| 1 | About this Guide ................................................................. 27  |
|  | Objectives ........................................................................ 27  |
|  | Audience .......................................................................... 27  |
|  | Conventions ....................................................................... 27  |
|  | Information Symbols .......................................................... 28  |
|  | Related Documents ............................................................. 28  |
| 2 | Configuration Fundamentals ............................................... 29  |
|  | Accessing the Command Line ................................................ 29  |
|  | CLI Modes .......................................................................... 30  |
|  | Navigating CLI Modes .......................................................... 31  |
|  | The do Command ................................................................... 34  |
|  | Undoing Commands ................................................................ 34  |
|  | Obtaining Help .................................................................... 35  |
|  | Entering and Editing Commands ............................................ 35  |
|  | Command History .................................................................. 36  |
|  | Filtering show Command Outputs ......................................... 37  |
|  | Multiple Users in Configuration mode .................................... 38  |
| 3 | Getting Started ................................................................. 39  |
|  | Console access ..................................................................... 39  |
|  | Serial console ..................................................................... 39  |
|  | Default Configuration ............................................................ 40  |
|  | Configure a Host Name .......................................................... 41  |
|  | Access the System Remotely .................................................. 41  |
|  | Access the C-Series, E-Series, S-Series, and the Z-Series Remotely 41  |
|  | Access the S-Series Remotely ................................................ 43  |
|  | Configure the Enable Password ............................................. 44  |
|  | Configuration File Management ............................................. 44  |
|  | Copy Files to and from the System ........................................ 45  |
|  | Save the Running-configuration ............................................. 46  |
|  | Configure the Overload bit for Startup Scenario ....................... 47  |
|  | View Files .......................................................................... 48  |
|  | File System Management ...................................................... 49  |
|  | View command history ......................................................... 50  |
|  | Upgrading FTOS ................................................................. 50  |
| 4 | Management ....................................................................... 51  |
|  | Configure Privilege Levels .................................................... 51  |
|  | Create a Custom Privilege Level ........................................... 51  |
|  | Apply a Privilege Level to a Username .................................... 54  |
|  | Apply a Privilege Level to a Terminal Line .............................. 54  |
Configure Logging ................................................. 55
Log Messages in the Internal Buffer ..................... 55
  Configuration Task List for System Log Management ... 55
Disable System Logging ......................................... 56
Send System Messages to a Syslog Server ............... 56
  Configure a Unix System as a Syslog Server .......... 56
Change System Logging Settings ......................... 57
Display the Logging Buffer and the Logging Configuration . 57
Configure a UNIX logging facility level ............... 58
Synchronize log messages .................................. 60
Enable timestamp on syslog messages .................. 60
File Transfer Services .................................... 61
  Configuration Task List for File Transfer Services ... 61
Terminal Lines ............................................... 63
  Deny and Permit Access to a Terminal Line .......... 63
  Configure Login Authentication for Terminal Lines ... 63
Time out of EXEC Privilege Mode ....................... 65
Telnet to Another Network Device ..................... 65
Lock CONFIGURATION mode ................................ 66
  Viewing the Configuration Lock Status ................. 67
Recovering from a Forgotten Password on the S4810 .... 67
  Recovering from a Forgotten Enable Password on the S4810 . 68
Recovering from a Failed Start on the S4810 .......... 68

5  802.1ag ....................................................... 71
Ethernet CFM .................................................. 71
Maintenance Domains ....................................... 72
Maintenance Points .......................................... 72
Maintenance End Points ..................................... 73
Implementation Information ............................... 74
Configure CFM ............................................... 74
  Related Configuration Tasks ............................. 74
Enable Ethernet CFM ....................................... 75
Create a Maintenance Domain ............................ 75
Create a Maintenance Association ...................... 76
Create Maintenance Points ............................... 76
  Create a Maintenance End Point ....................... 76
  Create a Maintenance Intermediate Point ........... 77
MP Databases ............................................... 77
Continuity Check Messages .............................. 78
  Enable CCM ............................................... 79
  Enable Cross-checking ................................ 80
Loopback Message and Response ....................... 80
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-Byte AS Numbers</td>
<td>186</td>
</tr>
<tr>
<td>AS4 Number Representation</td>
<td>186</td>
</tr>
<tr>
<td>AS Number Migration</td>
<td>189</td>
</tr>
<tr>
<td>BGP4 Management Information Base (MIB)</td>
<td>191</td>
</tr>
<tr>
<td>Important Points to Remember</td>
<td>191</td>
</tr>
<tr>
<td>Configuration Information</td>
<td>192</td>
</tr>
<tr>
<td>BGP Configuration</td>
<td>193</td>
</tr>
<tr>
<td>Configuration Task List for BGP</td>
<td>194</td>
</tr>
<tr>
<td>MBGP Configuration</td>
<td>236</td>
</tr>
<tr>
<td>BGP Regular Expression Optimization</td>
<td>237</td>
</tr>
<tr>
<td>Debugging BGP</td>
<td>237</td>
</tr>
<tr>
<td>Storing Last and Bad PDUs</td>
<td>238</td>
</tr>
<tr>
<td>Capturing PDUs</td>
<td>239</td>
</tr>
<tr>
<td>PDU Counters</td>
<td>240</td>
</tr>
<tr>
<td>Sample Configurations</td>
<td>240</td>
</tr>
<tr>
<td>10 Bare Metal Provisioning 3.0 (BMP 3.0)</td>
<td>251</td>
</tr>
<tr>
<td>Overview</td>
<td>251</td>
</tr>
<tr>
<td>Prerequisites</td>
<td>252</td>
</tr>
<tr>
<td>Important Information</td>
<td>252</td>
</tr>
<tr>
<td>BMP Process Overview</td>
<td>252</td>
</tr>
<tr>
<td>Preparing BMP</td>
<td>253</td>
</tr>
<tr>
<td>DHCP Server</td>
<td>253</td>
</tr>
<tr>
<td>File Server</td>
<td>255</td>
</tr>
<tr>
<td>Domain Name Server</td>
<td>255</td>
</tr>
<tr>
<td>Reload Modes</td>
<td>255</td>
</tr>
<tr>
<td>BMP Mode</td>
<td>256</td>
</tr>
<tr>
<td>Normal Mode</td>
<td>257</td>
</tr>
<tr>
<td>DHCP Configuration</td>
<td>257</td>
</tr>
<tr>
<td>FTOS Image Retrieval</td>
<td>257</td>
</tr>
<tr>
<td>Scripts</td>
<td>257</td>
</tr>
<tr>
<td>Pre-configuration Scripts</td>
<td>257</td>
</tr>
<tr>
<td>Post-configuration Scripts</td>
<td>258</td>
</tr>
<tr>
<td>Auto-execution Scripts</td>
<td>258</td>
</tr>
<tr>
<td>Configuration Tasks</td>
<td>259</td>
</tr>
<tr>
<td>System boot and set-up behavior in BMP Mode</td>
<td>260</td>
</tr>
<tr>
<td>BMP mode: Boot and Set-up Behavior</td>
<td>261</td>
</tr>
<tr>
<td>Reload using the Auto-execution Script (Normal mode only)</td>
<td>265</td>
</tr>
<tr>
<td>Script Examples</td>
<td>266</td>
</tr>
<tr>
<td>Auto-execution Script - Normal mode</td>
<td>266</td>
</tr>
<tr>
<td>Pre-configuration Script - BMP Mode</td>
<td>269</td>
</tr>
<tr>
<td>Chapter</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>11</td>
<td>Content Addressable Memory (CAM)</td>
</tr>
<tr>
<td></td>
<td>Content Addressable Memory</td>
</tr>
<tr>
<td></td>
<td>CAM Profiles</td>
</tr>
<tr>
<td></td>
<td>Microcode</td>
</tr>
<tr>
<td></td>
<td>CAM Profiling for ACLs</td>
</tr>
<tr>
<td></td>
<td>Boot Behavior</td>
</tr>
<tr>
<td></td>
<td>Example: EF Line Card with EG Chassis Profile (Card Problem)</td>
</tr>
<tr>
<td></td>
<td>Example: EH Line Card with EG Chassis Profile (Card Problem)</td>
</tr>
<tr>
<td></td>
<td>When to Use CAM Profiling</td>
</tr>
<tr>
<td></td>
<td>Important Points to Remember</td>
</tr>
<tr>
<td></td>
<td>Differences Between EtherScale and TeraScale</td>
</tr>
<tr>
<td></td>
<td>Select CAM Profiles</td>
</tr>
<tr>
<td></td>
<td>CAM Allocation</td>
</tr>
<tr>
<td></td>
<td>Test CAM Usage</td>
</tr>
<tr>
<td></td>
<td>View CAM Profiles</td>
</tr>
<tr>
<td></td>
<td>View CAM-ACL settings</td>
</tr>
<tr>
<td></td>
<td>View CAM Usage</td>
</tr>
<tr>
<td></td>
<td>Configure IPv4Flow Sub-partitions</td>
</tr>
<tr>
<td></td>
<td>Configure Ingress Layer 2 ACL Sub-partitions</td>
</tr>
<tr>
<td></td>
<td>Return to the Default CAM Configuration</td>
</tr>
<tr>
<td></td>
<td>CAM Optimization</td>
</tr>
<tr>
<td></td>
<td>Applications for CAM Profiling</td>
</tr>
<tr>
<td></td>
<td>LAG Hashing</td>
</tr>
<tr>
<td></td>
<td>LAG Hashing based on Bidirectional Flow</td>
</tr>
<tr>
<td></td>
<td>CAM profile for the VLAN ACL group feature</td>
</tr>
<tr>
<td></td>
<td>Troubleshoot CAM Profiling</td>
</tr>
<tr>
<td></td>
<td>CAM Profile Mismatches</td>
</tr>
<tr>
<td></td>
<td>QoS CAM Region Limitation</td>
</tr>
<tr>
<td>12</td>
<td>Control Plane Policing (CoPP)</td>
</tr>
<tr>
<td></td>
<td>Overview</td>
</tr>
<tr>
<td></td>
<td>Configure Control Plane Policing</td>
</tr>
<tr>
<td></td>
<td>Configure CoPP for protocols</td>
</tr>
<tr>
<td></td>
<td>Configure CoPP for CPU queues</td>
</tr>
<tr>
<td></td>
<td>Show commands</td>
</tr>
<tr>
<td>13</td>
<td>Data Center Bridging (DCB)</td>
</tr>
<tr>
<td></td>
<td>Ethernet Enhancements in Data Center Bridging</td>
</tr>
<tr>
<td></td>
<td>Priority-Based Flow Control</td>
</tr>
<tr>
<td></td>
<td>Enhanced Transmission Selection</td>
</tr>
<tr>
<td></td>
<td>Data Center Bridging Exchange Protocol (DCBx)</td>
</tr>
<tr>
<td></td>
<td>Data Center Bridging in a Traffic Flow</td>
</tr>
</tbody>
</table>
Managing ECMP Group Paths ......................................................... 398

17 Enabling FIPS Cryptography ....................................................... 399
   Preparing the System ............................................................ 399
   Enabling FIPS Mode ............................................................. 399
   Generating Host-Keys ........................................................... 400
   Monitoring FIPS Mode Status ............................................... 401
   Disabling the FIPS Mode ....................................................... 401

18 FIP Snooping ................................................................. 403
   Fibre Channel over Ethernet .................................................. 403
   Ensuring Robustness in a Converged Ethernet Network .......... 403
   FIP Snooping on Ethernet Bridges ....................................... 405
   FIP Snooping in a Switch Stack .......................................... 407
   Configuring FIP Snooping ..................................................... 407
      Enabling the FIP Snooping Feature .................................. 408
      Enabling FIP Snooping on VLANs .................................... 408
      Configuring the FC-MAP Value ....................................... 408
      Configuring a Port for a Bridge-to-Bridge Link ................. 409
      Configuring a Port for a Bridge-to-FCF Link .................... 409
      Impact on Other Software Features ................................ 409
      FIP Snooping Prerequisites .............................................. 409
      FIP Snooping Restrictions .............................................. 410
      Configuration Procedure ............................................... 410
   Displaying FIP Snooping Information .................................... 411
   FIP Snooping Configuration Example .................................... 416

19 Force10 Resilient Ring Protocol (FRRP) .................................. 419
   Protocol Overview .......................................................... 419
      Ring Status ................................................................. 420
      Multiple FRRP Rings ..................................................... 421
      Important FRRP Points ................................................. 421
      Important FRRP Concepts ............................................. 422
   Implementing FRRP .......................................................... 423
   FRRP Configuration ........................................................ 423
   Troubleshooting FRRP ....................................................... 428
      Configuration Checks .................................................... 428
   Sample Configuration and Topology .................................... 428

20 GARP VLAN Registration Protocol (GVRP) ............................. 431
   Protocol Overview .......................................................... 431
      Important Points to Remember ....................................... 431
### Configuring IGMP Snooping

- Enabling IGMP Immediate-leave
- Disabling Multicast Flooding
- Specifying a Port as Connected to a Multicast Router
- Configuring the Switch as Querier
- Fast Convergence after MSTP Topology Changes
- Designating a Multicast Router Interface

## 23 Interfaces

### Interface Types
- View Basic Interface Information
- Enable a Physical Interface

### Physical Interfaces
- Configuration Task List for Physical Interfaces
- Overview of Layer Modes
- Configure Layer 2 (Data Link) Mode
- Configure Layer 3 (Network) Mode

### Management Interfaces
- Configure Management Interfaces on the E-Series, C-Series and the S4810
- Configure Management Interfaces on the S-Series

### VLAN Interfaces

### Loopback Interfaces

### Null Interfaces

### Port Channel Interfaces

### Bulk Configuration
- Interface Range
- Bulk Configuration Examples

### Interface Range Macros
- Define the Interface Range
- Choose an Interface-range Macro

### Monitor and Maintain Interfaces
- Maintenance using TDR

### Splitting QSFP ports to SFP+ ports

### Link Debounce Timer
- Important Points to Remember about Link Debounce Timer
- Assign a debounce time to an interface
- Show debounce times in an interface
- Disable ports when one only SFM is available (E300 only)
- Disable port on one SFM

### Link Dampening
- Important Points to Remember
- Enable Link Dampening
Microsoft Clustering ..........................................................627
Default Behavior ..............................................................628
Configuring the Switch for Microsoft Server Clustering ..........628
Enable and Disable VLAN Flooding ......................................629
Configuring Redundant Pairs .............................................630
Important Points about Configuring Redundant Pairs ..........631
Restricting Layer 2 Flooding ..............................................633
Far-end Failure Detection ..................................................633
FEFD state changes ..........................................................635
Important Points to Remember ..........................................636
Configuring FEFD ............................................................636
Debugging FEFD ...............................................................637

30 Link Layer Discovery Protocol (LLDP). ........................639
802.1AB (LLDP) Overview .................................................639
Protocol Data Units ..........................................................639
Optional TLVs .................................................................641
Management TLVs ............................................................641
TIA-1057 (LLDP-MED) Overview ......................................643
TIA Organizationally Specific TLVs ....................................643
Configuring LLDP ............................................................647
Related Configuration Tasks ..............................................647
Important Points to Remember ..........................................648
LLDP Compatibility ..........................................................648
CONFIGURATION versus INTERFACE Configurations ..........648
Enabling LLDP ...............................................................649
Disabling and Undoing LLDP ..............................................649
Advertising TLVs .............................................................650
Viewing the LLDP Configuration .......................................651
Viewing Information Advertised by Adjacent LLDP Agents ....652
Configuring LLDPDU Intervals ..........................................653
Configuring Transmit and Receive Mode ............................654
Configuring a Time to Live ................................................655
Debugging LLDP ...............................................................656
Relevant Management Objects .........................................657

