps preserve critical status information for cards which experience a task crash.

Message 21 Task Crash Detection

```
%E48TF:0 %TME-2-TASK SUSPENDED: LINECARD TASK SUSPENDED. SAVING FAILURE  RECORD IN PROGRESS
```

Writing a line card core dump to memory on the RPM requires 5 to 30 minutes. During this time, the physical interfaces remain in their current state, which is normally operationally up. This behavior assumes that most FTOS agent tasks running on the line card CPU continue to operate correctly and that you want to maximize uptime by having packets continue to flow while FTOS writes the core dump.

If you want to failover to a redundant system when a line card exception occurs, use the command **logging coredump linecard port-shutdown** (Figure 663) to shut down ports during a core dump so that the backup system can take over.

Line card core dumps are disabled by default. To enable line card core dumps and specify the shutdown mode:

Step	Task	Command Syntax	Command Mode
1	Enable line card core dumps and specify the shutdown mode.	**logging coredump linecard** {*all* | {*0-13*}} [**port-shutdown** | **no-port-shutdown**]	EXEC Privilege

Note: In the absence of port-shutdown and no-port-shutdown, the option no-port-shutdown is applied.

Once the core dump file has been created with the **logging coredump** command, the file can be deleted from the Standby RPM flash although the space is not released. The CP kernel core dump file space cannot be recovered by deleting the files. You must format the flash drive to recover the space.

Message 22 Internal File Transfer of Core Dump to the RPM Complete

```
%E48TF:0 %TME-2-TASK SUSPENDED: SAVING FAILURE RECORD COMPLETED
```

To locate the line card core dump file:

Step	Task	Syntax	Command Mode
1	Change the directory to CORE_DUMP_DIR.	**cd CORE_ DUMP_DIR**	EXEC Privilege
2	View the files in the directory.	**dir**	EXEC Privilege

Figure 61-12. Display the core dump directory.

```
FTOS#dir
Directory of flash:/CORE_DUMP_DIR
1  drwx       8192   May 12 2025 03:39:54 +00:00
2  drwx      32768   Jan 01 1980 00:00:00 +00:00
3  -r-x  134217728   Jul 21 2008 22:45:12 +00:00 f10cp.kcore.gz
5  -rwx      16384   Dec 03 2008 16:47:52 +00:00 f10cp.kcore.mini.txt
6  -rwx  154555270   Nov 20 2008 18:57:38 +00:00 f10lp13.core.gz
Line card core dumps use a file naming convention of f10lp{slot#}.core.gz.
```

62

S-Series Debugging and Diagnostics

The chapter contains the following major sections:

Offline diagnostics

The offline diagnostics test suite is useful for isolating faults and debugging hardware.The diagnostics tests are grouped into three levels:

- **Level 0**—Level 0 diagnostics check for the presence of various components and perform essential path verifications. In addition, they verify the identification registers of the components on the board.
- **Level 1**—A smaller set of diagnostic tests. Level 1 diagnostics perform status/self-test for all the components on the board and test their registers for appropriate values. In addition, they perform extensive tests on memory devices (e.g., SDRAM, flash, NVRAM, EEPROM, and CPLD) wherever possible.
- **Level 2**—The full set of diagnostic tests. Level 2 diagnostics are used primarily for on-board loopback tests and more extensive component diagnostics. Various components on the board are put into loopback mode, and test packets are transmitted through those components. These diagnostics also perform snake tests using VLAN configurations.

Important Points to Remember

- You can only perform offline diagnostics on an offline standalone unit or offline member unit of a stack of three or more. You cannot perform diagnostics on the management or standby unit in a stack of two or more (Message 1).

Message 1 Offline Diagnostics on Master/Standby Error

```
Running Diagnostics on master/standby unit is not allowed on stack.
```

- Perform offline diagnostics on one stack member at a time.
- Diagnostics only test connectivity, not the entire data path.
- Diagnostic results are stored on the flash of the unit on which you performed the diagnostics.
- When offline diagnostics are complete, the unit or stack member reboots automatically.

Running Offline Diagnostics

1. Place the unit in the offline state using the **offline stack-unit** command from EXEC Privilege mode, as shown in Figure 62-1. YOu cannot enter the command on a Master or Standby stack unit.

The system reboots when the off-line diagnostics complete. This is an automatic process. A warning message appears when the **offline stack-unit** command is implemented.

```
Warning - Diagnostic execution will cause stack-unit to reboot after completion of
diags.
Proceed with Offline-Diags [confirm yes/no]:y
```

Figure 62-1. Taking an S-Series Stack Unit Offline

```
FTOS#offline stack-unit 2
Warning - Diagnostic execution will cause stack-unit to reboot after completion of diags.
Proceed with Offline-Diags [confirm yes/no]:y
5w6d12h: %STKUNIT0-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 2 down - stack unit offline
5w6d12h: %STKUNIT0-M:CP %IFMGR-1-DEL_PORT: Removed port: Gi 2/1-48
FTOS#5w6d12h: %STKUNIT1-S:CP %IFMGR-1-DEL_PORT: Removed port: Gi 2/1-48
```

2. Use the **show system brief** command from EXEC Privilege mode to confirm offline status, as shown in Figure 62-2.

Figure 62-2. Verifying the Offline/Online Status of an S-Series Stack Unit

```
FTOS#show system brief | no-more

Stack MAC : 00:01:e8:d6:02:39

--  Stack Info  --
Unit  UnitType     Status         ReqTyp      CurTyp      Version     Ports
-------------------------------------------------------------------------------
0 Standby      online         S25V        S25V        4.7.7.220   28
  1   Management   offline S50N        S50N        4.7.7.220   52
  2   Member       online         S25P        S25P        4.7.7.220   28
3   Member       not present
  4   Member       not present
  5   Member       not present
  6   Member       not present
  7   Member       not present

--  Module Info  --
Unit  Module No   Status        Module Type        Ports
-------------------------------------------------------------------------------
  0   0           online        S50-01-10GE-2C      2
  0   1           online        S50-01-12G-2S       2
  1   0           online        S50-01-10GE-2P      2
  1   1           online        S50-01-12G-2S       2
  2   0           not present   No Module           0
  2   1           offline       S50-01-12G-2S       2

--  Power Supplies  --
Unit   Bay   Status       Type
-------------------------------------------------------------------------------
  0     0     up            AC
  0     1     absent
  1     0     up            AC
  1     1     absent
  2     0     up            AC
  2     1     absent
```

3. Start diagnostics on the unit using the command **diag**, as shown in Figure 62-3. When the tests are complete, the system displays syslog Message 2, and automatically reboots the unit. Diagnostic results are printed to a file in the flash using the filename format *TestReport-SU-<stack-unit>.txt*.

Message 2 Offline Diagnostics Complete

```
FTOS#00:09:32 : Diagnostic test results are stored on file: flash:/TestReport-SU-1.txt
00:09:37: %S50N:1 %DIAGAGT-6-DA_DIAG_DONE: Diags finished on stack unit 1
Diags completed... Rebooting the system now!!!
```

As shown in Figure 62-3 and Figure 62-4, log messages differ somewhat when diagnostics are done on a standalone unit and on a stack member.

Figure 62-3. Running Offline Diagnostics on an S-Series Standalone Unit

```
FTOS#diag stack-unit 1 alllevels
Warning - diagnostic execution will cause multiple link flaps on the peer side - advisable
to shut directly connected ports
Proceed with Diags [confirm yes/no]: yes
00:03:35: %S50N:1 %DIAGAGT-6-DA_DIAG_STARTED: Starting diags on stack unit 1
00:03:35 : Approximate time to complete these Diags ... 6 Min
S50N#00:09:32 : Diagnostic test results are stored on file: flash:/TestReport-SU-0.txt
00:09:37: %S50N:0 %DIAGAGT-6-DA_DIAG_DONE: Diags finished on stack unit 0
Diags completed... Rebooting the system now!!!

[reboot output omitted]

S50N#00:01:35: %STKUNIT0-M:CP %SYS-5-CONFIG_I: Configured from console by  console
dir
Directory of flash:

  1  drw-      16384   Jan 01 1980 00:00:00 +00:00 .
  2  drwx       1536   Feb 29 1996 00:05:22 +00:00 ..
  3  drw-        512   Aug 15 1996 23:09:48 +00:00 TRACE_LOG_DIR
  4  d---        512   Aug 15 1996 23:09:52 +00:00 ADMIN_DIR
  5  -rw-       3854   Sep 24 1996 03:43:46 +00:00 startup-config
  6  -rw-      12632   Nov 05 2008 17:15:16 +00:00 TestReport-SU-1.txt

flash: 3104256 bytes total (3086336 bytes free)
```

Figure 62-4 shows the output of the master and member units when you run offline diagnostics on a member unit.

Figure 62-4. Running Offline Diagnostics on an S-Series Stack Member

```
[output from master unit]
FTOS#diag stack-unit 2
Warning - the stack unit will be pulled out of the stack for diagnostic execution
Proceed with Diags [confirm yes/no]: yes
Warning - diagnostic execution will cause multiple link flaps on the peer side - advisable
to shut directly connected ports
Proceed with Diags [confirm yes/no]: yes
FTOS#00:03:13: %S25P:2 %DIAGAGT-6-DA_DIAG_STARTED: Starting diags on stack unit 2
00:03:13 : Approximate time to complete these Diags ... 6 Min
00:03:13 : Diagnostic test results will be stored on stack unit 2 file: flash:/
TestReport-SU-2.txt
FTOS#00:03:35: %STKUNIT1-M:CP %CHMGR-2-STACKUNIT_DOWN: Stack unit 2 down - card removed
00:08:50: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 2 present
00:09:00: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 2 (type S25P, 28 ports)
00:09:00: %S25P:2 %CHMGR-0-PS_UP: Power supply 0 in unit 2 is up
00:09:00: %STKUNIT1-M:CP %CHMGR-5-STACKUNITUP: Stack unit 2 is up
[output from the console of the unit in which diagnostics are performed]
FTOS(stack-member-2)#
Diagnostic test results are stored on file: flash:/TestReport-SU-2.txt
Diags completed... Rebooting the system now!!!
```

4. View the results of the diagnostic tests using the command **show file flash://** from EXEC Privilege mode, as shown in Figure 62-5.

Figure 62-5. Viewing the Results of Offline Diagnostics on a Standalone Unit

```
FTOS#show file flash://TestReport-SU-0.txt

*********************************S-Series Diagnostics*******************
Stack Unit Board Serial Number : DL267160098
CPU Version : MPC8541, Version: 1.1
PLD Version : 5
Diag image based on build : E_MAIN4.7.7.206
Stack Unit Board Voltage levels - 3.300000 V, 2.500000 V, 1.800000 V, 1.250000 V, 1.200000
V, 2.000000 V
Stack Unit Board temperature : 26 Degree C
Stack Unit Number : 0

**************************Stack Unit EEPROM INFO*******************************

********MFG INFO******************

Data in Chassis Eeprom Mfg Info is listed as...
Vendor Id: 07
Country Code: 01
Date Code: 12172007
Serial Number: DL267160098
Part Number: 7590003600
Product Revision: B
Product Order Number: ${

*************************** LEVEL 0 DIAGNOSTICS*************************

Test 0 - CPLD Presence Test ..................................... PASS
 Hardware PCB Revision is - Revision B
Test 1 - CPLD Hardware PCB Revision Test ........................ PASS
Test 2.000 - CPLD Fan-0 Presence Test ........................... PASS
Test 2.001 - CPLD Fan-1 Presence Test ........................... PASS
Test 2.002 - CPLD Fan-2 Presence Test ........................... PASS
Test 2.003 - CPLD Fan-3 Presence Test ........................... PASS
Test 2.004 - CPLD Fan-4 Presence Test ........................... PASS
Test 2.005 - CPLD Fan-5 Presence Test ........................... PASS
Test 3.000 - CPLD Power Bay-0 Presence Test ..................... PASS
Test 3.001 - CPLD Power Bay-1 Presence Test .....................     NOT PRESENT
Test 4 - SDRAM Access Test ...................................... PASS
Test 5 - CPU Access Test ........................................ PASS
Test 6 - I2C Temp Access Test CPU Board ......................... PASS
Test 7 - I2C Temp Access Test Main Board ........................ PASS
Test 8 - RTC Access Test ........................................ PASS
--More--
```

Trace logs

In addition to the syslog buffer, FTOS buffers trace messages which are continuously written by various FTOS software tasks to report hardware and software events and status information. Each trace message provides the date, time, and name of the FTOS process. All messages are stored in a ring buffer and can be saved to a file either manually or automatically upon failover.

Auto Save on Crash or Rollover

Exception information on for master or standby units is stored in the **flash:/TRACE_LOG_DIR** directory. This directory contains files that save trace information when there has been a task crash or timeout.

On a master unit, the **TRACE_LOG_DIR** files can be reached by FTP or by using the **show file** command from the **flash://TRACE_LOG_DIR** directory.

On a Standby unit, the **TRACE_LOG_DIR** files can be reached only by using the **show file** command from the **flash://TRACE_LOG_DIR** directory.

Note: Non-management member units do not support this functionality.

Hardware watchdog timer

The **hardware watchdog** command automatically reboots an FTOS switch/router when a unit is unresponsive. This is a last resort mechanism intended to prevent a manual power cycle.

Command	Description
hardware watchdog	Enable the hardware watchdog mechanism.

Buffer tuning

Buffer Tuning allows you to modify the way your switch allocates buffers from its available memory, and helps prevent packet drops during a temporary burst of traffic.

The S-Series ASICs implement the key functions of queuing, feature lookups, and forwarding lookups in hardware.

- Forwarding Processor (FP) ASICs provide Ethernet MAC functions, queueing and buffering, as well as store feature and forwarding tables for hardware-based lookup and forwarding decisions. 1G and 10G interfaces use different FPs.

Table 62-1 describes the type and number of ASICs per platform.

Table 62-1. ASICS by Platform

Hardware	FP	CSF
S50N, S50V	2	0
S25V, S25P, S25N	1	0

You can tune buffers at three locations, as shown in Figure 62-6.

1. CSF – Output queues going from the CSF.
2. FP Uplink—Output queues going from the FP to the CSF IDP links.
3. Front-End Link—Output queues going from the FP to the front-end PHY.

All ports support eight queues, 4 for data traffic and 4 for control traffic. All 8 queues are tunable.