31 Multicast Source Discovery Protocol (MSDP) ................663
Protocol Overview ..........................................................663
Anycast RP .....................................................................664
Implementation Information ..............................................664
Configuring Multicast Source Discovery Protocol ...............665
Related Configuration Tasks ..............................................665
Enable MSDP .................................................................670
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manage the Source-active Cache</td>
<td>670</td>
</tr>
<tr>
<td>View the Source-active Cache</td>
<td>671</td>
</tr>
<tr>
<td>Limit the Source-active Cache</td>
<td>671</td>
</tr>
<tr>
<td>Clear the Source-active Cache</td>
<td>671</td>
</tr>
<tr>
<td>Enable the Rejected Source-active Cache</td>
<td>671</td>
</tr>
<tr>
<td>Accept Source-active Messages that fail the RFP Check</td>
<td>672</td>
</tr>
<tr>
<td>Limit the Source-active Messages from a Peer</td>
<td>674</td>
</tr>
<tr>
<td>Prevent MSDP from Caching a Local Source</td>
<td>675</td>
</tr>
<tr>
<td>Prevent MSDP from Caching a Remote Source</td>
<td>676</td>
</tr>
<tr>
<td>Prevent MSDP from Advertising a Local Source</td>
<td>677</td>
</tr>
<tr>
<td>Log Changes in Peership States</td>
<td>678</td>
</tr>
<tr>
<td>Terminate a Peership</td>
<td>678</td>
</tr>
<tr>
<td>Clear Peer Statistics</td>
<td>679</td>
</tr>
<tr>
<td>Debug MSDP</td>
<td>680</td>
</tr>
<tr>
<td>MSDP with Anycast RP</td>
<td>680</td>
</tr>
<tr>
<td>Reducing Source-active Message Flooding</td>
<td>682</td>
</tr>
<tr>
<td>Specify the RP Address Used in SA Messages</td>
<td>682</td>
</tr>
<tr>
<td>MSDP Sample Configurations</td>
<td>686</td>
</tr>
</tbody>
</table>

32 Multiple Spanning Tree Protocol (MSTP)                              | 691  |
| Protocol Overview                                                   | 691  |
| Implementation Information                                          | 692  |
| Configure Multiple Spanning Tree Protocol                           | 692  |
| Related Configuration Tasks                                         | 692  |
| Enable Multiple Spanning Tree Globally                              | 693  |
| Add and Remove Interfaces                                           | 693  |
| Create Multiple Spanning Tree Instances                             | 694  |
| Influence MSTP Root Selection                                       | 695  |
| Interoperate with Non-FTOS Bridges                                  | 695  |
| Modify Global Parameters                                            | 696  |
| Modify Interface Parameters                                         | 697  |
| Configure an EdgePort                                               | 698  |
| Flush MAC Addresses after a Topology Change                         | 699  |
| MSTP Sample Configurations                                          | 700  |
| Debugging and Verifying MSTP Configuration                          | 704  |

33 Multicast Features                                                 | 707  |
| Implementation Information                                          | 707  |
| Enable IP Multicast                                                 | 707  |
| Multicast with ECMP                                                 | 708  |
| Implementation Information                                          | 708  |
| First Packet Forwarding for Lossless Multicast                      | 709  |
| Multicast Policies                                                  | 710  |
IPv4 Multicast Policies .......................................................... 710
IPv6 Multicast Policies .......................................................... 715
Multicast Traceroute ............................................................. 717

34 Open Shortest Path First (OSPFv2 and OSPFv3) .................. 719
Protocol Overview ............................................................. 720
  Autonomous System (AS) Areas ........................................... 720
  Area Types ........................................................................ 721
  Networks and Neighbors ..................................................... 722
  Router Types ..................................................................... 722
  Designated and Backup Designated Routers ......................... 724
  Link-State Advertisements (LSAs) ........................................ 725
  Virtual Links ..................................................................... 726
  Router Priority and Cost ....................................................... 726
Implementing OSPF with FTOS ................................................. 727
  Graceful Restart ................................................................. 728
  Fast Convergence (OSPFv2, IPv4 only) ............................... 729
  Multi-Process OSPF (OSPFv2, IPv4 only) ............................ 729
  Processing SNMP and Sending SNMP Traps ......................... 730
  RFC-2328 Compliant OSPF Flooding .................................... 730
  OSPF ACK Packing ............................................................ 731
  OSPF Adjacency with Cisco Routers .................................... 731
Configuration Information ...................................................... 732
Configuration Task List for OSPFv2 (OSPF for IPv4) .............. 732
  Enable OSPFv2 ................................................................. 733
  Enable Multi-Process OSPF ............................................... 735
  Assign an OSPFv2 area ....................................................... 736
  Enable OSPFv2 on interfaces ............................................. 737
  Configure stub areas ........................................................ 739
  Configure OSPF Stub-Router Advertisement ......................... 740
  Enable passive interfaces ................................................... 741
  Enable fast-convergence .................................................... 742
  Change OSPFv2 parameters on interfaces ......................... 743
  Enable OSPFv2 authentication ......................................... 745
  Enable OSPFv2 graceful restart ........................................ 745
  Configure virtual links ...................................................... 747
  Filter routes .................................................................... 748
  Redistribute routes .......................................................... 749
  Troubleshooting OSPFv2 ................................................... 750
Sample Configurations for OSPFv2 ........................................ 753
  Basic OSPFv2 Router Topology .......................................... 753
Configuration Task List for OSPFv3 (OSPF for IPv6) .............. 754
  Enable IPv6 Unicast Routing ............................................. 755
Assign IPv6 addresses on an interface ........................................ 755
Assign Area ID on interface ...................................................... 755
Assign OSPFv3 Process ID and Router ID Globally ......................... 756
Configure stub areas .............................................................. 756
Configure Passive-Interface ..................................................... 757
Redistribute routes ............................................................... 758
Configure a default route ....................................................... 758
Enable OSPFv3 graceful restart ................................................. 759
OSPFv3 Authentication Using IPsec ........................................... 762
Troubleshooting OSPFv3 .......................................................... 772

35 PIM Sparse-Mode (PIM-SM) ..................................................... 775
   Implementation Information .................................................... 775
   Protocol Overview ............................................................. 775
   Requesting Multicast Traffic ................................................. 776
   Refusing Multicast Traffic ................................................... 776
   Sending Multicast Traffic .................................................... 776
   Important Points to Remember ............................................. 777
   Configure PIM-SM .............................................................. 777
   Related Configuration Tasks .................................................. 777
   Enable PIM-SM ................................................................. 778
   Configurable S,G Expiry Timers ............................................. 779
   Configure a Static Rendezvous Point ..................................... 780
   Override Bootstrap Router Updates ....................................... 781
   Configure a Designated Router ............................................. 781
   Create Multicast Boundaries and Domains ............................... 782
   PIM-SM Graceful Restart .................................................... 782
   Monitoring PIM ................................................................. 783

36 Port Monitoring ................................................................. 785
   Important Points to Remember ............................................. 785
   Port Monitoring on E-Series .................................................. 786
   E-Series TeraScale ............................................................ 786
   E-Series ExaScale ............................................................. 787
   Port Monitoring on C-Series and S-Series ............................... 787
   Configuring Port Monitoring ............................................... 790
   Flow-based Monitoring ....................................................... 791

37 Private VLANs (PVLAN) ....................................................... 793
   Private VLAN Concepts ....................................................... 793
   Private VLAN Commands ..................................................... 795
   Private VLAN Configuration Task List ................................... 796
### Implementation Information
- Configuration Information
  - Configuration Task List for RIP
  - RIP Configuration Example

### Remote Monitoring (RMON)
- Implementation
- Fault Recovery

### Rapid Spanning Tree Protocol (RSTP)
- Protocol Overview
- Configuring Rapid Spanning Tree
  - Related Configuration Tasks
  - Important Points to Remember
  - RSTP and VLT
- Configure Interfaces for Layer 2 Mode
- Enable Rapid Spanning Tree Protocol Globally
- Add and Remove Interfaces
- Modify Global Parameters
- Modify Interface Parameters
- Configure an EdgePort
- Influence RSTP Root Selection
- SNMP Traps for Root Elections and Topology Changes
- Fast Hellos for Link State Detection

### Software-Defined Networking (SDN)

### Security
- AAA Accounting
  - Configuration Task List for AAA Accounting
- AAA Authentication
  - Configuration Task List for AAA Authentication
- AAA Authorization
  - Privilege Levels Overview
  - Configuration Task List for Privilege Levels
- RADIUS
  - RADIUS Authentication and Authorization
  - Configuration Task List for RADIUS
- TACACS+
  - Configuration Task List for TACACS+
  - TACACS+ Remote Authentication and Authorization
  - Command Authorization
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual IP</td>
<td>973</td>
</tr>
<tr>
<td>Failover Roles</td>
<td>974</td>
</tr>
<tr>
<td>MAC Addressing on S-Series Stacks</td>
<td>974</td>
</tr>
<tr>
<td>Stacking LAG</td>
<td>976</td>
</tr>
<tr>
<td>Supported Stacking Topologies</td>
<td>976</td>
</tr>
<tr>
<td>High Availability on S-Series Stacks</td>
<td>977</td>
</tr>
<tr>
<td>Management Access on S-Series Stacks</td>
<td>978</td>
</tr>
<tr>
<td>Important Points to Remember - S4810 Stacking</td>
<td>979</td>
</tr>
<tr>
<td>S-Series Stacking Installation Tasks</td>
<td>980</td>
</tr>
<tr>
<td>Create an S-Series Stack</td>
<td>980</td>
</tr>
<tr>
<td>Add Units to an Existing S-Series Stack</td>
<td>985</td>
</tr>
<tr>
<td>Split an S-Series Stack</td>
<td>989</td>
</tr>
<tr>
<td>S-Series Stacking Configuration Tasks</td>
<td>989</td>
</tr>
<tr>
<td>Assign Unit Numbers to Units in an S-Series Stack</td>
<td>989</td>
</tr>
<tr>
<td>Create a Virtual Stack Unit on an S-Series Stack</td>
<td>990</td>
</tr>
<tr>
<td>Display Information about an S-Series Stack</td>
<td>990</td>
</tr>
<tr>
<td>Influence Management Unit Selection on an S-Series Stack</td>
<td>992</td>
</tr>
<tr>
<td>Manage Redundancy on an S-Series Stack</td>
<td>993</td>
</tr>
<tr>
<td>Reset a Unit on an S-Series Stack</td>
<td>993</td>
</tr>
<tr>
<td>Verifying a Stack Configuration</td>
<td>993</td>
</tr>
<tr>
<td>LED Status Indicators on an S4810 Stack</td>
<td>993</td>
</tr>
<tr>
<td>Display Status of Stacking Ports</td>
<td>994</td>
</tr>
<tr>
<td>Removing Units or Front End Ports from a Stack</td>
<td>995</td>
</tr>
<tr>
<td>Remove a Unit from an S-Series Stack</td>
<td>995</td>
</tr>
<tr>
<td>Remove Front End Port Stacking</td>
<td>996</td>
</tr>
<tr>
<td>Troubleshoot an S-Series Stack</td>
<td>997</td>
</tr>
<tr>
<td>Recover from Stack Link Flaps</td>
<td>997</td>
</tr>
<tr>
<td>Recover from a Card Problem State on an S-Series Stack</td>
<td>997</td>
</tr>
<tr>
<td>Recover from a Card Mismatch State on an S-Series Stack</td>
<td>999</td>
</tr>
<tr>
<td>49 Storm Control</td>
<td>1001</td>
</tr>
<tr>
<td>Configure Storm Control</td>
<td>1001</td>
</tr>
<tr>
<td>Configure storm control from INTERFACE mode</td>
<td>1001</td>
</tr>
<tr>
<td>Configure storm control from CONFIGURATION mode</td>
<td>1002</td>
</tr>
<tr>
<td>50 Spanning Tree Protocol (STP)</td>
<td>1003</td>
</tr>
<tr>
<td>Protocol Overview</td>
<td>1003</td>
</tr>
<tr>
<td>Configuring Spanning Tree</td>
<td>1003</td>
</tr>
<tr>
<td>Related Configuration Tasks</td>
<td>1003</td>
</tr>
<tr>
<td>Important Points to Remember</td>
<td>1004</td>
</tr>
<tr>
<td>Configuring Interfaces for Layer 2 Mode</td>
<td>1004</td>
</tr>
<tr>
<td>Enabling Spanning Tree Protocol Globally</td>
<td>1006</td>
</tr>
<tr>
<td>Adding an Interface to the Spanning Tree Group</td>
<td>1008</td>
</tr>
</tbody>
</table>
53 Upgrade Procedures .................................................. 1051
  Find the upgrade procedures .................................... 1051
  Get Help with upgrades .......................................... 1051

54 Virtual LANs (VLAN) .............................................. 1053
  Default VLAN ..................................................... 1054
  Port-Based VLANs ................................................ 1055
  VLANs and Port Tagging ......................................... 1055
  Configuration Task List for VLANs ......................... 1056
  VLAN Interface Counters ....................................... 1060
  Native VLANs .................................................... 1061
  Enable Null VLAN as the Default VLAN .................. 1062

55 Virtual Link Trunking (VLT) .................................. 1063
  Overview .......................................................... 1063
  Enhanced VLT .................................................... 1065
  VLT Concepts Terminology ...................................... 1065
  Configuring Virtual Link Trunking ......................... 1066
    Important Points to Remember .............................. 1066
    Configuration Notes ......................................... 1067
  RSTP and VLT ................................................... 1071
  VLT Bandwidth Monitoring .................................... 1071
  VLT and IGMP Snooping ......................................... 1072
  VLT Port Delayed Restoration .................................. 1072
  PIM-Sparse Mode Support on VLT .............................. 1073
  RSTP Configuration ............................................. 1074
  VLT Configuration Procedure .................................. 1075
  Verifying a VLT Configuration .............................. 1089
  Sample Configuration: Virtual Link Trunking ............ 1091
  Troubleshooting VLT ............................................. 1094

56 Virtual Router Redundancy Protocol (VRRP) .............. 1097
  VRRP Overview .................................................. 1097
  VRRP Benefits .................................................. 1099
  VRRP Implementation .......................................... 1099
  VRRP Configuration ........................................... 1101
    Configuration Task List for VRRP ......................... 1101
    VRRP initialization delay .................................... 1111
  Sample Configurations ........................................ 1112
  VRRP for IPv4 Configuration .................................. 1112
  VRRP for IPv6 Configuration .................................. 1114
  VRRP in VRF Configuration ................................... 1116
57 Standards Compliance ............................................. 1123
  IEEE Compliance .................................................. 1123
  RFC and I-D Compliance ......................................... 1124
  MIB Location ..................................................... 1134

58 Index ................................................................. 1135
About this Guide

Objectives

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

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 Chapter 57, 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:

<table>
<thead>
<tr>
<th>Convention</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>keyword</td>
<td>Keywords are in bold and should be entered in the CLI as listed.</td>
</tr>
<tr>
<td>parameter</td>
<td>Parameters are in italics and require a number or word to be entered in the CLI.</td>
</tr>
<tr>
<td>{X}</td>
<td>Keywords and parameters within braces must be entered in the CLI.</td>
</tr>
<tr>
<td>[X]</td>
<td>Keywords and parameters within brackets are optional.</td>
</tr>
<tr>
<td>x</td>
<td>y</td>
</tr>
</tbody>
</table>
Information Symbols

Table 1-1 describes symbols contained in this guide.

Table 1-1. Information Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Warning</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CES</td>
<td>Platform Specific Feature</td>
<td>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.</td>
</tr>
<tr>
<td>ETEX</td>
<td>E-Series Specific Feature/Command</td>
<td>If a feature or command applies to only one of the E-Series platforms, a separate symbol calls this to attention: E for the TeraScale or EX for the ExaScale.</td>
</tr>
<tr>
<td>S4810</td>
<td>S4810</td>
<td>This symbol indicates that the selected feature is supported on the S4810 but not on other S-Series systems.</td>
</tr>
<tr>
<td>*</td>
<td>Exception</td>
<td>This symbol is a note associated with some other text on the page that is marked with an asterisk.</td>
</tr>
</tbody>
</table>

Related Documents

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

- FTOS Command Reference
- Installing the System
- FTOS Release Notes
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 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.

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

telnet 172.31.1.53
Trying 172.31.1.53...
Connected to 172.31.1.53.
Escape character is '^]'.
Login: username
Password:
FTOS> EXEC mode prompt
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 34). 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 44.

- **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 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.
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

<table>
<thead>
<tr>
<th>CLI Command Mode</th>
<th>Prompt</th>
<th>Access Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXEC</td>
<td>FTOS&gt;</td>
<td>Access the router through the console or Telnet.</td>
</tr>
<tr>
<td>EXEC Privilege</td>
<td>FTOS#</td>
<td>• From EXEC mode, enter the command <strong>enable</strong>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• From any other mode, use the command <strong>end</strong>.</td>
</tr>
<tr>
<td>CONFIGURATION</td>
<td>FTOS(conf)#</td>
<td>• From EXEC privilege mode, enter the command <strong>configure</strong>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• From every mode except EXEC and EXEC Privilege, enter the command <strong>exit</strong>.</td>
</tr>
</tbody>
</table>

**Note:** Access all of the following modes from CONFIGURATION mode.

<table>
<thead>
<tr>
<th>INTERFACE modes</th>
<th>Prompt</th>
<th>Access Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARCHIVE</td>
<td>FTOS(conf-archive)</td>
<td>archive</td>
</tr>
<tr>
<td>AS-PATH ACL</td>
<td>FTOS(config-as-path)#</td>
<td>ip as-path access-list</td>
</tr>
<tr>
<td>Gigabit Ethernet Interface</td>
<td>FTOS(conf-if-gi-0/0)#</td>
<td></td>
</tr>
<tr>
<td>10 Gigabit Ethernet Interface</td>
<td>FTOS(conf-if-te-0/0)#</td>
<td></td>
</tr>
<tr>
<td>Interface Range</td>
<td>FTOS(conf-if-range)#</td>
<td></td>
</tr>
<tr>
<td>Loopback Interface</td>
<td>FTOS(conf-if-lo-0)#</td>
<td></td>
</tr>
<tr>
<td>Management Ethernet Interface</td>
<td>FTOS(conf-if-ma-0/0)#</td>
<td></td>
</tr>
<tr>
<td>Null Interface</td>
<td>FTOS(conf-if-nu-0)#</td>
<td></td>
</tr>
<tr>
<td>Port-channel Interface</td>
<td>FTOS(conf-if-po-0)#</td>
<td></td>
</tr>
<tr>
<td>SONET Interface</td>
<td>FTOS(conf-if-so-0/0)#</td>
<td></td>
</tr>
<tr>
<td>VLAN Interface</td>
<td>FTOS(conf-if-vl-0)#</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IP ACCESS-LIST</th>
<th>Prompt</th>
<th>Access Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>STANDARD ACCESS-LIST</td>
<td>FTOS(config-std-nacl)#</td>
<td>ip access-list standard</td>
</tr>
<tr>
<td>EXTENDED ACCESS-LIST</td>
<td>FTOS(config-ext-nacl)#</td>
<td>ip access-list extended</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LINE</th>
<th>Prompt</th>
<th>Access Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUXILIARY</td>
<td>FTOS(config-line-aux)#</td>
<td></td>
</tr>
<tr>
<td>CONSOLE</td>
<td>FTOS(config-line-console)#</td>
<td></td>
</tr>
<tr>
<td>VIRTUAL TERMINAL</td>
<td>FTOS(config-line-vty)#</td>
<td>line</td>
</tr>
</tbody>
</table>
Table 2-1. FTOS Command Modes (continued)

<table>
<thead>
<tr>
<th>CLI Command Mode</th>
<th>Prompt</th>
<th>Access Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>STANDARD ACCESS-LIST</td>
<td>FTOS(config-std-macl)#</td>
<td>mac access-list standard</td>
</tr>
<tr>
<td>EXTENDED ACCESS-LIST</td>
<td>FTOS(config-ext-macl)#</td>
<td>mac access-list extended</td>
</tr>
<tr>
<td>MULTIPLE SPANNING TREE</td>
<td>FTOS(config-mstp)#</td>
<td>protocol spanning-tree mstp</td>
</tr>
<tr>
<td>OPENFLOW</td>
<td>FTOS(conf-of-instance of-id)#</td>
<td>openflow of-instance of-id</td>
</tr>
<tr>
<td></td>
<td></td>
<td>of-id represents the OpenFlow instance ID.</td>
</tr>
<tr>
<td>Per-VLAN SPANNING TREE Plus</td>
<td>FTOS(config-pvst)#</td>
<td>protocol spanning-tree pvst</td>
</tr>
<tr>
<td>PREFIX-LIST</td>
<td>FTOS(conf-nprefixl)#</td>
<td>ip prefix-list</td>
</tr>
<tr>
<td>RAPID SPANNING TREE</td>
<td>FTOS(config-rstp)#</td>
<td>protocol spanning-tree rstp</td>
</tr>
<tr>
<td>REDIRECT</td>
<td>FTOS(conf-redirect-list)#</td>
<td>ip redirect-list</td>
</tr>
<tr>
<td>ROUTE-MAP</td>
<td>FTOS(config-route-map)#</td>
<td>route-map</td>
</tr>
<tr>
<td>ROUTER BGP</td>
<td>FTOS(conf-router_bgp)#</td>
<td>router bgp</td>
</tr>
<tr>
<td>ROUTER ISIS</td>
<td>FTOS(conf-router_isis)#</td>
<td>router isis</td>
</tr>
<tr>
<td>ROUTER OSPF</td>
<td>FTOS(conf-router_ospf)#</td>
<td>router ospf</td>
</tr>
<tr>
<td>ROUTER RIP</td>
<td>FTOS(conf-router_rip)#</td>
<td>router rip</td>
</tr>
<tr>
<td>SPANNING TREE</td>
<td>FTOS(config-span)#</td>
<td>protocol spanning-tree 0</td>
</tr>
<tr>
<td>TRACE-LIST</td>
<td>FTOS(conf-trace-acl)#</td>
<td>ip trace-list</td>
</tr>
</tbody>
</table>

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

**Figure 2-3. Changing CLI Modes**

```
FTOS(conf)#protocol spanning-tree 0
FTOS(config-span)#
```

New command prompt
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

```
FTOS(conf)#do show linecard all
-- Line cards --
Slot Status NxtBoot ReqTyp CurTyp Version Ports
---------------------------------------------------------------------------
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
```

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

```
FTOS(conf)#interface gigabitethernet 4/17
FTOS(conf-if-gi-4/17)#ip address 192.168.10.1/24
FTOS(conf-if-gi-4/17)#show config
!
interface GigabitEthernet 4/17
  ip address 192.168.10.1/24
  no shutdown
FTOS(conf-if-gi-4/17)#no ip address
FTOS(conf-if-gi-4/17)#show config
!
interface GigabitEthernet 4/17
```

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

```
FTOS#?  “?” at prompt for list of commands
calendar        Manage the hardware calendar
                Change current directory
change          Change subcommands
clear           Reset functions
clock           Manage the system clock
configure       Configuring from terminal
                Copy from one file to another
dump            Debug functions
                --More--
```

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

Figure 2-7. Keyword? Command Example

```
FTOS(conf)#cl?  partial keyword plus “[space]?” for matching keywords
class-map
clock
FTOS(conf)#cl
```

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

Figure 2-8. Keyword ? Command Example

```
FTOS(conf)#clock ?  keyword plus “[space]?” for compatible keywords
summer-time     Configure summer (daylight savings) time
timezone        Configure time zone
FTOS(conf)#clock
```

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

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

### 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-insensitive. 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.
Getting Started

This chapter contains the following major sections:

- Default Configuration
- Configure a Host Name
- Access the System Remotely
- Configure the Enable Password
- Configuration File Management
- File System Management

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 EXEC mode prompt.

For details on using the Command Line Interface (CLI), refer to Accessing the Command Line in the Configuration Fundamentals chapter.

Console access

The S4810 has 2 management ports available for system access: a serial console port and an Out-of-Bounds (OOB) port.

Serial console

The RJ-45/RS-232 console port is labeled on the S4810 chassis. It is in the upper right-hand side, as you face the I/O side of the chassis.
To access the console port, follow the procedures below. Refer to Table 3-1 for the console port pinout.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Install an RJ-45 copper cable into the console port. Use a rollover (crossover) cable to connect the S4810 console port to a terminal server.</td>
</tr>
<tr>
<td>2</td>
<td>Connect the other end of the cable to the DTE terminal server.</td>
</tr>
<tr>
<td>3</td>
<td>Terminal settings on the console port cannot be changed in the software and are set as follows: 9600 baud rate, No parity, 8 data bits, 1 stop bit, No flow control</td>
</tr>
</tbody>
</table>

**Accessing the RJ-45 console port with a DB-9 adapter**

You can connect to the console using a RJ-45 to RJ-45 rollover cable and a RJ-45 to DB-9 female DTE adapter to a terminal server (for example, PC). Table 3-1 lists the pin assignments.

<table>
<thead>
<tr>
<th>S-Series Console Port</th>
<th>RJ-45 to RJ-45 Rollover Cable</th>
<th>RJ-45 to DB-9 Adapter</th>
<th>Terminal Server Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>RJ-45 pinout</td>
<td>RJ-45 Pinout</td>
<td>DB-9 Pin</td>
</tr>
<tr>
<td>RTS</td>
<td>1</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>NC</td>
<td>2</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>TxD</td>
<td>3</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>GND</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>GND</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>RxD</td>
<td>6</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>NC</td>
<td>7</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>CTS</td>
<td>8</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>

**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 FTOS). You must configure the system using the CLI.
Configure a Host Name

The host name appears in the prompt. The default host name is FTOS.

- 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create a new host name.</td>
<td>hostname name</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

The example below illustrates the hostname command.

```
FTOS(conf)#hostname R1
R1(conf)#
```

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, E-Series, S-Series (except for S25 and S50), and Z9000 have a dedicated management port and a management routing table that is separate from the IP routing table.
- The S-Series (except the S4810) does not have a dedicated management port, but is managed from any port. It does not have a separate management routing table.
- All Dell Force10 products can be managed via the front-end data ports as well.

**Access the C-Series, E-Series, S-Series, and the Z-Series Remotely**

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.
To configure the management port IP address:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter INTERFACE mode for the Management port.</td>
<td><code>interface ManagementEthernet slot/port</code>&lt;br&gt;• <code>slot</code> range: 0 to 1&lt;br&gt;• <code>port</code> range: 0</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Assign an IP address to the interface.</td>
<td><code>ip address ip-address/mask</code>&lt;br&gt;• <code>ip-address</code>: an address in dotted-decimal format (A.B.C.D).&lt;br&gt;• <code>mask</code>: a subnet mask in /prefix-length format (/ xx).</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>3</td>
<td>Enable the interface.</td>
<td><code>no shutdown</code></td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

**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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure a management route to the network from which you are accessing the system.</td>
<td><code>management route ip-address/mask gateway</code>&lt;br&gt;• <code>ip-address</code>: the network address in dotted-decimal format (A.B.C.D).&lt;br&gt;• <code>mask</code>: a subnet mask in /prefix-length format (/ xx).&lt;br&gt;• <code>gateway</code>: the next hop for network traffic originating from the management port.</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**Configure a Username and Password**

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

**Note:** Assign different IP addresses to each RPM's management port.
To configure a username and password:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure a username and password to access the system remotely.</td>
<td>username username password [encryption-type] password encryption-type</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

username username 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.

### 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 S4810 system uses management ports and should be configured similar to the C-Series and E-Series systems. Refer to [Access the C-Series, E-Series, S-Series, and the Z-Series Remotely](#) for more information.

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 the example below.

```
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
```

2. Configure a IP route with a default gateway using the command `ip route` from CONFIGURATION mode, as shown in the example below.

```
R5(conf)#ip route 10.11.32.0/23 10.11.131.254
```

3. Configure a username and password using the command `username` from CONFIGURATION mode, as shown in the example below.

```
R5(conf)#username admin pass FTOS
```
Configure the Enable Password

Access the EXEC Privilege mode using the `enable` command. The EXEC Privilege mode is unrestricted by default. Configure a password as a basic security measure. 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create a password to access EXEC Privilege mode.</td>
<td>`enable [password</td>
<td>secret] [level level] [encryption-type] password`</td>
</tr>
<tr>
<td></td>
<td><code>level</code> is the privilege level, is 15 by default, and is not required.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>encryption-type</code> specifies how you are inputting the password, is 0 by default, and is not required.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 0 is for inputting the password in clear text.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 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.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 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.</td>
<td></td>
</tr>
</tbody>
</table>

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 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 Line Reference Guide 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 3-2.
- 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 3-2.

### Table 3-2. Forming a copy Command

<table>
<thead>
<tr>
<th>source-file-url Syntax</th>
<th>destination-file-url Syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Local File Location</strong></td>
<td></td>
</tr>
<tr>
<td>Internal flash:</td>
<td></td>
</tr>
<tr>
<td>primary RPM</td>
<td><code>copy flash://filename</code></td>
</tr>
<tr>
<td>standby RPM</td>
<td>`copy rpm{0</td>
</tr>
<tr>
<td>External flash:</td>
<td></td>
</tr>
<tr>
<td>primary RPM</td>
<td>`copy rpm{0</td>
</tr>
<tr>
<td>standby RPM</td>
<td>`copy rpm{0</td>
</tr>
<tr>
<td><strong>USB Drive (E-Series ExaScale)</strong></td>
<td></td>
</tr>
<tr>
<td>USB drive on RPM0</td>
<td><code>copy rpm0usbflash://filepath</code></td>
</tr>
<tr>
<td>External USB drive</td>
<td><code>copy usbfash://filepath</code></td>
</tr>
<tr>
<td><strong>Remote File Location</strong></td>
<td></td>
</tr>
<tr>
<td>TFTP server</td>
<td>`copy tftp://{hostip</td>
</tr>
<tr>
<td>SCP server</td>
<td>`copy scp://{hostip</td>
</tr>
</tbody>
</table>

**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 systems. Refer to your system’s Release Notes for a list of approved USB vendors.

• The **usbflash** command is supported on Z9000. Refer to your system’s Release Notes for a list of approved USB vendors.

The following text is an example of using the **copy** command to save a file to an FTP server.

```
FTOS#copy flash://FTOS-EF-8.2.1.0.bin ftp://myusername:mypassword@10.10.10.10//FTOS/FTOS-EF-8.2.1.0
27952672 bytes successfully copied
```

The following text is an example of using the **copy** command to import a file to the Dell Force10 system from an FTP server.

```
core1#copy ftp://myusername:mypassword@10.10.10.10//FTOS/FTOS-EF-8.2.1.0.bin flash://
Destination file name [FTOS-EF-8.2.1.0.bin.bin]:
26292881 bytes successfully copied
```

### 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 but use the filenames **startup-configuration** and **running-configuration**. These commands assume that current directory is the internal flash, which is the system default.
Configure the Overload bit for Startup Scenario

For information on setting the router overload bit for a specific period of time after a switch reload is implemented, see the *FTOS Command Line Reference Guide*, Chapter 18 - Intermediate System to Intermediate System (IS-IS).
## 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>View a list of files on:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>the internal flash of an RPM</td>
<td>dir flash:</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td></td>
<td>the external flash of an RPM</td>
<td>dir slot:</td>
<td></td>
</tr>
</tbody>
</table>

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 the example below.

```bash
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 Jan 01 1980 00:18:28 diag
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>View the:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>contents of a file in the internal flash of an RPM</td>
<td>show file rpm{0</td>
<td>1};flash://filename</td>
</tr>
<tr>
<td></td>
<td>contents of a file in the external flash of an RPM</td>
<td>show file rpm{0</td>
<td>1};slot0://filename</td>
</tr>
<tr>
<td></td>
<td>running-configuration</td>
<td>show running-config</td>
<td></td>
</tr>
<tr>
<td></td>
<td>startup-configuration</td>
<td>show startup-config</td>
<td></td>
</tr>
</tbody>
</table>
View Configuration Files

Configuration files have three commented lines at the beginning of the file, as shown in the example below, 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.