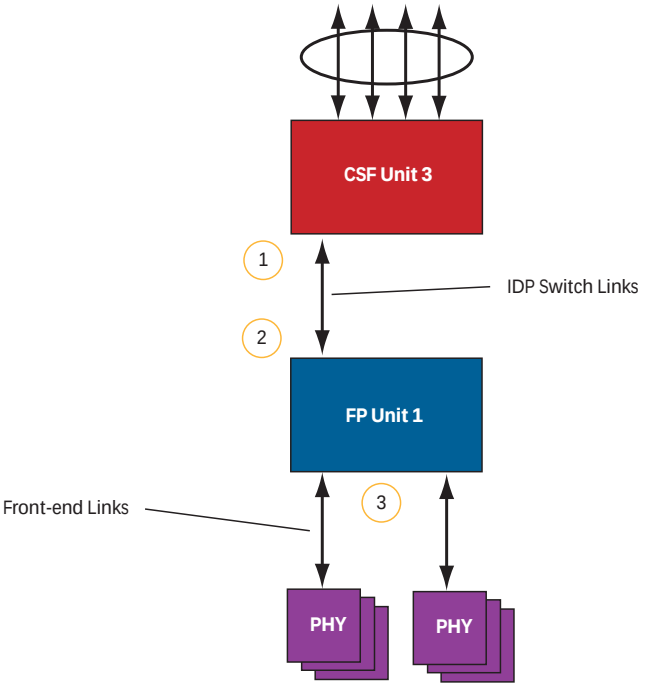
Physical memory is organized into cells of 128 bytes. The cells are organized into two buffer pools—dedicated buffer and dynamic buffer.

- **Dedicated buffer** is reserved memory that cannot be used by other interfaces on the same ASIC or by other queues on the same interface. This buffer is always allocated, and no dynamic recarving takes place based on changes in interface status. Dedicated buffers introduce a trade-off. They provide each interface with a guaranteed minimum buffer to prevent an overused and congested interface from starving all other interfaces. However, this minimum guarantee means the buffer manager does not reallocate the buffer to an adjacent congested interface, which means that in some cases, memory is underused.
- **Dynamic buffer** is shared memory that is allocated as needed, up to a configured limit. Using dynamic buffers provides the benefit of statistical buffer sharing. An interface requests dynamic buffers when its dedicated buffer pool is exhausted. The buffer manager grants the request based on three conditions:
- The number of used and available dynamic buffers
- The maximum number of cells that an interface can occupy
- Available packet pointers (2k per interface). Each packet is managed in the buffer using a unique packet pointer. Thus, each interface can manage up to 2k packets.

You can configure dynamic buffers per port on both 1G and 10G FPs and per queue on CSFs. By default, the FP dynamic buffer allocation is 10 times oversubscribed. For the 48-port 1G card:

- Dynamic Pool= Total Available Pool(16384 cells) – Total Dedicated Pool = 5904 cells
- Oversubscription ratio = 10
- Dynamic Cell Limit Per port = 59040/29 = 2036 cells

Figure 62-6. Buffer Tuning Points



Deciding to tune buffers

Dell Force10 recommends exercising caution when configuring any non-default buffer settings, as tuning can significantly affect system performance. The default values work for most cases.

As a guideline, consider tuning buffers if traffic is very bursty (and coming from several interfaces). In this case:

- Reduce the dedicated buffer on all queues/interfaces.
- Increase the dynamic buffer on all interfaces.
- Increase the cell pointers on a queue that you are expecting will receive the largest number of packets.

Buffer tuning commands

Task	Command	Command Mode
Define a buffer profile for the FP queues.	**buffer-profile fp fsqueue**	CONFIGURATION
Define a buffer profile for the CSF queues.	**buffer-profile csf csqueue**	CONFIGURATION
Change the dedicated buffers on a physical 1G interface.	**buffer dedicated**	BUFFER PROFILE
Change the maximum amount of dynamic buffers an interface can request.	**buffer dynamic**	BUFFER PROFILE
Change the number of packet-pointers per queue.	**buffer packet-pointers**	BUFFER PROFILE
Apply the buffer profile to a line card.	**buffer fp-uplink linecard**	CONFIGURATION
Apply the buffer profile to a CSF to FP link.	**buffer csf linecard**	CONFIGURATION

FTOS Behavior: If you attempt to apply a buffer profile to a non-existent port-pipe, FTOS displays the following message. However, the configuration still appears in the running-config.

```
%DIFFSERV-2-DSA_BUFF_CARVING_INVALID_PORT_SET: Invalid FP port-set 2 for linecard 2. Valid
range of port-set is <0-1>
```

Configuration changes take effect immediately and appear in the running configuration. Since under normal conditions all ports do not require the maximum possible allocation, the configured dynamic allocations can exceed the actual amount of available memory; this is called oversubscription. If you choose to oversubscribe the dynamic allocation, a burst of traffic on one interface might prevent other interfaces from receiving the configured dynamic allocation, which causes packet loss.

You cannot allocate more than the available memory for the dedicated buffers. If the system determines that the sum of the configured dedicated buffers allocated to the queues is more than the total available memory, the configuration is rejected, returning a syslog message similar to the following.

Table 62-2. Buffer Allocation Error

```
00:04:20: %S50N:0 %DIFFSERV-2-DSA_DEVICE_BUFFER_UNAVAILABLE: Unable to allocate dedicated buffers for
stack-unit 0, port pipe 0, egress port 25 due to unavailability of cells
```

FTOS Behavior: When you remove a buffer-profile using the command **no buffer-profile [fp | csf]** from CONFIGURATION mode, the buffer-profile name still appears in the output of **show buffer-profile [detail | summary]**. After a line card reset, the buffer profile correctly returns to the default values, but the profile name remains. Remove it from the **show buffer-profile** [**detail** | **summary**] command output by entering **no buffer** [**fp-uplink** |**csf**] **linecard port-set buffer-policy** from CONFIGURATION mode and **no buffer-policy** from INTERFACE mode.

Display the allocations for any buffer profile using the show commands in Figure 62-8. Display the default buffer profile using the command **show buffer-profile {summary | detail}** from EXEC Privilege mode, as shown in Figure 62-7.

Figure 62-7. Display the Default Buffer Profile

```
FTOS#show buffer-profile detail interface gigabitethernet 0/1
Interface Gi 0/1
Buffer-profile -
Dynamic buffer 194.88 (Kilobytes)
Queue#               Dedicated Buffer     Buffer Packets
                     (Kilobytes)
0                    2.50                 256
1                    2.50                 256
2                    2.50                 256
3                    2.50                 256
4                    9.38                 256
5                    9.38                 256
6                    9.38                 256
7                    9.38                 256
```

Figure 62-8. Displaying Buffer Profile Allocations

```
FTOS#show running-config interface tengigabitethernet 2/0 !
interface TenGigabitEthernet 2/0
no ip address
mtu 9252
switchport
no shutdown
buffer-policy myfsbufferprofile

FTOS#sho buffer-profile detail int gi 0/10
Interface Gi 0/10
Buffer-profile fsqueue-fp
Dynamic buffer 1256.00 (Kilobytes)
Queue#               Dedicated Buffer     Buffer Packets
                     (Kilobytes)
0                    3.00                 256
1                    3.00                 256
2                    3.00                 256
3                    3.00                 256
4                    3.00                 256
5                    3.00                 256
6                    3.00                 256
7                    3.00                 256

FTOS#sho buffer-profile detail fp-uplink stack-unit 0 port-set 0
Linecard 0 Port-set 0
Buffer-profile fsqueue-hig
Dynamic Buffer 1256.00 (Kilobytes)
Queue#               Dedicated Buffer     Buffer Packets
                     (Kilobytes)
0                    3.00                 256
1                    3.00                 256
2                    3.00                 256
3                    3.00                 256
4                    3.00                 256
5                    3.00                 256
6                    3.00                 256
7                    3.00                 256
```

Using a pre-defined buffer profile

FTOS provides two pre-defined buffer profiles, one for single-queue (i.e non-QoS) applications, and one for four-queue (i.e QoS) applications.