```
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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>View information about each file system.</td>
<td>show file-systems</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

The output of the command `show file-systems` in the example below shows the total capacity, amount of free memory, file structure, media type, read/write privileges for each storage device in use.

```
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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change the default directory.</td>
<td>cd directory</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

In the example below, 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.

```
FTOS#cd slot0:
FTOS#copy running-config test
FTOS#copy run test
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
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 timestamp 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 the example below.

```
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
```

Upgrading FTOS

**Note:** To upgrade FTOS, see the Release Notes for the version you want to load on the system.
This chapter explains the different protocols or services used to manage the Dell Force10 system including:

- Configure Privilege Levels
- Configure Logging
- File Transfer Services
- Terminal Lines
- Lock CONFIGURATION mode
- Recovering from a Forgotten Password on the S4810
- Recovering from a Failed Start on the S4810

## 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remove a command from the list of available commands in EXEC mode.</td>
<td>`privilege exec level level {command</td>
<td></td>
</tr>
<tr>
<td>Move a command from EXEC Privilege to EXEC mode.</td>
<td>`privilege exec level level {command</td>
<td></td>
</tr>
<tr>
<td>Allow access to CONFIGURATION mode.</td>
<td><code>privilege exec level level configure</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
The configuration in the following example 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
- allows access to INTERFACE and LINE modes are allowed with no commands

```
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#
```

- 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
```

- 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
```

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allow access to INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode. Specify all keywords in the command.</td>
<td>privilege configure level level</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Allow access to a CONFIGURATION, INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode command.</td>
<td>privilege {configure</td>
<td>interface</td>
</tr>
</tbody>
</table>

Task Command Syntax Command Mode

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allow access to INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode. Specify all keywords in the command.</td>
<td>privilege configure level level</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Allow access to a CONFIGURATION, INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode command.</td>
<td>privilege {configure</td>
<td>interface</td>
</tr>
</tbody>
</table>

The configuration in the following example 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
- allows access to INTERFACE and LINE modes are allowed with no commands

```
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#
```

- 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
```

- 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
```

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allow access to INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode. Specify all keywords in the command.</td>
<td>privilege configure level level</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Allow access to a CONFIGURATION, INTERFACE, LINE, ROUTE-MAP, and/or ROUTER mode command.</td>
<td>privilege {configure</td>
<td>interface</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure a privilege level for a user.</td>
<td>username username privilege level</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**Apply a Privilege Level to a Terminal Line**

To set a privilege level for a terminal line:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure a privilege level for a terminal line.</td>
<td>privilege level level</td>
<td>LINE</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disable all logging except on the console.</td>
<td>no logging on</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Disable logging to the logging buffer.</td>
<td>no logging buffer</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Disable logging to terminal lines.</td>
<td>no logging monitor</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Disable console logging.</td>
<td>no logging console</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

Log Messages in the Internal Buffer

All error messages, except those beginning with %BOOTUP (Message), 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
• Send System Messages to a Syslog Server
## 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disable all logging except on the console.</td>
<td>no logging on</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Disable logging to the logging buffer.</td>
<td>no logging buffer</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Disable logging to terminal lines.</td>
<td>no logging monitor</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Disable console logging.</td>
<td>no logging console</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

## Send System Messages to a Syslog Server

Send system messages to a syslog server by specifying the server with the following command:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify the server to which you want to send system messages. You can configure up to eight syslog servers.</td>
<td>logging {ip-address</td>
<td>hostname}</td>
</tr>
</tbody>
</table>

## 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/ftos.log
- on a 5.7 SunOS UNIX system, add the line: local7.debugging /var/adm/ftos.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 settings of the system logging by changing the severity level and the storage location. The default is to log all messages up to debug level, that is, all system messages. By changing the severity level in the logging commands, you control the number of system messages logged.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify the minimum severity level for logging to the logging buffer.</td>
<td>logging buffered level</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Specify the minimum severity level for logging to the console.</td>
<td>logging console level</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Specify the minimum severity level for logging to terminal lines.</td>
<td>logging monitor level</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Specifying the minimum severity level for logging to a syslog server.</td>
<td>logging trap level</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Specify the minimum severity level for logging to the syslog history table.</td>
<td>logging history level</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify the size of the logging buffer.</td>
<td>logging buffered size</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td><strong>Note</strong>: 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.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specify the number of messages that FTOS saves to its logging history table.</td>
<td>logging history size size</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

To change one of the settings for logging system messages, use any or all of the following commands in the CONFIGURATION mode:

To view the logging buffer and configuration, use the show logging command in the EXEC privilege mode as shown in the example for Display the Logging Buffer and the Logging Configuration.

To change the severity level of messages logged to a syslog server, use the following command in the CONFIGURATION mode:

To view the logging configuration, use the show running-config logging command in the EXEC privilege mode as shown in the example for Configure a UNIX logging facility level.

Display the Logging Buffer and the Logging Configuration

Display the current contents of the logging buffer and the logging settings for the system, use the show logging command in the EXEC privilege mode as shown in the example below.

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

To view any changes made, use the `show running-config logging` command in the EXEC privilege mode as shown in the example for Configure a UNIX logging facility level.

**Configure a UNIX logging facility level**

You can save system log messages with a UNIX system logging facility.
To configure a UNIX logging facility level, use the following command in the CONFIGURATION mode:

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

To view nondefault settings, use the `show running-config logging` command in the EXEC mode as shown in the example below.

```
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 FTOS to filter and consolidate the system messages for a specific line by synchronizing the message output. Only the messages with a severity at or below the set level appear. This feature works on the terminal and console connections available on the system.

To synchronize log messages, use these commands in the following sequence starting in the CONFIGURATION mode:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 1    | line {console 0 | vty number [end-number] | aux 0} | CONFIGURATION | 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 and an end-number. |
| 2    | logging synchronous [level severity-level | all] [limit] | LINE | Configure a level and set the maximum number of messages to be printed. Configure the following optional parameters:  
  • 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. |

To view the logging synchronous configuration, use the show config command in the LINE mode.

Enable timestamp on syslog messages

By default, syslog messages 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, use the following command syntax in the CONFIGURATION mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| service timestamps [log | debug] [datetime [localtime] [msec] [show-timezone] | uptime] | CONFIGURATION | 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 last boot. If neither parameter is specified, FTOS configures uptime. |
To view the configuration, use the `show running-config logging` command in the EXEC privilege mode.

To disable time stamping on syslog messages, enter `no service timestamps [log | debug]`.

## File Transfer Services

With FTOS, you can configure the system to transfer files over the network using File Transfer Protocol (FTP). One FTP application is copying the system image files over an interface on to the system; however, FTP is not supported on VLAN interfaces.

For more information on FTP, refer to RFC 959, *File Transfer Protocol*.

> **Note:** To transmit large files, Dell Force10 recommends configuring the switch as an FTP server.

### Configuration Task List for File Transfer Services

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

- Enable FTP server (mandatory)
- Configure FTP server parameters (optional)
- Configure FTP client parameters (optional)

### Enable FTP server

To enable the system as an FTP server, use the following command in the CONFIGURATION mode:

```
Command Syntax     Command Mode     Purpose
-----------------------------------------------------------------------
ftp-server enable  CONFIGURATION    Enable FTP on the system.
```

To view FTP configuration, use the `show running-config ftp` command in the EXEC privilege mode as shown in the example below.

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

### Configure FTP server parameters

After the FTP server is enabled on the system, you can configure different parameters.
To configure FTP server parameters, use any or all of the following commands in the CONFIGURATION mode:

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

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

To view the FTP configuration, use the `show running-config ftp` command in EXEC privilege mode.

### Configure FTP client parameters

To configure FTP client parameters, use the following commands in the CONFIGURATION mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ip ftp source-interface interface</code></td>
<td>CONFIGURATION</td>
<td>Enter the following keywords and slot/port or number information:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a loopback interface, enter the keyword loopback followed by a number between 0 and 16383.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a SONET interface, enter the keyword sonet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10-Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a VLAN interface, enter the keyword vlan followed by a number from 1 to 4094.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 40-Gigabit Ethernet interface, enter the keyword fortyGigE followed by the slot/port information.</td>
</tr>
<tr>
<td><code>ip ftp password password</code></td>
<td>CONFIGURATION</td>
<td>Configure a password.</td>
</tr>
<tr>
<td><code>ip ftp username name</code></td>
<td>CONFIGURATION</td>
<td>Enter username to use on FTP client.</td>
</tr>
</tbody>
</table>

To view FTP configuration, use the `show running-config ftp` command in the EXEC privilege mode as shown in the example for Enable FTP server.
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

Dell Force10 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apply an ACL to a VTY line.</td>
<td>ip access-class access-list</td>
<td>LINE</td>
</tr>
</tbody>
</table>

To view the configuration, enter the `show config` command in the LINE mode, as shown in the example below.

```
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 line 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 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>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.</td>
<td>`aaa authentication login {method-list-name</td>
<td>default} [method-1] [method-2] [method-3] [method-4] [method-5] [method-6]`</td>
</tr>
<tr>
<td>2</td>
<td>Apply the method list from Step 1 to a terminal line.</td>
<td>`login authentication {method-list-name</td>
<td>default}`</td>
</tr>
<tr>
<td>3</td>
<td>If you used the line authentication method in the method list you applied to the terminal line, configure a password for the terminal line.</td>
<td><code>password</code></td>
<td>LINE</td>
</tr>
</tbody>
</table>

In the example below, VTY lines 0-2 use a single authentication method, `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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set the number of minutes and seconds.</td>
<td>exec-timeout minutes [seconds]</td>
<td>LINE</td>
</tr>
<tr>
<td>Default: 10 minutes on console, 30 minutes on VTY. Disable EXEC timeout by setting the timeout period to 0.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return to the default timeout values.</td>
<td>no exec-timeout</td>
<td>LINE</td>
</tr>
</tbody>
</table>

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

```
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:

Note: The system allows 120 telnet sessions per minute, allowing the login and logout of 10 telnet sessions 12 times in a minute. If the system reaches this non-practical limit, the telnet service will be stopped for 10 minutes. Console and SSH service may be used to access the system during downtime.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
<td>telnet-peer-rpm</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>
| 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. | telnet [ip-address] | EXEC Privilege |

```
FTOS# telnet 10.11.80.203
Trying 10.11.80.203...
Connected to 10.11.80.203.
Exit character is '^['].
```
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.

```
FTOS(conf)#configuration mode exclusive auto
BATMAN(conf)#exit
3d23h35m: %RPM0-P:CP %SYS-5-CONFIG_I: Configured from console by console
FTOS#config
! Locks configuration mode exclusively.
FTOS(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)
```
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 on the S4810

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Log onto the system via console.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Power-cycle the chassis by switching off all of the power modules and then switching them back on.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Hit any key to abort the boot process. You enter uBoot immediately, as indicated by the <code>=&gt;</code> prompt.</td>
<td>hit any key</td>
<td>(during bootup)</td>
</tr>
<tr>
<td>4</td>
<td>Set the system parameters to ignore the startup configuration file when the system reloads.</td>
<td><code>setenv stconfigignore true</code></td>
<td>uBoot</td>
</tr>
<tr>
<td>5</td>
<td>To save the changes use the <code>saveenv</code> command.</td>
<td><code>saveenv</code></td>
<td>uBoot</td>
</tr>
<tr>
<td>6</td>
<td>Reload the system.</td>
<td><code>reset</code></td>
<td>uBoot</td>
</tr>
<tr>
<td>7</td>
<td>Copy startup-config.bak to the running config.</td>
<td><code>copy flash://startup-config.bak running-config</code></td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>8</td>
<td>Remove all authentication statements you might have for the console.</td>
<td><code>no authentication login</code> <code>no password</code></td>
<td>LINE</td>
</tr>
</tbody>
</table>
Recovering from a Forgotten Enable Password on the S4810

If you forget the enable password:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Log onto the system via console.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Power-cycle the chassis by switching off all of the power modules and then switching them back on.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Hit any key to abort the boot process. You enter uBoot immediately, as indicated by the =&gt; prompt.</td>
<td>hit any key</td>
<td>(during bootup)</td>
</tr>
<tr>
<td>4</td>
<td>Set the system parameters to ignore the enable password when the system reloads.</td>
<td>setenv enablepwdignore true</td>
<td>uBoot</td>
</tr>
<tr>
<td>5</td>
<td>Reload the system.</td>
<td>reset</td>
<td>uBoot</td>
</tr>
<tr>
<td>6</td>
<td>Configure a new enable password.</td>
<td>enable {secret</td>
<td>password}</td>
</tr>
<tr>
<td>7</td>
<td>Save the running-config to the startup-config.</td>
<td>copy running-config startup-config</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

Recovering from a Failed Start on the S4810

A system that does not start correctly might be attempting to boot from a corrupted FTOS image or from a mis-specified location. In that case, you can restart the system and interrupt the boot process to point the system to another boot location. Use the setenv command, as described below. For details on the setenv command, its supporting commands, and other commands that can help recover from a failed start, see the Boot User chapter in the FTOS Command Line Reference for the S4810.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Power-cycle the chassis (pull the power cord and reinsert it).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Hit any key to abort the boot process. You enter uBoot immediately, as indicated by the =&gt; prompt.</td>
<td>hit any key</td>
<td>(during bootup)</td>
</tr>
<tr>
<td>Step</td>
<td>Task</td>
<td>Command Syntax</td>
<td>Command Mode</td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>3</td>
<td>Assign the new location to the FTOS image to be used when the system reloads.</td>
<td>`setenv [primary_image f10boot location</td>
<td>secondary_image f10boot location</td>
</tr>
<tr>
<td>4</td>
<td>Assign an IP address to the Management Ethernet interface.</td>
<td><code>setenv ipaddre address</code></td>
<td>uBoot</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Assign an IP address as the default gateway for the system.</td>
<td><code>setenv gatewayip address</code></td>
<td>uBoot</td>
</tr>
<tr>
<td>7</td>
<td>Reload the system.</td>
<td><code>reset</code></td>
<td>uBoot</td>
</tr>
</tbody>
</table>
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 with regard to OAM for Metro/Carrier Ethernet 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 the illustration below.

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.

**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.

**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.
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. Refer to Configure Ingress Layer 2 ACL Sub-partitions.
2. Enable Ethernet CFM.
3. Create a Maintenance Domain.
4. Create a Maintenance Association.
5. Create Maintenance Points.
6. Use CFM tools:
   a. Continuity Check Messages
   b. Loopback Message and Response
   c. Linktrace Message and Response

Related Configuration Tasks

- Enable CFM SNMP Traps.
- Display Ethernet CFM Statistics
Enable Ethernet CFM

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spawn the CFM process. No CFM configuration is allowed until the CFM</td>
<td>ethernet cfm</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>process is spawned.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disable Ethernet CFM without stopping the CFM process.</td>
<td>disable</td>
<td>ETHERNET CFM</td>
</tr>
</tbody>
</table>

Create a Maintenance Domain

Connectivity Fault Management (CFM) divides a network into hierarchical maintenance domains, as shown in the illustration in Maintenance Domains.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create maintenance domain.</td>
<td>domain name md-level number Range: 0-7</td>
<td>ETHERNET CFM</td>
</tr>
<tr>
<td>2</td>
<td>Display maintenance domain information.</td>
<td>show ethernet cfm domain [name</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

FTOS# show ethernet cfm domain

<table>
<thead>
<tr>
<th>Domain Name: customer</th>
<th>Level: 7</th>
<th>Total Service: 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Services</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA-Name</td>
<td>VLAN</td>
<td>CC-Int</td>
</tr>
<tr>
<td>My_MA</td>
<td>200</td>
<td>10s</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Domain Name: praveen</th>
<th>Level: 6</th>
<th>Total Service: 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Services</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA-Name</td>
<td>VLAN</td>
<td>CC-Int</td>
</tr>
<tr>
<td>Your_MA</td>
<td>100</td>
<td>10s</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create maintenance association.</td>
<td>service name vlan vlan-id</td>
<td>ECFM DOMAIN</td>
</tr>
</tbody>
</table>

Create 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 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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create an MEP.</td>
<td>ethernet cfm mep {up-mep</td>
<td>down-mep} domain {name</td>
</tr>
<tr>
<td>Display configured MEPs and MIPs.</td>
<td>show ethernet cfm maintenance-points local [mep</td>
<td>mip]</td>
</tr>
</tbody>
</table>
**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.

**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

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display the MEP Database.</td>
<td>show ethernet cfm maintenance-points remote detail [active</td>
<td>domain {level</td>
</tr>
</tbody>
</table>

```
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: Force10
- MEP Interface status: Up
- MEP Port status: Forwarding
- Receive RDI: FALSE
- MP Status: Active

Display the MIP Database.  show ethernet cfm mipdb

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set the amount of time that data from a missing MEP is kept in the Continuity Check Database.</td>
<td>database hold-time minutes Default: 100 minutes Range: 100-65535 minutes</td>
<td>ECFM DOMAIN</td>
</tr>
</tbody>
</table>

**MP Database Persistence**

**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 process 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, "Continuity Check Message Processing," in 802.1ag.

### Table 5-1. Continuity Check Message Processing

<table>
<thead>
<tr>
<th>Frames at</th>
<th>Frames from</th>
<th>UP-MEP Action</th>
<th>Down-MEP Action</th>
<th>MIP Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than my level</td>
<td>Bridge-relay side or Wire side</td>
<td>Drop</td>
<td>Drop</td>
<td>Drop</td>
</tr>
<tr>
<td>My level</td>
<td>Bridge-relay side</td>
<td>Consume</td>
<td>Drop</td>
<td>Add to MIP-DB and forward</td>
</tr>
<tr>
<td></td>
<td>Wire side</td>
<td>Drop</td>
<td>Consume</td>
<td></td>
</tr>
<tr>
<td>Greater than my level</td>
<td>Bridge-relay side or Wire side</td>
<td>Forward</td>
<td>Forward</td>
<td>Forward</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enable CCM.</td>
<td>no ccm disable</td>
<td>ECFM DOMAIN</td>
</tr>
<tr>
<td></td>
<td>Default: Disabled</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Configure the transmit interval (mandatory). The interval specified applies to all MEPs in the domain.</td>
<td>ccm transmit-interval seconds</td>
<td>ECFM DOMAIN</td>
</tr>
<tr>
<td></td>
<td>Default: 10 seconds</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Enable Cross-checking

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

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Send a Loopback message.</td>
<td>ping ethernet domain name ma-name ma-name remote</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td></td>
<td>mep-id</td>
<td>mac-addr mac-address</td>
</tr>
</tbody>
</table>

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.
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.

### 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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Send a Linktrace message. Since the LTM is a Multicast message sent to the entire ME, there is no need to specify a destination.</td>
<td>traceroute ethernet domain</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Enable Link Trace caching.</td>
<td>traceroute cache</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Set the amount of time a trace result is cached.</td>
<td>traceroute cache hold-time minutes Default: 100 minutes Range: 10-65535 minutes</td>
<td>ETHERNET CFM</td>
</tr>
<tr>
<td>Set the size of the Link Trace Cache.</td>
<td>traceroute cache size entries Default: 100 Range: 1 - 4095 entries</td>
<td>ETHERNET CFM</td>
</tr>
<tr>
<td>Display the Link Trace Cache.</td>
<td>show ethernet cfm traceroute-cache</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Hops</th>
<th>Host</th>
<th>IngressMAC</th>
<th>Ingr Action</th>
<th>Relay Action</th>
<th>FWD Status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>00:00:00:01:e8:53:4a:f8</td>
<td>00:01:e8:52:4a:f8</td>
<td>IngOK</td>
<td>RlyHit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>00:00:00:01:e8:52:4a:f8</td>
<td></td>
<td>Terminal MEP</td>
<td></td>
</tr>
</tbody>
</table>

Delete all Link Trace Cache entries.
clear ethernet cfm traceroute-cache EXEC Privilege
Enable CFM SNMP Traps.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable SNMP trap messages for Ethernet CFM.</td>
<td>snmp-server enable traps cfm</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

A Trap is sent only when one of the five highest priority defects occur, as shown in Table 5-2, "ECFM SNMP Traps," in 802.1ag.

Table 5-2. ECFM SNMP Traps

<table>
<thead>
<tr>
<th>Cross-connect defect</th>
<th>%ECFM-5-ECFM_XCON_ALARM: Cross connect fault detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error-CCM defect</td>
<td>%ECFM-5-ECFM_ERROR_ALARM: Error CCM Defect detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000</td>
</tr>
<tr>
<td>MAC Status defect</td>
<td>%ECFM-5-ECFM_MAC_STATUS_ALARM: MAC Status Defect detected by MEP 1 in Domain provider at Level 4 VLAN 3000</td>
</tr>
<tr>
<td>Remote CCM defect</td>
<td>%ECFM-5-ECFM_REMOTE_ALARM: Remote CCM Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000</td>
</tr>
<tr>
<td>RDI defect</td>
<td>%ECFM-5-ECFM_RDI_ALARM: RDI Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000</td>
</tr>
</tbody>
</table>

Three values are providing 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

<table>
<thead>
<tr>
<th>MPID</th>
<th>Domain Name</th>
<th>Level</th>
<th>Type</th>
<th>Port</th>
<th>CCM-Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>cfm0</td>
<td>7</td>
<td>MEP</td>
<td>Gi 4/10</td>
<td>Enabled</td>
</tr>
<tr>
<td></td>
<td>test0</td>
<td>10</td>
<td>DOWN</td>
<td>00:01:e8:59:23:45</td>
<td></td>
</tr>
</tbody>
</table>

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

Domain Name: My_Name
MD Index: 1
Level: 0
Total Service: 1
Services

<table>
<thead>
<tr>
<th>MA-Index</th>
<th>MA-Name</th>
<th>VLAN</th>
<th>CC-Int</th>
<th>X-CHK Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>test</td>
<td>0</td>
<td>1s</td>
<td>enabled</td>
</tr>
</tbody>
</table>

Domain Name: Your_Name
MD Index: 2
Level: 2
Total Service: 1
Services

<table>
<thead>
<tr>
<th>MA-Index</th>
<th>MA-Name</th>
<th>VLAN</th>
<th>CC-Int</th>
<th>X-CHK Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>test</td>
<td>100</td>
<td>1s</td>
<td>enabled</td>
</tr>
</tbody>
</table>
## Display Ethernet CFM Statistics

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display MEP CCM statistics.</td>
<td>show ethernet cfm statistics [domain {name</td>
<td>level} vlan-id vlan-id mpid mpid]</td>
</tr>
</tbody>
</table>

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

Display CFM statistics by port.

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
802.1X

802.1X is supported on platforms: \([E \ C \ S \ Z \ S4810]\)

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 packets on the network until its identity can be verified (through a username and password, for example). 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.

![Diagram of EAP frame encapsulation](image)

The illustration above and the illustration below show how EAP frames are encapsulated in Ethernet and RADIUS frames.
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 authenticator is the gatekeeper 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. **Note**: FTOS supports 802.1X with EAP-MD5, EAP-OTP, EAP-TLS, EAP-TTLS, PEAPv0, PEAPv1, and MS-CHAPv2 with PEAP.
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

---

**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.
RADIUS Attributes for 802.1 Support

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

- **Attribute 31—Calling-station-id**: relays the supplicant MAC address to the authentication server.
- **Attribute 41—NAS-Port-Type**: NAS-port physical port type. 15 indicates Ethernet.
- **Attribute 61—NAS-Port**: the physical port number by which the authenticator is connected to the supplicant.
- **Attribute 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 one-step process:

1. **Enabling 802.1X**.

Related Configuration Tasks

- Configuring Request Identity Re-transmissions
- Forcibly Authorizing or Unauthorizing a Port
- Re-authenticating a Port
- Configuring Timeouts
- Configuring a Guest VLAN
- Configuring an Authentication-fail VLAN
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.
- 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.

Enabling 802.1X

802.1X must be enabled globally.

To enable 802.1X:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enable 802.1X globally.</td>
<td>dot1x authentication</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Enter INTERFACE mode on an interface or a range of interfaces.</td>
<td>interface [range]</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>3</td>
<td>Enable 802.1X on the supplicant interface only.</td>
<td>dot1x authentication</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

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 the example below.

FTOS#show running-config | find dot1x
dot1x authentication
!
[output omitted]
!
interface TenGigabitEthernet 2/1
no ip address
dot1x authentication
no shutdown
!
FTOS#

View 802.1X configuration information for an interface using the command show dot1x interface, as shown in the example below.

FTOS#show dot1x interface TenGigabitEthernet 2/1

802.1x information on Te 2/1:
-----------------------------
Dot1x Status: Enable
Port Control: AUTO
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
Mac-Auth-Bypass: Disable
Mac-Auth-Bypass Only: Disable
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: 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, or there might be a physical layer problem.
To configure the amount of time that the authenticator waits before re-transmitting an EAP Request Identity frame:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure the amount of time that the authenticator waits before re-transmitting an EAP Request Identity frame.</td>
<td>dot1x tx-period number</td>
<td>INTERFACE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range: 1 - 65535 (1 year)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default: 30</td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure a maximum number of times that a Request Identity frame can be re-transmitted by the authenticator.</td>
<td>dot1x max-eap-req number</td>
<td>INTERFACE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range: 1- 10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default: 2</td>
<td></td>
</tr>
</tbody>
</table>

The example in Configuring a Quiet Period after a Failed Authentication 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure the amount of time that the authenticator waits to re-transmit a Request Identity frame after a failed authentication.</td>
<td>dot1x quiet-period seconds</td>
<td>INTERFACE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range: 1- 65535</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default: 60</td>
<td></td>
</tr>
</tbody>
</table>

The example below 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

FTOS(conf-if-range-Te-0/0)#dot1x tx-period 90
FTOS(conf-if-range-Te-0/0)#dot1x max-eap-req 10
FTOS(conf-if-range-Te-0/0)#dot1x quiet-period 120
FTOS#show dot1x interface TenGigabitEthernet 2/1
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.

- **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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Place a port in the ForceAuthorized, ForceUnauthorized, or Auto state.</td>
<td>dot1x port-control {force-authorized</td>
<td>force-unauthorized</td>
</tr>
</tbody>
</table>

The example below shows configuration information for a port that has been force-authorized.

FTOS(conf-if-Te-0/0)#dot1x port-control force-authorized
FTOS(conf-if-Te-0/0)#show dot1x interface TenGigabitEthernet 0/0
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-authenticates 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure the authenticator to periodically re-authenticate the supplicant.</td>
<td><code>dot1x reauthentication [interval] seconds</code>&lt;br&gt;Range: 1-65535&lt;br&gt;Default:3600</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>
To configure a maximum number of re-authentications:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure the maximum number of times that the supplicant can be reauthenticated.</td>
<td><code>dot1x reauth-max number</code></td>
<td>INTERFACE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range: 1-10&lt;br&gt;Default: 2</td>
<td></td>
</tr>
</tbody>
</table>

```
FTOS(conf-if-Te-0/0)#dot1x reauthentication interval 7200
FTOS(conf-if-Te-0/0)#dot1x reauth-max 10
FTOS(conf-if-Te-0/0)#do show dot1x interface TenGigabitEthernet 0/0
```

802.1x information on Te 0/0:
---------------------------------
Dot1x Status:       Enable
Port Control:       FORCE_AUTHORIZED
Port Auth Status:   UNAUTHORIZED
Re-Authentication:  Enable
Untagged VLAN id:   None
Tx Period:          90 seconds
Quiet Period:       120 seconds
ReAuth Max:         10
Supplicant Timeout: 30 seconds
Server Timeout:     30 seconds
Re-Auth Interval:   7200 seconds
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. This amount of time that the authenticator waits for a response can be configured.

To terminate the authentication process due to an unresponsive supplicant:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Terminate the authentication process due to an unresponsive supplicant.</td>
<td><code>dot1x supplicant-timeout seconds</code></td>
<td>INTERFACE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range: 1-300&lt;br&gt;Default: 30</td>
<td></td>
</tr>
</tbody>
</table>
To terminate the authentication process due to an unresponsive authentication server:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Terminate the authentication process due to an unresponsive</td>
<td>dot1x server-timeout seconds</td>
<td>INTERFACE</td>
</tr>
<tr>
<td></td>
<td>unresponsive authentication server.</td>
<td>Range: 1-300</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default: 30</td>
<td></td>
</tr>
</tbody>
</table>

The example below shows configuration information for a port for which the authenticator terminates the authentication process for an unresponsive supplicant or server after 15 seconds.

FTOS(conf-if-Te-0/0)#dot1x port-control force-authorized
FTOS(conf-if-Te-0/0)#do show dot1x interface TenGigabitEthernet 0/0

802.1x information on Te 0/0:
---
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
FTOS supports dynamic VLAN assignment when using 802.1X. The basis for VLAN assignment is RADIUS attribute 81, Tunnel-Private-Group-ID. Dynamic VLAN assignment uses the 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, and 3) the RADIUS server authenticates the request and returns a RADIUS ACCEPT message with the VLAN assignment using Tunnel-Private-Group-ID.

### Dynamic VLAN Assignment with Port Authentication

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure 802.1x globally (refer to Enabling 802.1X) along with relevant RADIUS server configurations (refer to the illustration in Dynamic VLAN Assignment with Port Authentication).</td>
</tr>
<tr>
<td>2</td>
<td>Make the interface a switchport so that it can be assigned to a VLAN.</td>
</tr>
<tr>
<td>3</td>
<td>Create the VLAN to which the interface will be assigned.</td>
</tr>
<tr>
<td>4</td>
<td>Connect the supplicant to the port configured for 802.1X.</td>
</tr>
<tr>
<td>5</td>
<td>Verify that the port has been authorized and placed in the desired VLAN (refer to the illustration in Dynamic VLAN Assignment with Port Authentication).</td>
</tr>
</tbody>
</table>

The illustration below shows the configuration on the Dell Force10 system before connecting the end-user device in black and blue text, and after connecting the device in red text. The blue text corresponds to the preceding numbered steps on dynamic VLAN assignment with 802.1X.
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 within a determined amount of time \((\text{reauth-max} + 1) \times \text{tx-period}\), (refer to Configuring Timeouts) 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 the example below.

```
FTOS(conf-if-Te-2/1)#dot1x guest-vlan 200
FTOS(conf-if-Te 2/1))#show config

! interface TenGigabitEthernet 21
    switchport
    dot1x guest-vlan 200
    no shutdown
FTOS(conf-if-Te 2/1))#
```

View your configuration using the command `show config` from INTERFACE mode, as shown in the example above, or using the command `show dot1x interface` command from EXEC Privilege mode as shown in the second example below.
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). 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 the example below. Configure the maximum number of authentication attempts by the authenticator using the keyword `max-attempts` with this command.

```
FTOS(conf-if-Te-2/1)#dot1x guest-vlan 200
FTOS(conf-if-Te 2/1)#show config

interface TenGigabitEthernet 2/1
switchport
dot1x authentication
dot1x guest-vlan 200
no shutdown
FTOS(conf-if-Te-2/1)#

FTOS(conf-if-Te-2/1)#dot1x auth-fail-vlan 100 max-attempts 5
FTOS(conf-if-Te-2/1)#show config

interface TenGigabitEthernet 2/1
switchport
dot1x authentication
dot1x guest-vlan 200
dot1x auth-fail-vlan 100 max-attempts 5
no shutdown
FTOS(conf-if-Te-2/1)#
```

View your configuration using the command `show config` from INTERFACE mode, as shown in the example in Configuring a Guest VLAN, or using the command `show dot1x interface` command from EXEC Privilege mode as shown in the example below.

```
FTOS(conf-if-Te 2/1)#dot1x port-control force-authorized
FTOS(conf-if-Te 2/1)#show dot1x interface TenGigabitEthernet 2/1

802.1x information on Te 2/1:
-----------------------------
Dot1x Status:                Enable
Port Control:                FORCE_AUTHORIZED
Port Auth Status:            UNAUTHORIZED
Re-Authentication:           Disable
Untagged VLAN id:            None
Guest VLAN:                  Disabled
Guest VLAN id:               200
Auth-Fail VLAN:              Disabled
```
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
Access Control Lists (ACLs)

This chapter describes the Access Control Lists (ACLs), prefix lists, and route-maps.