Task	Command	Mode
Apply one of two pre-defined buffer profiles for all port pipes in the system.	**buffer-profile global [1Q|4Q]**	CONFIGURATION

You must reload the system for the global buffer profile to take effect (Message 3).

Message 3 Reload After Applying Global Buffer Profile

```
% Info: For the global pre-defined buffer profile to take effect, please save the config and reload the
system.
```

FTOS Behavior: After you configure **buffer-profile global 1Q**, Message 3 is displayed during every bootup. Only one reboot is required for the configuration to take effect; afterwards this bootup message may be ignored.

FTOS Behavior: The buffer profile does not returned to the default, **4Q**, if you configure **1Q**, save the running-config to the startup-config, and then delete the startup-config and reload the chassis. The only way to return to the default buffer profile is to explicitly configure **4Q**, and then reload the chassis.

The **buffer-profile global** command fails if you have already applied a custom buffer profile on an interface.

Message 4 Global Buffer Profile Error

```
% Error: User-defined buffer profile already applied. Failed to apply global pre-defined buffer profile.
Please remove all user-defined buffer profiles.
```

Similarly, when **buffer-profile global** is configured, you cannot not apply a buffer profile on any single interface.

Message 5 Global Buffer Profile Error

```
% Error: Global pre-defined buffer profile already applied. Failed to apply user-defined buffer profile
on interface Gi 0/1. Please remove global pre-defined buffer profile.
```

If the default buffer profile (4Q) is active, FTOS displays an error message instructing you to remove the default configuration using the command **no buffer-profile global**.

Sample buffer profile configuration

The two general types of network environments are sustained data transfers and voice/data. Dell Force10 recommends a single-queue approach for data transfers, as shown in Figure 62-9.

Figure 62-9. Single Queue Application for S50N with Default Packet Pointers

```
!
buffer-profile fp fsqueue-fp
buffer dedicated queue0 3 queue1 3 queue2 3 queue3 3 queue4 3 queue5 3 queue6 3 queue7 3
buffer dynamic 1256
!
 buffer-profile fp fsqueue-hig
 buffer dedicated queue0 3 queue1 3 queue2 3 queue3 3 queue4 3 queue5 3 queue6 3 queue7 3
 buffer dynamic 1256

!
buffer fp-uplink stack-unit 0 port-set 0 buffer-policy fsqueue-hig
buffer fp-uplink stack-unit 0 port-set 1 buffer-policy fsqueue-hig
!
Interface range gi 0/1 - 48
buffer-policy fsqueue-fp

FTOS#sho run int gi 0/10
!
interface GigabitEthernet 0/10
 no ip address
```

Troubleshooting packet loss

The **show hardware stack-unit** command, introduced in FTOS 7.7.1.0 is intended primarily to troubleshoot packet loss. FTOS 7.7.1.1 augmented the statistics reported by existing command options (see Dataplane Statistics on page 1234), added commands and command options, and added **clear** commands to refresh those counters, as listed here:

- **show hardware stack-unit cpu data-plane statistics**
- **show hardware stack-unit cpu party-bus statistics**
- **show hardware stack-unit 0-7 drops unit 0-1 port 0-27**
- **show hardware stack-unit 0-7 stack-port 0-52**
- **show hardware stack-unit 0-7 unit 0-1 {counters | details | port-stats [detail] | register}**:
- **show hardware {layer2| layer3} acl stack-unit 0-7 stack 0-1**
- **show hardware layer3 qos stack-unit 0-7 port-set 0-1**
- **show hardware system-flow layer2 stack-unit 0-7 port-set 0-1 [counters]**
- **clear hardware stack-unit 0-7 counters**
- **clear hardware stack-unit 0-7 unit 0-1 counters**
- **clear hardware stack-unit 0-7 cpu data-plane statistics**
- **clear hardware stack-unit 0-7 cpu party-bus statistics**
- **clear hardware stack-unit 0-7 stack-port 0-52**

Displaying Drop Counters

The **show hardware stack-unit 0–7 drops [unit 0–1 [port 0–27]]** command assists in identifying which stack unit, port pipe, and port is experiencing internal drops, as shown in Figure 62-10 and Figure 62-11.

Figure 62-10. Displaying Drop Counter Statistics

```
FTOS#show hardware stack-unit 0 drops
UNIT No: 0
Total Ingress Drops :0
Total IngMac Drops :0
Total Mmu Drops :0
Total EgMac Drops :0
Total Egress Drops :0
UNIT No: 1
Total Ingress Drops :0
Total IngMac Drops :0
Total Mmu Drops :0
Total EgMac Drops :0
Total Egress Drops :0

FTOS#show hardware stack-unit 0 drops unit 0
Port# :Ingress Drops :IngMac Drops :Total Mmu Drops :EgMac Drops :Egress
Drops
1 0 0 0 0 0
2 0 0 0 0 0
3 0 0 0 0 0
4 0 0 0 0 0
5 0 0 0 0 0
6 0 0 0 0 0
7 0 0 0 0 0
8 0 0 0 0 0
```

Display drop counters with the **show hardware stack-unit drops unit port** command:

Figure 62-11. Displaying Drop Counters

```
FTOS#show hardware stack-unit 0 drops unit 0 port 1
 --- Ingress Drops      ---
Ingress Drops                 : 30
IBP CBP Full Drops            : 0
PortSTPnotFwd Drops           : 0
IPv4 L3 Discards              : 0
Policy Discards               : 0
Packets dropped by FP         : 14
(L2+L3) Drops                 : 0
Port bitmap zero Drops        : 16
Rx VLAN Drops                 : 0

--- Ingress MAC counters---
Ingress FCSDrops              : 0
Ingress MTUExceeds            : 0

--- MMU Drops          ---
HOL DROPS                     : 0
TxPurge CellErr               : 0
Aged Drops                    : 0

--- Egress MAC counters---
Egress FCS Drops              : 0

--- Egress FORWARD PROCESSOR Drops   ---
IPv4 L3UC Aged & Drops        : 0
TTL Threshold Drops           : 0
INVALID VLAN CNTR Drops       : 0
L2MC Drops                    : 0
PKT Drops of ANY Conditions   : 0
Hg MacUnderflow               : 0
TX Err PKT Counter            : 0
```

Dataplane Statistics

The **show hardware stack-unit cpu data-plane statistics** command provides insight into the packet types coming to the CPU. As shown in Figure 62-12, the command output has been augmented, providing detailed RX/TX packet statistics on a per-queue basis. The objective is to see whether CPU-bound traffic is internal (so-called party bus or IPC traffic) or network control traffic, which the CPU must process.