Access Control Lists (ACLs) are supported on platforms: E C S Z S4810

**Ingress** IP and MAC ACLs are supported on platforms: E C S Z S4810

**Egress** IP and MAC ACLs are supported on platforms: E S Z S4810

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 Layer 2.

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 Content Addressable Memory (CAM) for complete CAM profiling information.

This chapter covers the following topics:

- IP Access Control Lists (ACLs)
  - CAM Profiling, CAM Allocation, and CAM Optimization
  - Implementing ACLs on FTOS
- IP Fragment Handling
- Configure a standard IP ACL
- Configure an extended IP ACL
- Configuring Layer 2 and Layer 3 ACLs on an Interface
- Assign an IP ACL to an Interface
- Configuring Ingress ACLs
- Configuring Egress ACLs
• Configuring ACLs to Loopback
  • Applying an ACL on Loopback Interfaces
• IP Prefix Lists
• ACL Resequencing
• Route Maps

**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 and on all platforms.

**Note:** Hot Lock ACLs are supported for 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** and **S**

**CAM optimization** is supported on platforms **C** **S**
CAM Profiling

**CAM optimization** is supported on platforms E T.

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 7-1 lists the sub-partition and the percentage of the Layer 2 ACL CAM partition that FTOS allocates to each by default.

### Table 7-1. Layer 2 ACL CAM Sub-partition Sizes

<table>
<thead>
<tr>
<th>Partition</th>
<th>% Allocated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sysflow</td>
<td>6</td>
</tr>
<tr>
<td>L2ACL</td>
<td>14</td>
</tr>
<tr>
<td>*PVST</td>
<td>50</td>
</tr>
<tr>
<td>QoS</td>
<td>12</td>
</tr>
<tr>
<td>L2PT</td>
<td>13</td>
</tr>
<tr>
<td>FRRP</td>
<td>5</td>
</tr>
</tbody>
</table>

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 and S4810.

Allocate space for IPV6 ACLs on the 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. The example below gives a sample of the output shown when executing the command. The status column indicates whether or not the policy can be enabled.

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

<table>
<thead>
<tr>
<th>Linecard</th>
<th>Portpipe</th>
<th>CAM Partition</th>
<th>Available CAM</th>
<th>Estimated CAM per Port</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>IPv4Flow</td>
<td>232</td>
<td>0</td>
<td>Allowed</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>IPv6Flow</td>
<td>0</td>
<td>0</td>
<td>Allowed</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>IPv4Flow</td>
<td>232</td>
<td>0</td>
<td>Allowed</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>IPv6Flow</td>
<td>0</td>
<td>0</td>
<td>Allowed</td>
</tr>
</tbody>
</table>

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 the example below, 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 the example below. 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.

```
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
```
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 to set up extended ACLs.

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

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

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>ip access-list standard access-listname</code></td>
<td>CONFIGURATION</td>
<td>Enter IP ACCESS LIST mode by naming a standard IP access list.</td>
</tr>
<tr>
<td>2</td>
<td>`seq sequence-number {deny</td>
<td>permit} [source [mask]</td>
<td>any</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• log and monitor options are supported on E-Series only.</td>
</tr>
</tbody>
</table>
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 in EXEC Privilege mode as shown in the example below.

```
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#
```

The example below 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.

```
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ip access-list standard access-list-name</td>
<td>CONFIGURATION</td>
<td>Create a standard IP ACL and assign it a unique name.</td>
</tr>
<tr>
<td>2</td>
<td>{deny</td>
<td>permit} {source [mask]</td>
<td>any</td>
</tr>
<tr>
<td></td>
<td>[log] [order] [monitor] [fragments]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

The example below 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.

FTOS(config-route-map)#ip access standard kigali
FTOS(config-std-nacl)#permit 10.1.0.0/16
FTOS(config-std-nacl)#show config

![Command Output]

To view all configured IP ACLs, use the show ip accounting access-list command in the EXEC Privilege mode as shown in the example below.

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 an extended ACL 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>ip access-list extended access-list-name</code></td>
<td>CONFIGURATION</td>
<td>Enter the IP ACCESS LIST mode by creating an extended IP ACL.</td>
</tr>
<tr>
<td>2</td>
<td><code>seq sequence-number</code>&lt;br&gt;`{deny</td>
<td>permit}<code>&lt;br&gt;</code>{ip-protocol-number</td>
<td>icmp</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>ip access-list extended access-list-name</code></td>
<td>CONFIGURATION</td>
<td>Create an extended IP ACL and assign it a unique name.</td>
</tr>
</tbody>
</table>
| 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>ip access-list extended access-list-name</code></td>
<td>CONFIGURATION</td>
<td>Create a extended IP ACL and assign it a unique name.</td>
</tr>
</tbody>
</table>
| 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.
The following example 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.

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 config

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>`{deny</td>
<td>permit} {source mask</td>
<td>any</td>
</tr>
<tr>
<td>`{deny</td>
<td>permit} tcp {source mask</td>
<td>any</td>
</tr>
<tr>
<td>`{deny</td>
<td>permit} udp {source mask</td>
<td>any</td>
</tr>
</tbody>
</table>

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.

The following example 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.

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 in the EXEC Privilege mode as shown in the first example in Configure a standard IP ACL.

**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 instead 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 7-2. L2 and L3 ACL Filtering on Switched Packets**

<table>
<thead>
<tr>
<th>L2 ACL Behavior</th>
<th>L3 ACL Behavior</th>
<th>Decision on Targeted Traffic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deny</td>
<td>Deny</td>
<td>Denied by L3 ACL</td>
</tr>
<tr>
<td>Deny</td>
<td>Permit</td>
<td>Permitted by L3 ACL</td>
</tr>
</tbody>
</table>
Assign an IP ACL to an Interface

Ingress IP ACLs are supported on platforms: C
Ingress and Egress IP ACLs are supported on platform: E and S

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:

- Configuring Ingress ACLs
- Configuring Egress ACLs
- Configuring ACLs to Loopback

For more information on Layer-3 interfaces, refer to Interfaces.

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interface interface slot/port</td>
<td>CONFIGURATION</td>
<td>Enter the interface number.</td>
</tr>
<tr>
<td>2</td>
<td>ip address ip-address</td>
<td>INTERFACE</td>
<td>Configure an IP address for the interface, placing it in Layer-3 mode.</td>
</tr>
</tbody>
</table>
To view which IP ACL is applied to an interface, use the `show config` command in the INTERFACE mode as shown below or the `show running-config` command in the EXEC 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)#
```

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

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Create an ACL that uses rules with the count option. See Configure a standard IP ACL</td>
</tr>
<tr>
<td>2</td>
<td>Apply the ACL as an inbound or outbound ACL on an interface. See Assign an IP ACL to an Interface</td>
</tr>
<tr>
<td>3</td>
<td>View the number of packets matching the ACL using the <code>show ip accounting access-list</code> command from EXEC Privilege mode.</td>
</tr>
</tbody>
</table>
**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 in the EXEC Privilege mode as shown below. This example also shows applying the ACL, applying rules to the newly created access group, and viewing the access list:

```
FTOS(conf)#interface gige 0/0
FTOS(conf-if-gige0/0)#ip access-group abcd in
FTOS(conf-if-gige0/0)#show config
!
gigethernet 0/0
 no ip address
 ip access-group abcd in
 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
 seq 15 permit 1.1.1.2
```

**Configuring Egress ACLs**

**Egress ACLs** are supported on platforms E and 54810.

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 in the EXEC Privilege mode as shown in the example below. This example also shows viewing the configuration, applying rules to the newly created access group, and viewing the access list:

```
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
  seq 15 permit 1.1.1.2
```

### 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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apply Egress ACLs to IPv4 system traffic.</td>
<td><code>ip control-plane [egress filter]</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Apply Egress ACLs to IPv6 system traffic.</td>
<td><code>ipv6 control-plane [egress filter]</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Create a Layer 3 ACL using permit rules with the <code>count</code> option to describe the desired CPU traffic</td>
<td>`permit ip {source mask</td>
<td>any</td>
</tr>
</tbody>
</table>

**Note:** The `ip control-plane [egress filter]` and the `ipv6 control-plane [egress filter]` commands are not supported on S4810 systems.
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interface loopback 0</td>
<td>CONFIGURATION</td>
<td>Only loopback 0 is supported for the loopback ACL.</td>
</tr>
<tr>
<td>2</td>
<td>ip access-list [standard</td>
<td>CONFIGURATION</td>
<td>Apply rules to the new ACL.</td>
</tr>
<tr>
<td></td>
<td>extended] name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ip access-group name in</td>
<td>INTERFACE</td>
<td>Apply an ACL to traffic entering loopback.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• in: configure the ACL to filter incoming traffic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Note: ACLs for loopback can only be applied to incoming traffic.</td>
</tr>
</tbody>
</table>
To apply ACLs on loopback, use the **ip access-group** command in the INTERFACE mode as shown in the example below. This example also shows the interface configuration status, adding rules to the access group, and displaying the list of rules in the ACL:

```plaintext
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
  permit 1.1.1.2

Note: Refer to the section VTY Line Local Authentication and Authorization in the Security chapter.

### IP Prefix Lists

Prefix Lists are supported on platforms: ![Supported Platforms]

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
- Use a prefix list for route redistribution

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:

- `ip prefix-list prefix-name`

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>ip prefix-list prefix-name</code></td>
<td>CONFIGURATION</td>
<td>Create a prefix list and assign it a unique name. You are in the PREFIX LIST mode.</td>
</tr>
</tbody>
</table>
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.

The example below 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.

```
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>ip prefix-list prefix-name</code></td>
<td>CONFIGURATION</td>
<td>Create a prefix list and assign it a unique name.</td>
</tr>
</tbody>
</table>
The example below 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.

```
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show ip prefix-list detail [prefix-name]</td>
<td>EXEC Privilege</td>
<td>Show detailed information about configured Prefix lists.</td>
</tr>
<tr>
<td>show ip prefix-list summary [prefix-name]</td>
<td>EXEC Privilege</td>
<td>Show a table of summarized information about configured Prefix lists.</td>
</tr>
</tbody>
</table>

```
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)
```
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>router rip</td>
<td>CONFIGURATION</td>
<td>Enter RIP mode</td>
</tr>
<tr>
<td>distribute-list prefix-list-name in [interface]</td>
<td>CONFIG-ROUTER-RIP</td>
<td>Apply a configured prefix list to incoming routes. You can specify an interface. If you enter the name of a nonexistent prefix list, all routes are forwarded.</td>
</tr>
<tr>
<td>distribute-list prefix-list-name out [interface</td>
<td>connected</td>
<td>static</td>
</tr>
</tbody>
</table>

To view the configuration, use the `show config` command in the ROUTER RIP mode as shown in the example below or the `show running-config rip` command in the EXEC 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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>router ospf</td>
<td>CONFIGURATION</td>
<td>Enter OSPF mode</td>
</tr>
</tbody>
</table>
To view the configuration, use the `show config` command in the ROUTER OSPF mode as shown in the example below or the `show running-config ospf` command in the EXEC 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

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 7-3 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 7-4. 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 7-3. ACL Resequencing Example (Insert New Rules)

| seq 5 permit any host 1.1.1.1 |
| seq 6 permit any host 1.1.1.2 |
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 7-5. You must specify the list name, starting number, and increment when using these commands.

```markdown
Table 7-5. Resequencing ACLs and Prefix Lists

<table>
<thead>
<tr>
<th>List</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv4, IPv6, or MAC ACL</td>
<td>resequence access-list {ipv4</td>
<td>ipv6</td>
</tr>
<tr>
<td>IPv4 or IPv6 prefix-list</td>
<td>resequence prefix-list {ipv4</td>
<td>ipv6} {prefix-list-name StartingSeqNum Step-to-Increment}</td>
</tr>
</tbody>
</table>
```

The following example shows the resequencing of an IPv4 access-list beginning with the number 2 and incrementing by 2.

```bash
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
```
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 a rule. These two mechanisms allow remarks to retain their original position in the list.

For example, in the following example, 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.
Route Maps

Route-maps are supported on platforms: C E S Z

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 (mandatory)
• Configure route map filters (optional)
• Configure a route map for route redistribution (optional)
• Configure a route map for route tagging (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:

```
route-map map-name [permit | deny] [sequence-number]
```

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 as shown in the example below.

```
FTOS(config-route-map)#show config
!
routemap 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. The following text shows an example with two instances of a route map.

```
FTOS#show route-map
route-map zakho, permit, sequence 10
  Match clauses:
  Set clauses:
route-map zakho, permit, sequence 20
  Match clauses:
    interface GigabitEthernet 0/1
  Set clauses:
    tag 35
    level stub-area
FTOS#
```

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 as shown in the following example.

```
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#

The following text 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.

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
```
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>match as-path as-path-name</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Match routes with the same AS-PATH numbers.</td>
</tr>
<tr>
<td><code>match community community-list-name [exact]</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Match routes with COMMUNITY list attributes in their path.</td>
</tr>
</tbody>
</table>
**Command Syntax** | **Command Mode** | **Purpose**
---|---|---
**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.
- For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` followed by the slot/port information.

**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.
To configure a set condition, use any or all of the following commands in the ROUTE-MAP mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>match tag tag-value</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Match routes with a specific tag.</td>
</tr>
<tr>
<td><code>set as-path prepend as-number [ ... as-number]</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Add an AS-PATH number to the beginning of the AS-PATH</td>
</tr>
<tr>
<td><code>set automatic-tag</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Generate a tag to be added to redistributed routes.</td>
</tr>
<tr>
<td>`set level {backbone</td>
<td>level-1</td>
<td>level-1-2</td>
</tr>
<tr>
<td><code>set local-preference value</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Specify a value for the BGP route’s LOCAL_PREF attribute.</td>
</tr>
<tr>
<td>`set metric {+</td>
<td>-</td>
<td>metric-value}`</td>
</tr>
<tr>
<td>`set metric-type {external</td>
<td>internal</td>
<td>type-1</td>
</tr>
<tr>
<td><code>set next-hop ip-address</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Assign an IP address as the route’s next hop.</td>
</tr>
<tr>
<td><code>set ipv6 next-hop ip-address</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Assign an IPv6 address as the route’s next hop.</td>
</tr>
<tr>
<td>`set origin {egp</td>
<td>igp</td>
<td>incomplete}`</td>
</tr>
<tr>
<td><code>set tag tag-value</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Specify a tag for the redistributed routes.</td>
</tr>
<tr>
<td><code>set weight value</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Specify a value as the route’s weight.</td>
</tr>
</tbody>
</table>

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 the following example, 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.

```
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
```

**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.

```
routing 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 the following example, 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.

```
! 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. The following example 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.

```
!  
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!
```
Bidirectional Forwarding Detection (BFD)

Bidirectional Forwarding Detection (BFD) is supported only on platforms: E C Z S4810

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. The C-Series and E-Series support BFD on OSPF, IS-IS, VLANs, VRRP, LAGs, and physical ports based on the IETF internet draft document draft-ietf-bfd-base-03. On the S4810, BFD is supported on dynamic routing protocols such as OSPF, IS-IS and BGP.
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. The following illustration shows the complete encapsulation of a BFD control packet inside an IPv4 packet.
Figure 8.1. BFD in IPv4 Packet Format
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:
Bidirectional Forwarding Detection (BFD)

- **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 the illustration below 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 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.

Session state changes

The illustration below 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.
Important Points to Remember

- BFD for line card ports is hitless, but is not hitless for VLANs since they are instantiated on the RPM.
- FTOS supports a maximum of 100 sessions per BFD agent on C-Series and E-Series. 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). On the S4810, FTOS supports 128 sessions per stack unit at 200 minimum transmit and receive intervals with a multiplier of 3, and 64 sessions at 100 minimum transmit and receive intervals with a multiplier of 4.
- 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, IS-IS, (E-Series, Z-Series and S4810 only), BGP (S4810 only), and VRRP (not on S4810) protocols as BFD clients.