Figure 62-12. Displaying Dataplane Statistics

```
FTOS#show hardware stack-unit 2 cpu data-plane statistics

bc pci driver statistics for device:
 rxHandle        :0
 noMhdr          :0
 noMbuf          :0
 noClus          :0
 recvd           :0
 dropped         :0
 recvToNet       :0
 rxError         :0
 rxDatapathErr   :0
 rxPkt(COS0)     :0
 rxPkt(COS1)     :0
 rxPkt(COS2)     :0
 rxPkt(COS3)     :0
 rxPkt(COS4)     :0
 rxPkt(COS5)     :0
 rxPkt(COS6)     :0
 rxPkt(COS7)     :0
 rxPkt(UNIT0)    :0
 rxPkt(UNIT1)    :0
 rxPkt(UNIT2)    :0
 rxPkt(UNIT3)    :0
 transmitted     :0
 txRequested     :0
 noTxDesc        :0
 txError         :0
 txReqTooLarge   :0
 txInternalError :0
 txDatapathErr   :0
 txPkt(COS0)     :0
 txPkt(COS1)     :0
 txPkt(COS2)     :0
 txPkt(COS3)     :0
 txPkt(COS4)     :0
 txPkt(COS5)     :0
 txPkt(COS6)     :0
 txPkt(COS7)     :0
 txPkt(UNIT0)    :0
 txPkt(UNIT1)    :0
 txPkt(UNIT2)    :0
 txPkt(UNIT3)    :0
```

The **show hardware stack-unit cpu party-bus statistics** command displays input and output statistics on the party bus, which carries inter-process communication traffic between CPUs, as shown in Figure 62-13.

Figure 62-13. Displaying Party Bus Statistics

```
FTOS#sh hardware stack-unit 2 cpu party-bus statistics
Input Statistics:
    27550 packets, 2559298 bytes
    0 dropped, 0 errors
Output Statistics:
    1649566 packets, 1935316203 bytes
    0 errors
```

Displaying Stack Port Statistics

The **show hardware stack-unit stack-port** command displays input and output statistics for a stack-port interface, as shown in Figure 62-14.

Figure 62-14. Displaying Stack Unit Statistics

```
FTOS#show hardware stack-unit 2 stack-port 49
Input Statistics:
     27629 packets, 3411731 bytes
     0 64-byte pkts, 27271 over 64-byte pkts, 207 over 127-byte pkts
     17 over 255-byte pkts, 56 over 511-byte pkts, 78 over 1023-byte pkts
     0 Multicasts, 5 Broadcasts
     0 runts, 0 giants, 0 throttles
     0 CRC, 0 overrun, 0 discarded
Output Statistics:
     1649714 packets, 1948622676 bytes, 0 underruns
     0 64-byte pkts, 27234 over 64-byte pkts, 107970 over 127-byte pkts
     34 over 255-byte pkts, 504838 over 511-byte pkts, 1009638 over 1023-byte pkts
     0 Multicasts, 0 Broadcasts, 1649714 Unicasts
     0 throttles, 0 discarded, 0 collisions
Rate info (interval 45 seconds):
     Input 00.00 Mbits/sec,          2 packets/sec, 0.00% of line-rate
     Output 00.06 Mbits/sec,          8 packets/sec, 0.00% of line-rate
FTOS#
```

Displaying Stack Member Counters

The **show hardware stack-unit 0–7 {counters | details | port-stats [detail] | register}** command displays internal receive and transmit statistics, based on the selected command option. A sample of the output is shown for the **counters** option in Figure 62-15.

Figure 62-15. Displaying Stack Unit Counters

```
RIPC4.ge0       :                1,202             +1,202
RUC.ge0         :                1,224             +1,217
RDBGC0.ge0      :                   34                +24
RDBGC1.ge0      :                  366               +235
RDBGC5.ge0      :                   16                +12
RDBGC7.ge0      :                   18                +12
GR64.ge0        :                5,176                +24
GR127.ge0       :                1,566             +1,433
GR255.ge0       :                    4                 +4
GRPKT.ge0       :                1,602             +1,461
GRBYT.ge0       :              117,600           +106,202
GRMCA.ge0       :                  366               +235
GRBCA.ge0       :                   12                 +9
GT64.ge0        :                    4                 +3
GT127.ge0       :                  964               +964
GT255.ge0       :                    4                 +4
GT511.ge0       :                    1                 +1
GTPKT.ge0       :                  973               +972
GTBCA.ge0       :                    1                 +1
GTBYT.ge0       :               71,531            +71,467
RUC.cpu0        :                  972               +971
TDBGC6.cpu0     :                1,584             +1,449=
```

Application core dumps

Application core dumps are *disabled* by default. A core dump file can be very large. Due to memory requirements the file can only be sent directly to an FTP server. It is not stored on the local flash. Enable full application core dumps with the following:

Task	Command Syntax	Command Mode
Enable RPM core dumps and specify the shutdown mode. You may specify an IPv4 or IPv6 address for the server.	**logging coredump server**	CONFIGURATION

Undo this command using the **no logging coredump server.**

Mini core dumps

FTOS supports mini core dumps on the for application and kernel crashes. The mini core dump apply to Master, Standby and Member units.

Application and kernel mini core dumps are always enabled. The mini core dumps contain the stack space and some other very minimal information that can be used to debug a crash. These files are small files and are written into flash until space is exhausted. When the flash is full, the write process is stopped.

Note: If the Hardware watchdog timer is enabled and it triggers, no kernel core dump is generated because this is considered a power-off. In all other kernel crashes, a kernel mini core dump is be generated. When an application crashes, only the application coredump is generated.

A mini core dump contains critical information in the event of a crash. Mini core dump files are located in **flash:/ (root dir)**. The application mini core file name format is **f10StkUnit<***Stack_unit_no***>.<***Application name***>.acore.mini.txt**. The kernel mini core file name format is **f10StkUnit<***Stack_unit_no***>.kcore.mini.txt.** Sample files names are shown in Figure 62-16 and sample file text is shown in Figure 62-17.

Figure 62-16. Mini application core file naming example

```
FTOS#dir
Directory of flash:

  1  drw-      16384   Jan 01 1980 00:00:00 +00:00 .
  2  drwx       1536   Sep 03 2009 16:51:02 +00:00 ..
  3  drw-        512   Aug 07 2009 13:05:58 +00:00 TRACE_LOG_DIR
  4  d---        512   Aug 07 2009 13:06:00 +00:00 ADMIN_DIR
  5  -rw-       8693   Sep 03 2009 16:50:56 +00:00 startup-config
  6  -rw-       8693   Sep 03 2009 16:44:22 +00:00 startup-config.bak
  7  -rw-        156   Aug 28 2009 16:16:10 +00:00 f10StkUnit0.mrtm.acore.mini.txt
  8  -rw-        156   Aug 28 2009 17:17:24 +00:00 f10StkUnit0.vrrp.acore.mini.txt
  9  -rw-        156   Aug 28 2009 18:25:18 +00:00 f10StkUnit0.sysd.acore.mini.txt
 10  -rw-        156   Aug 28 2009 19:07:36 +00:00 f10StkUnit0.frrp.acore.mini.txt
 11  -rw-        156   Aug 31 2009 16:18:50 +00:00 f10StkUnit2.sysd.acore.mini.txt
 12  -rw-        156   Aug 29 2009 14:28:34 +00:00 f10StkUnit0.ipm1.acore.mini.txt
 13  -rw-        156   Aug 31 2009 16:14:56 +00:00 f10StkUnit0.acl.acore.mini.txt

flash: 3104256 bytes total (2959872 bytes free)
FTOS#
```

When a member or standby unit crashes, the mini core file gets uploaded to master unit. When the master unit crashes, the mini core file is uploaded to new master.

Figure 62-17. Mini core text file example

```
                      VALID MAGIC
------------------------PANIC STRING -----------------
panic string is :<null>
----------------------STACK TRACE START---------------
0035d60c <f10_save_mmu+0x120>:
00274f8c <panic+0x144>:
0024e2b0 <db_fncall+0x134>:
0024dee8 <db_command+0x258>:
0024d9c4 <db_command_loop+0xc4>:
002522b0 <db_trap+0x158>:
0026a8d0 <mi_switch+0x1b0>:
0026a00c <bpendtsleep>:
-----------------------STACK TRACE END----------------

---------------------------FREE MEMORY---------------
uvmexp.free = 0x2312
```

The panic string contains key information regarding the crash. Several panic string types exist, and they are displayed in regular english text to enable easier understanding of the crash cause.

63

Standards Compliance

This appendix contains the following sections:

- IEEE Compliance
- RFC and I-D Compliance
- MIB Location

Note: Unless noted, when a standard cited here is listed as supported by FTOS, FTOS also supports predecessor standards. One way to search for predecessor standards is to use the http://tools.ietf.org/ website. Click on "**Browse and search IETF documents**", enter an RFC number, and inspect the top of the resulting document for obsolescence citations to related RFCs.

IEEE Compliance

- 802.1AB — LLDP
- 802.1D — Bridging, STP
- 802.1p — L2 Prioritization
- 802.1Q — VLAN Tagging, Double VLAN Tagging, GVRP
- 802.1s — MSTP
- 802.1w — RSTP
- 802.1X — Network Access Control (Port Authentication)
- 802.3ab — Gigabit Ethernet (1000BASE-T)
- 802.3ac — Frame Extensions for VLAN Tagging
- 802.3ad — Link Aggregation with LACP
- 802.3ae — 10 Gigabit Ethernet (10GBASE-W, 10GBASE-X)
- 802.3af — Power over Ethernet
- 802.3ak — 10 Gigabit Ethernet (10GBASE-CX4)
- 802.3i — Ethernet (10BASE-T)
- 802.3u — Fast Ethernet (100BASE-FX, 100BASE-TX)
- 802.3x — Flow Control
- 802.3z — Gigabit Ethernet (1000BASE-X)
- ANSI/TIA-1057— LLDP-MED
- Dell Force10 — FRRP (Force10 Redundant Ring Protocol)

- Dell Force10 — PVST+
- SFF-8431 — SFP+ Direct Attach Cable (10GSFP+Cu)
- MTU — 9,252 bytes

RFC and I-D Compliance

The following standards are supported by FTOS, and are grouped by related protocol. The columns showing support by platform indicate which version of FTOS first supports the standard.

Note: Checkmarks (✓) in the E-Series column indicate that FTOS support was added before FTOS version 7.5.1.

General Internet Protocols

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
768	User Datagram Protocol	7.6.1	7.5.1	✓	8.1.1
793	Transmission Control Protocol	7.6.1	7.5.1	✓	8.1.1
854	Telnet Protocol Specification	7.6.1	7.5.1	✓	8.1.1
959	File Transfer Protocol (FTP)	7.6.1	7.5.1	✓	8.1.1
1321	The MD5 Message-Digest Algorithm	7.6.1	7.5.1	✓	8.1.1
1350	The TFTP Protocol (Revision 2)	7.6.1	7.5.1	✓	8.1.1
1661	The Point-to-Point Protocol (PPP)			✓	
1989	PPP Link Quality Monitoring			✓	
1990	The PPP Multilink Protocol (MP)			✓	
1994	PPP Challenge Handshake Authentication Protocol (CHAP)			✓	
2474	Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers	7.7.1	7.5.1	✓	8.1.1
2615	PPP over SONET/SDH			✓	
2615	PPP over SONET/SDH			✓	
2698	A Two Rate Three Color Marker			✓	8.1.1
3164	The BSD syslog Protocol	7.6.1	7.5.1	✓	8.1.1
draft-ietf-bfd-base-03	Bidirectional Forwarding Detection		7.6.1	✓	8.1.1

General IPv4 Protocols

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
791	Internet Protocol	7.6.1	7.5.1	✓	8.1.1
792	Internet Control Message Protocol	7.6.1	7.5.1	✓	8.1.1
826	An Ethernet Address Resolution Protocol	7.6.1	7.5.1	✓	8.1.1
1027	Using ARP to Implement Transparent Subnet Gateways	7.6.1	7.5.1	✓	8.1.1
1035	DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION (client)	7.6.1	7.5.1	✓	8.1.1
1042	A Standard for the Transmission of IP Datagrams over IEEE 802 Networks	7.6.1	7.5.1	✓	8.1.1
1191	Path MTU Discovery	7.6.1	7.5.1	✓	8.1.1
1305	Network Time Protocol (Version 3) Specification, Implementation and Analysis	7.6.1	7.5.1	✓	8.1.1
1519	Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy	7.6.1	7.5.1	✓	8.1.1
1542	Clarifications and Extensions for the Bootstrap Protocol	7.6.1	7.5.1	✓	8.1.1
1812	Requirements for IP Version 4 Routers	7.6.1	7.5.1	✓	8.1.1
2131	Dynamic Host Configuration Protocol	7.6.1	7.5.1	✓	8.1.1
2338	Virtual Router Redundancy Protocol (VRRP)	7.6.1	7.5.1	✓	8.1.1
3021	Using 31-Bit Prefixes on IPv4 Point-to-Point Links	7.7.1	7.7.1	7.7.1	8.1.1
3046	DHCP Relay Agent Information Option	7.8.1	7.8.1		
3069	VLAN Aggregation for Efficient IP Address Allocation	7.8.1	7.8.1		
3128	Protection Against a Variant of the Tiny Fragment Attack	7.6.1	7.5.1	✓	8.1.1

General IPv6 Protocols

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
1886	DNS Extensions to support IP version 6	7.8.1	7.8.1	✓	8.2.1
1981 (Partial)	Path MTU Discovery for IP version 6	7.8.1	7.8.1	✓	8.2.1
2460	Internet Protocol, Version 6 (IPv6) Specification	7.8.1	7.8.1	✓	8.2.1
2461 (Partial)	Neighbor Discovery for IP Version 6 (IPv6)	7.8.1	7.8.1	✓	8.2.1
2462 (Partial)	IPv6 Stateless Address Autoconfiguration	7.8.1	7.8.1	✓	8.2.1
2463	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification	7.8.1	7.8.1	✓	8.2.1
2464	Transmission of IPv6 Packets over Ethernet Networks	7.8.1	7.8.1	✓	8.2.1
2675	IPv6 Jumbograms	7.8.1	7.8.1	✓	8.2.1
3587	IPv6 Global Unicast Address Format	7.8.1	7.8.1	✓	8.2.1
4291	Internet Protocol Version 6 (IPv6) Addressing Architecture	7.8.1	7.8.1	✓	8.2.1

Border Gateway Protocol (BGP)

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
1997	BGP Communities Attribute	7.8.1	7.7.1	✓	8.1.1
2385	Protection of BGP Sessions via the TCP MD5 Signature Option	7.8.1	7.7.1	✓	8.1.1
2439	BGP Route Flap Damping	7.8.1	7.7.1	✓	8.1.1
2545	Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing		7.8.1	✓	8.2.1
2796	BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)	7.8.1	7.7.1	✓	8.1.1
2842	Capabilities Advertisement with BGP-4	7.8.1	7.7.1	✓	8.1.1
2858	Multiprotocol Extensions for BGP-4	7.8.1	7.7.1	✓	8.1.1
2918	Route Refresh Capability for BGP-4	7.8.1	7.7.1	✓	8.1.1
3065	Autonomous System Confederations for BGP	7.8.1	7.7.1	✓	8.1.1
4360	BGP Extended Communities Attribute	7.8.1	7.7.1	7.6.1	8.1.1
4893	BGP Support for Four-octet AS Number Space	7.8.1	7.7.1	7.7.1	8.1.1
5396	Textual Representation of Autonomous System (AS) Numbers	8.1.2	8.1.2	8.1.2	8.2.1
draft-ietf-idr-bgp4-20	A Border Gateway Protocol 4 (BGP-4)	7.8.1	7.7.1	✓	8.1.1
draft-ietf-idr-restart-06	Graceful Restart Mechanism for BGP	7.8.1	7.7.1	✓	8.1.1