Configuring Bidirectional Forwarding Detection

The remainder of this chapter is divided into the following sections:

- Configuring BFD for Physical Ports
- Configuring BFD for Static Routes
- Configuring BFD for OSPF
- Configuring BFD for IS-IS
- Configuring BFD for BGP
- Configuring BFD for VRRP
- Configuring BFD for VLANs
- Configuring BFD for Port-Channels
- Configuring Protocol Liveness
- Troubleshooting BFD

Configuring BFD for Physical Ports

Configuring BFD for Physical Ports is supported on C-Series and E-Series only.

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. Enabling BFD globally.
2. Establish a session with a next-hop neighbor.

Related configuration tasks

- Viewing physical port session parameters.
- Disabling and re-enabling BFD.

Enabling BFD globally

BFD must be enabled globally on both routers, as shown in the illustration in Establishing a session on physical ports.

To enable BFD globally:

```
R1(conf)#bfd enable
R1(conf)#do show running-config bfd

bfd enable
```

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 the illustration below. The configuration parameters do not need to match.
To establish a session:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter interface mode</td>
<td>interface</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Assign an IP address to the interface if one is not already assigned.</td>
<td>ip address ip-address</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

Verify that the session is established using the command `show bfd neighbors`, as shown in the example below.

R1(config-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)

<table>
<thead>
<tr>
<th>LocalAddr</th>
<th>RemoteAddr</th>
<th>Interface</th>
<th>State</th>
<th>Rx-int</th>
<th>Tx-int</th>
<th>Mult</th>
<th>Clients</th>
</tr>
</thead>
<tbody>
<tr>
<td>* 2.2.2.1</td>
<td>2.2.2.2</td>
<td>Gi 4/24</td>
<td>Up</td>
<td>100</td>
<td>100</td>
<td>3</td>
<td>c</td>
</tr>
</tbody>
</table>

The example below for the command `show bfd neighbors detail` shows more specific information about BFD sessions.

R1(config-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:
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)

### Viewing physical port session parameters

BFD sessions are configured with default intervals and a default role (active). Dell Force10 recommends maintaining the default values.

View session parameters using the `show bfd neighbors detail` command.

```
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disable BFD on an interface.</td>
<td>no bfd enable</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

**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 AdDn 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enable BFD on an interface.</td>
<td>bfd enable</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

**Configuring BFD for Static Routes**

Configuring BFD for Static Routes is supported on C-Series and E-Series only.

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. **Enabling BFD globally.**
2. On the local system, establish a session with the next hop of a static route. Refer to Configuring BFD for Static Routes.
3. On the remote system, establish a session with the physical port that is the origin of the static route. Refer to Establishing a session on physical ports.

**Related configuration tasks**

- Changing static route session parameters.
- Disabling BFD for static routes.

**Establishing sessions for static routes**

Sessions are established for all neighbors that are the next hop of a static route.

To establish a BFD session:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Establish BFD sessions for all neighbors that are the next hop of a static route.</td>
<td>ip route bfd</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

Verify that sessions have been created for static routes using the command show bfd neighbors, as shown in the example below.

```
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)

<table>
<thead>
<tr>
<th>LocalAddr</th>
<th>RemoteAddr</th>
<th>Interface</th>
<th>State</th>
<th>Rx-int</th>
<th>Tx-int</th>
<th>Mult</th>
<th>Clients</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.2.1</td>
<td>2.2.2.2</td>
<td>Gi 4/24</td>
<td>Up</td>
<td>100</td>
<td>100</td>
<td>4</td>
<td>R</td>
</tr>
</tbody>
</table>

View detailed session information using the command `show bfd neighbors detail`, as shown in the example in Verifying BFD sessions with BGP neighbors using show bfd neighbors detail.

**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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Change parameters for all static route sessions.</td>
<td>`ip route bfd interval milliseconds min_rx milliseconds multiplier value role [active</td>
<td>passive]`</td>
</tr>
</tbody>
</table>

View session parameters using the command `show bfd neighbors detail`, as shown in the example in Verifying BFD sessions with BGP neighbors using show bfd neighbors detail.

**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).

To disable BFD for static routes:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disable BFD for static routes.</td>
<td><code>no ip route bfd</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**Configuring BFD for OSPF**

BFD for OSPF is only supported on platforms: E C Z S 10

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. Enabling BFD globally.
2. Establishing sessions with OSPF neighbors.

Related configuration tasks

- Changing OSPF session parameters.
- Disabling BFD for OSPF.

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Establish sessions with all OSPF neighbors.</td>
<td>bfd all-neighbors</td>
<td>ROUTER-OSPF</td>
</tr>
</tbody>
</table>
To establish BFD for all OSPF neighbors on a single interface:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Establish sessions with all OSPF neighbors on a single interface.</td>
<td>ip ospf bfd all-neighbors</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

View the established sessions using the command `show bfd neighbors`, as shown in the example below.

```
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)

<table>
<thead>
<tr>
<th>LocalAddr</th>
<th>RemoteAddr</th>
<th>Interface</th>
<th>State</th>
<th>Rx-int</th>
<th>Tx-int</th>
<th>Multi Clients</th>
</tr>
</thead>
<tbody>
<tr>
<td>* 2.2.2.2</td>
<td>2.2.2.1</td>
<td>Gi 2/1</td>
<td>Up</td>
<td>100</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>* 2.2.3.1</td>
<td>2.2.3.2</td>
<td>Gi 2/2</td>
<td>Up</td>
<td>100</td>
<td>100</td>
<td>3</td>
</tr>
</tbody>
</table>
```

**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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Change parameters for OSPF sessions.</td>
<td>bfd all-neighbors interval milliseconds min_rx milliseconds multiplier value role [active</td>
<td>passive]</td>
</tr>
</tbody>
</table>

To change parameters for OSPF sessions on an interface:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Change parameters for all OSPF sessions on an interface.</td>
<td>ip ospf bfd all-neighbors interval milliseconds min_rx milliseconds multiplier value role [active</td>
<td>passive]</td>
</tr>
</tbody>
</table>

View session parameters using the command `show bfd neighbors detail`, as shown in the example in Displaying BFD for BGP Information.
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). 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disable BFD sessions with all OSPF neighbors.</td>
<td>no bfd all-neighbors</td>
<td>ROUTER-OSPF</td>
</tr>
</tbody>
</table>

To disable BFD sessions with all OSPF neighbors out of an interface:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disable BFD sessions with all OSPF neighbors out of an interface</td>
<td>ip ospf bfd all-neighbors disable</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

Configuring BFD for IS-IS

BFD for IS-IS is supported on platforms:

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.
2. Establish sessions for all or particular IS-IS neighbors.

Related configuration tasks
- Change session parameters.
- Disable BFD sessions for IS-IS.

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.
To establish BFD with all IS-IS neighbors:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Establish sessions with all IS-IS neighbors.</td>
<td>bfd all-neighbors</td>
<td>ROUTER-ISIS</td>
</tr>
</tbody>
</table>

To establish BFD with all IS-IS neighbors out of a single interface:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Establish sessions with all IS-IS neighbors out of an interface.</td>
<td>isis bfd all-neighbors</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

View the established sessions using the command show bfd neighbors, as shown in Figure 8-3.

Figure 8-3. 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
* 2.2.2.2    2.2.2.1    Gi 2/1    Up      100     100     3
```

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Change parameters for all IS-IS sessions.</td>
<td>bfd all-neighbors interval <em>milliseconds</em> min_<em>rx</em> <em>milliseconds</em> multiplier value role [active</td>
<td>passive]</td>
</tr>
</tbody>
</table>

To change parameters for IS-IS sessions on an interface:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Change parameters for all IS-IS sessions out of an interface.</td>
<td>isis bfd all-neighbors interval <em>milliseconds</em> min_<em>rx</em> <em>milliseconds</em> multiplier value role [active</td>
<td>passive]</td>
</tr>
</tbody>
</table>

View session parameters using the command `show bfd neighbors detail`.

# 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). 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disable BFD sessions with all IS-IS neighbors.</td>
<td>no bfd all-neighbors</td>
<td>ROUTER-ISIS</td>
</tr>
</tbody>
</table>
To disable BFD sessions with all IS-IS neighbors out of an interface:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disable BFD sessions with all IS-IS neighbors out of an interface.</td>
<td>isis bfd all-neighbors disable</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

### Configuring BFD for BGP

BFD for BGP is only supported on platforms: C E Z S 4 & 10

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 configure the following settings:

1. Configure BGP on the routers that you want to interconnect as described in Chapter 9, Border Gateway Protocol.
2. Enable fast fall-over for BGP neighbors to reduce convergence time (neighbor fall-over command) as described in BGP fast fall-over.

### Establishing sessions with BGP neighbors

Before configuring BFD for BGP, you must first configure BGP on the routers that you want to interconnect. For more information, refer to Chapter 9, Border Gateway Protocol.

For example, the following illustration 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.
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 ExaScale, up to 100 simultaneous BFD sessions are supported.
- On an S4810, up to 128 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.

On C-Series and E-Series only, 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enable BFD globally.</td>
<td>bfd enable</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Specify the AS number and enter ROUTER BGP configuration mode.</td>
<td>router bgp as-number</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>3</td>
<td>Add a BGP neighbor or peer group in a remote AS.</td>
<td>neighbor {ip-address</td>
<td>peer-group name} remote-as as-number</td>
</tr>
<tr>
<td>4</td>
<td>Enable the BGP neighbor.</td>
<td>neighbor {ip-address</td>
<td>peer-group-name} no shutdown</td>
</tr>
<tr>
<td>5</td>
<td>Configure parameters for a BFD session established with all neighbors discovered by BGP.</td>
<td>bfd all-neighbors [interval milliseconds min_rx milliseconds multiplier value role {active</td>
<td>passive}]</td>
</tr>
<tr>
<td></td>
<td>Establish a BFD session with a specified BGP neighbor or peer group using the default BFD session parameters.</td>
<td>neighbor {ip-address</td>
<td>peer-group-name} bfd</td>
</tr>
</tbody>
</table>

**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, refer to Configure Peer Groups.

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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verify a BFD for BGP configuration.</td>
<td>show running-config bgp&lt;br&gt;Verifying a BFD for BGP Configuration</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>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.</td>
<td>show bfd neighbors [interface] [detail]&lt;br&gt;Verifying BFD sessions with BGP neighbors using show bfd neighbors and Verifying BFD sessions with BGP neighbors using show bfd neighbors detail</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Check to see if BFD is enabled for BGP connections.</td>
<td>show ip bgp summary&lt;br&gt;Displaying BFD for BGP status</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Displays routing information exchanged with BGP neighbors, including BFD for BGP sessions.</td>
<td>show ip bgp neighbors [ip-address]&lt;br&gt;Displaying Routing Sessions with BGP neighbors</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

The following examples show the BFD for BGP output displayed for these show commands.
Verifying a BFD for BGP Configuration

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

Verifying BFD sessions with BGP neighbors using show bfd neighbors

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

<table>
<thead>
<tr>
<th>LocalAddr</th>
<th>RemoteAddr</th>
<th>Interface</th>
<th>State</th>
<th>Rx-int</th>
<th>Tx-int</th>
<th>Mult</th>
<th>Clients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1.1.3</td>
<td>1.1.1.2</td>
<td>Te 6/0</td>
<td>Up</td>
<td>100</td>
<td>100</td>
<td>3</td>
<td>B</td>
</tr>
<tr>
<td>2.2.2.3</td>
<td>2.2.2.2</td>
<td>Te 6/1</td>
<td>Up</td>
<td>100</td>
<td>100</td>
<td>3</td>
<td>B</td>
</tr>
<tr>
<td>3.3.3.3</td>
<td>3.3.3.2</td>
<td>Te 6/2</td>
<td>Up</td>
<td>100</td>
<td>100</td>
<td>3</td>
<td>B</td>
</tr>
</tbody>
</table>

Verifying BFD sessions with BGP neighbors using show bfd neighbors detail

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

Displaying BFD Packet Counters

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

Displaying BFD for BGP status
R2# show ip bgp summary
BGP router identifier 10.0.0.1, local AS number 2
BGP table version is 0, main routing table version 0
BFD is enabled, Interval 100 Min_rx 100 Multiplier 3 Role Active
3 neighbor(s) using 24168 bytes of memory

<table>
<thead>
<tr>
<th>Neighbor</th>
<th>AS</th>
<th>MsgRcvd</th>
<th>MsgSent</th>
<th>TblVer</th>
<th>InQ</th>
<th>OutQ</th>
<th>Up/Down</th>
<th>State/Pfx</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1.1.2</td>
<td>1</td>
<td>282</td>
<td>281</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>00:38:12</td>
<td>0</td>
</tr>
<tr>
<td>2.2.2.2</td>
<td>1</td>
<td>273</td>
<td>273</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>04:32:26</td>
<td>0</td>
</tr>
<tr>
<td>3.3.3.2</td>
<td>1</td>
<td>282</td>
<td>281</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>00:38:12</td>
<td>0</td>
</tr>
</tbody>
</table>

Displaying Routing Sessions with BGP neighbors
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
...
Configuring BFD for VRRP

BFD for VRRP is only supported on platforms: E C

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. Refer to Enabling BFD globally.
2. Establish VRRP BFD sessions with all VRRP-participating neighbors.
3. On the master router, establish a VRRP BFD sessions with the backup routers. Refer to Establishing sessions with all VRRP neighbors.

Related configuration tasks

- Changing VRRP session parameters.
- Establishing sessions with OSPF neighbors.

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.
To establish sessions with all VRRP neighbors:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Establish sessions with all VRRP neighbors.</td>
<td>vrrp bfd all-neighbors</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Establish a session with a particular VRRP neighbor.</td>
<td>vrrp bfd neighbor ip-address</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

View the established sessions using the command `show bfd neighbors`, as shown in the example below.

```
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
Session state information is also shown in the show vrrp command output, as shown in the following example.

R1(config-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

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Change parameters for all VRRP sessions.</td>
<td>vrrp bfd all-neighbors interval milliseconds min_rx milliseconds multiplier value role [active</td>
<td>passive]</td>
</tr>
</tbody>
</table>

To change parameters for a particular VRRP session:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Change parameters for a particular VRRP session.</td>
<td>vrrp bfd neighbor ip-address interval milliseconds min_rx milliseconds multiplier value role [active</td>
<td>passive]</td>
</tr>
</tbody>
</table>
View session parameters using the command `show bfd neighbors detail`, as shown in the example in [Verifying BFD sessions with BGP neighbors using show bfd neighbors detail](#).

## 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).

To disable all VRRP sessions on an interface:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disable all VRRP sessions on an interface.</td>
<td><code>no vrrp bfd all-neighbors</code></td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

To disable all VRRP sessions in a particular VRRP group:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disable all VRRP sessions in a VRRP group.</td>
<td><code>bfd disable</code></td>
<td>VRRP</td>
</tr>
</tbody>
</table>

To disable a particular VRRP session:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disable a particular VRRP session on an interface.</td>
<td><code>no vrrp bfd neighbor ip-address</code></td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

## Configuring BFD for VLANs

Configuring BFD for VLANs is supported on platforms C E.

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. **Enabling BFD globally.**
2. **Establishing sessions with VLAN neighbors.**
Related configuration tasks

- Establishing sessions with OSPF neighbors.

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 the illustration below. The session parameters do not need to match.

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).

To disable BFD on a VLAN interface:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disable all sessions on a VLAN interface.</td>
<td>no bfd enable</td>
<td>INTERFACE VLAN</td>
</tr>
</tbody>
</table>

Configuring BFD for Port-Channels

Configuring BFD for port-channels is supported on platforms: [C E]

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. **Enabling BFD globally.**
2. **Establishing sessions on port-channels.**

**Related configuration tasks**

- Disabling BFD for port-channels.

**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 the example below. The session parameters do not need to match.

**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).

To disable BFD for a port-channel:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disable BFD for a port-channel.</td>
<td>no bfd enable</td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
</tbody>
</table>
**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).