Open Shortest Path First (OSPF)

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
1587	The OSPF Not-So-Stubby Area (NSSA) Option	7.6.1	7.5.1	✓	8.1.1
2154	OSPF with Digital Signatures	7.6.1	7.5.1	✓	8.1.1
2328	OSPF Version 2	7.6.1	7.5.1	✓	8.1.1
2370	The OSPF Opaque LSA Option	7.6.1	7.5.1	✓	8.1.1
2740	OSPF for IPv6		7.8.1	✓	8.2.1
3623	Graceful OSPF Restart	7.8.1	7.5.1	✓	8.1.1
4222	Prioritized Treatment of Specific OSPF Version 2 Packets and Congestion Avoidance	7.6.1	7.5.1	✓	8.1.1

Intermediate System to Intermediate System (IS-IS)

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
1142	OSI IS-IS Intra-Domain Routing Protocol (ISO DP 10589)			✓	8.1.1
1195	Use of OSI IS-IS for Routing in TCP/IP and Dual Environments			✓	8.1.1
2763	Dynamic Hostname Exchange Mechanism for IS-IS			✓	8.1.1
2966	Domain-wide Prefix Distribution with Two-Level IS-IS			✓	8.1.1
3373	Three-Way Handshake for Intermediate System to Intermediate System (IS-IS) Point-to-Point Adjacencies			✓	8.1.1
3567	IS-IS Cryptographic Authentication			✓	8.1.1
3784	Intermediate System to Intermediate System (IS-IS) Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)			✓	8.1.1
5120	M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)			7.8.1	8.2.1
5306	Restart Signaling for IS-IS			8.3.1	8.3.1
draft-ietf-isis-igp-p2p-over-lan-06	Point-to-point operation over LAN in link-state routing protocols			✓	8.1.1
draft-ietf-isis -ipv6-06	Routing IPv6 with IS-IS			7.5.1	8.2.1
draft-kaplan-isis-ext-eth-02	Extended Ethernet Frame Size Support			✓	8.1.1

Routing Information Protocol (RIP)

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
1058	Routing Information Protocol	7.8.1	7.6.1	✓	8.1.1
2453	RIP Version 2	7.8.1	7.6.1	✓	8.1.1

Multiprotocol Label Switching (MPLS)

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
2702	Requirements for Traffic Engineering Over MPLS				8.3.1
3031	Multiprotocol Label Switching Architecture				8.3.1
3032	MPLS Label Stack Encoding				8.3.1
3209	RSVP-TE: Extensions to RSVP for LSP Tunnels				8.3.1
3630	Traffic Engineering (TE) Extensions to OSPF Version 2				8.3.1
3784	Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)				8.3.1
3812	Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Management Information Base (MIB)				8.3.1
3813	Multiprotocol Label Switching (MPLS) Label Switching Router (LSR) Management Information Base (MIB)				8.3.1
4090	Fast Reroute Extensions to RSVP-TE for LSP Tunnels				8.3.1
4379	Detecting Multi-Protocol Label Switched Data Plane Failures (MPLS TE/LDP Ping & Traceroute				8.3.1
5036	LDP Specification				8.3.1
5063	Extensions to GMPLS Resource Reservation Protocol (RSVP) Graceful Restart				8.3.1

Multicast

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
1112	Host Extensions for IP Multicasting	7.8.1	7.7.1	✓	8.1.1
2236	Internet Group Management Protocol, Version 2	7.8.1	7.7.1	✓	8.1.1
2710	Multicast Listener Discovery (MLD) for IPv6			✓	8.2.1
3376	Internet Group Management Protocol, Version 3	7.8.1	7.7.1	✓	8.1.1
3569	An Overview of Source-Specific Multicast (SSM)	7.8.1 SSM for IPv4	7.7.1 SSM for IPv4	7.5.1 SSM for IPv4/ IPv6	8.2.1 SSM for IPv4
3618	Multicast Source Discovery Protocol (MSDP)			✓	8.1.1
3810	Multicast Listener Discovery Version 2 (MLDv2) for IPv6			✓	8.2.1
3973	Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised)			✓	
4541	Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches	7.6.1 (IGMPv1/v2)	7.6.1 (IGMPv1/v2)	✓ IGMPv1/v2/v3, MLDv1 Snooping	8.2.1 IGMPv1/v2/ v3, MLDv1 Snooping
draft-ietf-pim -sm-v2-new-05	Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)	7.8.1 PIM-SM for IPv4	7.7.1	✓ IPv4/ IPv6	8.2.1 PIM-SM for IPv4/IPv6

Network Management

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
1155	Structure and Identification of Management Information for TCP/IP-based Internets	7.6.1	7.5.1	✓	8.1.1
1156	Management Information Base for Network Management of TCP/IP-based internets	7.6.1	7.5.1	✓	8.1.1
1157	A Simple Network Management Protocol (SNMP)	7.6.1	7.5.1	✓	8.1.1
1212	Concise MIB Definitions	7.6.1	7.5.1	✓	8.1.1
1215	A Convention for Defining Traps for use with the SNMP	7.6.1	7.5.1	✓	8.1.1
1493	Definitions of Managed Objects for Bridges [except for the dot1dTpLearnedEntryDiscards object]	7.6.1	7.5.1	✓	8.1.1
1724	RIP Version 2 MIB Extension		7.5.1	✓	8.1.1
1850	OSPF Version 2 Management Information Base	7.6.1	7.5.1	✓	8.1.1
1901	Introduction to Community-based SNMPv2	7.6.1	7.5.1	✓	8.1.1
2011	SNMPv2 Management Information Base for the Internet Protocol using SMIv2	7.6.1	7.5.1	✓	8.1.1
2012	SNMPv2 Management Information Base for the Transmission Control Protocol using SMIv2	7.6.1	7.5.1	✓	8.1.1
2013	SNMPv2 Management Information Base for the User Datagram Protocol using SMIv2	7.6.1	7.5.1	✓	8.1.1
2024	Definitions of Managed Objects for Data Link Switching using SMIv2	7.6.1	7.5.1	✓	8.1.1
2096	IP Forwarding Table MIB	7.6.1	7.5.1	✓	8.1.1
2558	Definitions of Managed Objects for the Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH) Interface Type			✓	
2570	Introduction and Applicability Statements for Internet Standard Management Framework	7.6.1	7.5.1	✓	8.1.1
2571	An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks	7.6.1	7.5.1	✓	8.1.1
2572	Message Processing and Dispatching for the Simple Network Management Protocol (SNMP)	7.6.1	7.5.1	✓	8.1.1
2574	User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)	7.6.1	7.5.1	✓	8.1.1
2575	View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)	7.6.1	7.5.1	✓	8.1.1

Network Management (continued)

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
2576	Coexistence Between Version 1, Version 2, and Version 3 of the Internet-standard Network Management Framework	7.6.1	7.5.1	✓	8.1.1
2578	Structure of Management Information Version 2 (SMIv2)	7.6.1	7.5.1	✓	8.1.1
2579	Textual Conventions for SMIv2	7.6.1	7.5.1	✓	8.1.1
2580	Conformance Statements for SMIv2	7.6.1	7.5.1	✓	8.1.1
2618	RADIUS Authentication Client MIB, except the following four counters: radiusAuthClientInvalidServerAddresses radiusAuthClientMalformedAccessResponses radiusAuthClientUnknownTypes radiusAuthClientPacketsDropped	7.6.1	7.5.1	✓	8.1.1
2665	Definitions of Managed Objects for the Ethernet-like Interface Types	7.6.1	7.5.1	✓	8.1.1
2674	Definitions of Managed Objects for Bridges with Traffic Classes, Multicast Filtering and Virtual LAN Extensions	7.6.1	7.5.1	✓	8.1.1
2787	Definitions of Managed Objects for the Virtual Router Redundancy Protocol	7.6.1	7.5.1	✓	8.1.1
2819	Remote Network Monitoring Management Information Base: Ethernet Statistics Table, Ethernet History Control Table, Ethernet History Table, Alarm Table, Event Table, Log Table	7.6.1	7.5.1	✓	8.1.1
2863	The Interfaces Group MIB	7.6.1	7.5.1	✓	8.1.1
2865	Remote Authentication Dial In User Service (RADIUS)	7.6.1	7.5.1	✓	8.1.1
3273	Remote Network Monitoring Management Information Base for High Capacity Networks (64 bits): Ethernet Statistics High-Capacity Table, Ethernet History High-Capacity Table	7.6.1	7.5.1	✓	8.1.1
3416	Version 2 of the Protocol Operations for the Simple Network Management Protocol (SNMP)	7.6.1	7.5.1	✓	8.1.1
3418	Management Information Base (MIB) for the Simple Network Management Protocol (SNMP)	7.6.1	7.5.1	✓	8.1.1
3434	Remote Monitoring MIB Extensions for High Capacity Alarms, High-Capacity Alarm Table (64 bits)	7.6.1	7.5.1	✓	8.1.1
3580	IEEE 802.1X Remote Authentication Dial In User Service (RADIUS) Usage Guidelines	7.6.1	7.5.1	✓	8.1.1

Network Management (continued)

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
3815	Definitions of Managed Objects for the Multiprotocol Label Switching (MPLS), Label Distribution Protocol (LDP)				8.3.1
5060	Protocol Independent Multicast MIB	7.8.1	7.8.1	7.7.1	8.1.1
ANSI/TIA-1057	The LLDP Management Information Base extension module for TIA-TR41.4 Media Endpoint Discovery information	7.7.1	7.6.1	7.6.1	8.1.1
draft-grant-tacacs-02	The TACACS+ Protocol	7.6.1	7.5.1	✓	8.1.1
draft-ietf-idr-bgp4-mib-06	Definitions of Managed Objects for the Fourth Version of the Border Gateway Protocol (BGP-4) using SMIv2	7.8.1	7.7.1	✓	8.1.1
draft-ietf-isis-wg-mib-16	Management Information Base for Intermediate System to Intermediate System (IS-IS): isisSysObject (top level scalar objects) isisISAdjTable isisISAdjAreaAddrTable isisISAdjIPAddrTable isisISAdjProtSuppTable			✓	8.1.1
IEEE 802.1AB	Management Information Base module for LLDP configuration, statistics, local system data and remote systems data components.	7.7.1	7.6.1	7.6.1	8.1.1
IEEE 802.1AB	The LLDP Management Information Base extension module for IEEE 802.1 organizationally defined discovery information. (LLDP DOT1 MIB and LLDP DOT3 MIB)	7.7.1	7.6.1	7.6.1	8.1.1
IEEE 802.1AB	The LLDP Management Information Base extension module for IEEE 802.3 organizationally defined discovery information. (LLDP DOT1 MIB and LLDP DOT3 MIB)	7.7.1	7.6.1	7.6.1	8.1.1
ruzin-mstp-mib-02 (Traps)	Definitions of Managed Objects for Bridges with Multiple Spanning Tree Protocol	7.6.1	7.6.1	7.6.1	8.1.1
sFlow.org	sFlow Version 5	7.7.1	7.6.1	✓	8.1.1
sFlow.org	sFlow Version 5 MIB	7.7.1	7.6.1	✓	8.1.1
FORCE10-BGP4-V2-MIB	Dell Force10 BGP MIB (draft-ietf-idr-bgp4-mibv2-05)	7.8.1	7.7.1	✓	8.1.1
FORCE10-FIB-MIB	Dell Force10 CIDR Multipath Routes MIB (The IP Forwarding Table provides information that you can use to determine the egress port of an IP packet and troubleshoot an IP reachability issue. It reports the autonomous system of the next hop, multiple next hop support, and policy routing support)			7.6.1	8.1.1

Network Management (continued)

		FTOS support, per platform			
RFC#	**Full Name**	**S-Series**	**C-Series**	**E-Series TeraScale**	**E-Series ExaScale**
FORCE10-CS-CHASSIS-MIB	Dell Force10 C-Series Enterprise Chassis MIB		7.5.1		
FORCE10-IF-EXTENSION-MIB	Dell Force10 Enterprise IF Extension MIB (extends the Interfaces portion of the MIB-2 (RFC 1213) by providing proprietary SNMP OIDs for other counters displayed in the ”show interfaces” output)	7.6.1	7.6.1	7.6.1	8.1.1
FORCE10-LINKAGG-MIB	Dell Force10 Enterprise Link Aggregation MIB	7.6.1	7.5.1	✓	8.1.1
FORCE10-CHASSIS-MIB	Dell Force10 E-Series Enterprise Chassis MIB			✓	8.1.1
FORCE10-COPY-CONFIG-MIB	Dell Force10 File Copy MIB (supporting SNMP SET operation)	7.7.1	7.7.1	✓	8.1.1
FORCE10-MON-MIB	Dell Force10 Monitoring MIB	7.6.1	7.5.1	✓	8.1.1
FORCE10-PRODUCTS-MIB	Dell Force10 Product Object Identifier MIB	7.6.1	7.5.1	✓	8.1.1
FORCE10-SS-CHASSIS-MIB	Dell Force10 S-Series Enterprise Chassis MIB	7.6.1			
FORCE10-SMI	Dell Force10 Structure of Management Information	7.6.1	7.5.1	✓	8.1.1
FORCE10-SYSTEM-COMPONENT-MIB	Dell Force10 System Component MIB (enables the user to view CAM usage information)	7.6.1	7.5.1	✓	8.1.1
FORCE10-TC-MIB	Dell Force10 Textual Convention	7.6.1	7.5.1	✓	8.1.1
FORCE10-TRAP-ALARM-MIB	Dell Force10 Trap Alarm MIB	7.6.1	7.5.1	✓	8.1.1

MIB Location

Dell Force10 MIBs are under the **Force10 MIBs** subhead on the **Documentation** page of iSupport:

https://www.force10networks.com/csportal20/KnowledgeBase/Documentation.aspx

You also can obtain a list of selected MIBs and their OIDs at the following URL:

https://www.force10networks.com/csportal20/MIBs/MIB_OIDs.aspx

Some pages of iSupport require a login. To request an iSupport account, go to:

https://www.force10networks.com/CSPortal20/Support/AccountRequest.aspx

If you have forgotten or lost your account information, contact Dell Force10 TAC for assistance.

Index

Numerics

A

B

C

D

E

F

G

H

I

L

M

N

O

P

Q

R

S

T

U

V

W

X