To enable Protocol Liveness:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enable Protocol Liveness</td>
<td>bfd protocol-liveness</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**Troubleshooting BFD**

Examine control packet field values using the command `debug bfd detail`. The following example shows a three-way handshake using this command.

```
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, 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, 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`.

```
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:
   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:
```

Bidirectional Forwarding Detection (BFD) | 169
00:34:14 : Sent packet for session with neighbor 2.2.2.2 on Gi 4/24

The output for the command debug bfd event is the same as the log messages that appear on the console by default.
Border Gateway Protocol

Platforms support BGP according to the following table:

<table>
<thead>
<tr>
<th>FTOS version</th>
<th>Platform support</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv4: 8.3.11.2 IPv6: 9.0.0.0</td>
<td>Z9000</td>
</tr>
<tr>
<td>8.3.7.0</td>
<td>S4810</td>
</tr>
<tr>
<td>8.1.1.0</td>
<td>E-Series ExaScale</td>
</tr>
<tr>
<td>7.8.1.0</td>
<td>S-Series</td>
</tr>
<tr>
<td>7.7.1.0.</td>
<td>C-Series</td>
</tr>
<tr>
<td>pre-7.7.1.0.</td>
<td>E-Series TeraScale</td>
</tr>
</tbody>
</table>

This chapter is intended to provide a general description of Border Gateway Protocol version 4 (BGPv4) as it is supported in the Force10 Operating System (FTOS).

This chapter includes the following topics:

- Protocol Overview
  - Autonomous Systems (AS)
  - Sessions and Peers
  - Route Reflectors
  - Confederations
- BGP Attributes
  - Best Path Selection Criteria
  - Weight
  - Local Preference
  - Multi-Exit Discriminators (MEDs)
  - AS Path
  - Next Hop
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 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 9-1.

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 9-1, 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 9-1), 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 9-1. 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 router 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. Each BGP router within an AS must have iBGP sessions with all other BGP routers in the AS. For example, a BGP network within an AS needs to be in “full mesh.” As seen in Figure 9-2, 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 9-2. 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 process 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.

Once the Open message parameters are agreed between peers then the neighbor relation is established and is in Open confirm state. This is when the router receives and checks for agreement on the parameters of open messages to establish a session.

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.

**Note:** Route Reflectors (RRs) should not be used in the forwarding path. In iBGP, hierarchical RRs maintaining forwarding plane RRs could create routing loops.

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 9-3 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. Routers E and H are client peers of Router D; Routers B and C are nonclient peers of Router D.

**Figure 9-3. 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.
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.
Note: In 8.3.11.4, the `bgp bestpath as-path multipath-relax` command is disabled by default, preventing BGP from load-balancing a learned route across two or more eBGP peers. To enable load-balancing across different eBGP peers, enable the `bgp bestpath as-path multipath-relax` command.

A system error will result if the `bgp bestpath as-path ignore` command and the `bgp bestpath as-path multipath-relax` command are configured at the same time. Only enable one command at a time.

Figure 9-4 illustrates the decisions BGP goes through to select the best path. The list following the illustration details the path selection criteria.

**Figure 9-4. BGP Best Path Selection**

<table>
<thead>
<tr>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Largest Local Pref</td>
</tr>
<tr>
<td>Local Originated Path</td>
</tr>
<tr>
<td>Shortest AS Path</td>
</tr>
<tr>
<td>Lowest Origin Code</td>
</tr>
<tr>
<td>Lowest MED</td>
</tr>
<tr>
<td>Learned via EBGP</td>
</tr>
<tr>
<td>Lowest NEXT-HOP Cost</td>
</tr>
</tbody>
</table>

Tie Breakers:
- Short Cluster List
- Router ID Check
- From Lowest BGP ID
- Lowest Peering Addr

A Single Route is Selected and Installed in the Forwarding Table
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 is selected when synchronization is disabled and only an internal path remains.
9. FTOS deems the paths as equal and does not perform steps 9 through 11 listed below if the following criteria are met:
   - the IBGP multipath or EBGP multipath is configured (using the maximum-path command)
   - the bgp bestpath as-path multipath-relax command is not configured and the received paths are from the same AS with the same number of ASs in the AS path but have different NextHops
   - the bgp bestpath as-path multipath-relax command is configured and the learned paths are from different ASs but have equal AS path numbers
   - 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 external path originated from the BGP router with the lowest router ID. If both paths are external, prefer the oldest path (first received path). 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 9-4. For this example, assume that LOCAL_PREF is the only attribute applied. In Figure 9-5, 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.
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 9-4.

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 9-6, 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.
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. If the outbound route-map uses MED, it will overwrite the IGP MED.

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 9-7. The question mark (?) indicates an Origin code of INCOMPLETE. The lower case letter (i) indicates an Origin code of IGP.
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 9-8. Note that the Origin attribute is shown following the AS Path information.

Figure 9-8. AS Path attribute reported

<table>
<thead>
<tr>
<th>Network</th>
<th>Next Hop</th>
<th>Metric</th>
<th>LocPrf</th>
<th>Weight</th>
<th>Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>* 7.0.0.0/29</td>
<td>10.114.8.33</td>
<td>0</td>
<td>0</td>
<td>18508</td>
<td>?</td>
</tr>
<tr>
<td>* 7.0.0.0/30</td>
<td>10.114.8.33</td>
<td>0</td>
<td>0</td>
<td>18508</td>
<td>?</td>
</tr>
<tr>
<td>* 9.2.0.0/16</td>
<td>10.114.8.33</td>
<td>10</td>
<td>0</td>
<td>18508</td>
<td>701 i</td>
</tr>
</tbody>
</table>
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 Z

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. Therefore, you cannot redistribute Multiprotocol BGP routes into BGP.

Implementing BGP with FTOS

Additional Path (Add-Path) support

BGP Add-path is supported on platforms E Z S4810

The Add-path feature reduces convergence times by advertising multiple paths to its peers for the same address prefix without replacing existing paths with new ones. By default, a BGP speaker advertises only the best path to its peers for a given address prefix. If the best path becomes unavailable, the BGP speaker withdraws it path from its local RIB and recalculates a new best path. This requires both IGP and BGP convergence and can therefore be a lengthy process. BGP Add-path also helps switchover to the next new best path when the current best path is unavailable.
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 9-1 gives some examples of these rules.

### Table 9-1. Example MED advertisement

<table>
<thead>
<tr>
<th>Command Settings</th>
<th>BGP Local Routing Information Base</th>
<th>MED Advertised to Peer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WITH route-map metric-type internal</td>
<td>WITHOUT route-map metric-type internal</td>
</tr>
<tr>
<td>redistribute <strong>isis</strong> (IGP cost = 20)</td>
<td>MED: IGP cost 20</td>
<td>MED = 20</td>
</tr>
<tr>
<td>redistribute <strong>isis</strong> route-map set metric 50</td>
<td>MED: IGP cost 50</td>
<td>MED: 50</td>
</tr>
<tr>
<td>redistribute <strong>isis</strong> metric 100</td>
<td>MED: IGP cost 100</td>
<td>MED: 100</td>
</tr>
</tbody>
</table>

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 9-2. 4-Byte ASN Dot Format Examples

<table>
<thead>
<tr>
<th>Traditional Format</th>
<th>Dot Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>65001</td>
<td>0.65501</td>
</tr>
<tr>
<td>65536</td>
<td>1.0</td>
</tr>
<tr>
<td>100000</td>
<td>1.34464</td>
</tr>
<tr>
<td>4294967295</td>
<td>65535.65535</td>
</tr>
</tbody>
</table>

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-octet-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 9-2.

- 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 9-9 and Figure 9-10).
Figure 9-9. 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)#show conf
!
router bgp 100
  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 34558, local router ID is 172.30.1.57
<output truncated>
```
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 9-11 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.
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 9-11. 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 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 9-3 displays the default values for BGP on FTOS.

**Table 9-3. FTOS BGP Defaults**

<table>
<thead>
<tr>
<th>Item</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGP Neighbor Adjacency changes</td>
<td>All BGP neighbor changes are logged.</td>
</tr>
<tr>
<td>Fast External Fallover feature</td>
<td>Enabled</td>
</tr>
<tr>
<td>graceful restart feature</td>
<td>Disabled</td>
</tr>
<tr>
<td>Local preference</td>
<td>100</td>
</tr>
<tr>
<td>MED</td>
<td>0</td>
</tr>
<tr>
<td>Route Flap Damping Parameters</td>
<td>half-life = 15 minutes, reuse = 750, suppress = 2000, max-suppress-time = 60 minutes</td>
</tr>
<tr>
<td>Distance</td>
<td>external distance = 20, internal distance = 200, local distance = 200</td>
</tr>
<tr>
<td>Timers</td>
<td>keepalive = 60 seconds, holdtime = 180 seconds</td>
</tr>
<tr>
<td>Add-path</td>
<td>Disabled</td>
</tr>
</tbody>
</table>
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_PREF attribute
- Change NEXT_HOP attribute
- Change WEIGHT attribute
- Enable multipath
- Filter BGP routes
- Redistribute routes
- Enable additional paths
- 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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>router bgp as-number</td>
<td>CONFIGURATION</td>
<td>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.</td>
</tr>
<tr>
<td>1a</td>
<td>bgp four-octet-as-support</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Enable 4-Byte support for the BGP process. <strong>Note:</strong> 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.</td>
</tr>
<tr>
<td>1b</td>
<td>address-family {ipv4</td>
<td>ipv6}</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
<tr>
<td>2</td>
<td>neighbor {ip-address</td>
<td>peer-group name} remote-as as-number</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
</tbody>
</table>
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 9-12 shows the summary with a 2-Byte AS Number displayed; Figure 9-13 shows the summary with a 4-Byte AS Number displayed.

**Figure 9-12. Command example: show ip bgp summary (2-Byte AS Number displayed)**

```plaintext
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#
```

You must **Configure Peer Groups** before assigning it a remote AS.

**Note:** When you change the configuration of a BGP neighbor, always reset it by entering the `clear ip bgp` command in EXEC Privilege mode.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>`neighbor {ip-address</td>
<td>peer-group-name} no shutdown`</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
</tbody>
</table>
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 9-14) command in EXEC Privilege mode. For BGP neighbor configuration information, use the `show running-config bgp` command in EXEC Privilege mode (Figure 9-15). Note that the `showconfig` command in CONFIGURATION ROUTER BGP mode gives the same information as the `show running-config bgp`.

Figure 9-14 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`. 

```
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

<table>
<thead>
<tr>
<th>Neighbor</th>
<th>AS</th>
<th>MsgRcvd</th>
<th>MsgSent</th>
<th>TblVer</th>
<th>InQ</th>
<th>OutQ</th>
<th>Up/Down</th>
<th>State/Pfx</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.10.21.1</td>
<td>65123</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>never</td>
<td>Active</td>
</tr>
<tr>
<td>10.10.32.3</td>
<td>65123</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>never</td>
<td>Active</td>
</tr>
<tr>
<td>100.10.92.9</td>
<td>65192</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>never</td>
<td>Active</td>
</tr>
<tr>
<td>192.168.10.1</td>
<td>65123</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>never</td>
<td>Active</td>
</tr>
<tr>
<td>192.168.12.2</td>
<td>65123</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>never</td>
<td>Active</td>
</tr>
</tbody>
</table>
```
Figure 9-14. 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#
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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable ASPLAIN AS Number representation. Figure 9-16</td>
<td>bgp asnotation asplain</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
<tr>
<td>Note: ASPLAIN is the default method FTOS uses and does not appear in the configuration display.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enable ASDOT AS Number representation. Figure 9-17</td>
<td>bgp asnotation asdot</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
<tr>
<td>Enable ASDOT+ AS Number representation. Figure 9-18</td>
<td>bgp asnotation asdot+</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
</tbody>
</table>

**Figure 9-16. 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
  5ab3eb9a15ed02ff4f0df4500d6017873cfd9a267c04957
  neighbor 172.30.1.250 no shutdown
5332332 9911991 65057 18508 12182 7018 46164 i
```

**Figure 9-17. 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
  5ab3eb9a15ed02ff4f0df4500d6017873cfd9a267c04957
  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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>neighbor peer-group-name peer-group</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Create a peer group by assigning a name to it.</td>
</tr>
<tr>
<td>2</td>
<td>neighbor peer-group-name no shutdown</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Enable the peer group. By default, all peer groups are disabled</td>
</tr>
<tr>
<td>3</td>
<td>neighbor ip-address remote-as as-number</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Create a BGP neighbor.</td>
</tr>
<tr>
<td>4</td>
<td>neighbor ip-address no shutdown</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Enable the neighbor.</td>
</tr>
</tbody>
</table>
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 9-19 shows the creation of a peer group (zanzibar).
Use the `neighbor peer-group-name no shutdown` command in the CONFIGURATION ROUTER BGP mode to enable a peer group.

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 9-21) to view the status of peer groups.
Figure 9-21. 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 fast-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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>`neighbor {ip-address</td>
<td>peer-group-name} fall-over`</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
</tbody>
</table>

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 9-22 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.
Use the `show ip bgp peer-group` command to verify that fast fall-over is enabled on a peer-group.
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.

You can constrain the number of passive sessions accepted by the neighbor. The limit keyword allows you to set the total number of sessions the neighbor will accept, between 2 and 265. The default is 256 sessions.
Use these commands in the following sequence, starting in the CONFIGURATION ROUTER BGP mode to configure passive peering.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>neighbor peer-group-name peer-group passive limit</code></td>
<td>CONFIG-ROUTER-BGP</td>
<td>Configure a peer group that does not initiate TCP connections with other peers. Enter the limit keyword to restrict the number of sessions accepted.</td>
</tr>
<tr>
<td>2</td>
<td><code>neighbor peer-group-name subnet subnet-number mask</code></td>
<td>CONFIG-ROUTER-BGP</td>
<td>Assign a subnet to the peer group. The peer group will respond to OPEN messages sent on this subnet.</td>
</tr>
<tr>
<td>3</td>
<td><code>neighbor peer-group-name no shutdown</code></td>
<td>CONFIG-ROUTER-BGP</td>
<td>Enable the peer group.</td>
</tr>
<tr>
<td>4</td>
<td><code>neighbor peer-group-name remote-as as-number</code></td>
<td>CONFIG-ROUTER-BGP</td>
<td>Create and specify a remote peer for BGP neighbor.</td>
</tr>
</tbody>
</table>

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 201.

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>`neighbor {IP address</td>
<td>peer-group-name local-as as number [no prepend]`</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
</tbody>
</table>

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 9-24. 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)#
```

**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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>`neighbor {IP address</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Allow this neighbor ID to use the AS path the specified number of times.</td>
</tr>
<tr>
<td>peer-group-name} allowas-in number</td>
<td></td>
<td>Format:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IP Address: A.B.C.D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Peer Group Name: 16 characters</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number: 1-10</td>
</tr>
</tbody>
</table>

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.
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>bgp graceful-restart</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Enable graceful restart for the BGP node.</td>
</tr>
<tr>
<td>bgp graceful-restart [restart-time time-in-seconds]</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Set maximum restart time for all peers. Default is 120 seconds.</td>
</tr>
<tr>
<td>bgp graceful-restart [stale-path-time time-in-seconds]</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Set maximum time to retain the restarting peer’s stale paths. Default is 360 seconds.</td>
</tr>
<tr>
<td>bgp graceful-restart [role receiver-only]</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Local router supports graceful restart as a receiver only.</td>
</tr>
</tbody>
</table>

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`.
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 9-26).

**Figure 9-26. Command example: show ip bgp paths**

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>neighbor [ip-address</td>
<td>peer-group-name] graceful-restart [stale-path-time time-in-seconds]</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
</tbody>
</table>

**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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>ip as-path access-list as-path-name</code></td>
<td>CONFIGURATION</td>
<td>Assign a name to an AS-PATH ACL and enter AS-PATH ACL mode.</td>
</tr>
<tr>
<td>2</td>
<td>`{deny</td>
<td>permit} filter parameter`</td>
<td>CONFIG-AS-PATH</td>
</tr>
<tr>
<td>3</td>
<td><code>exit</code></td>
<td>AS-PATH ACL</td>
<td>Return to CONFIGURATION mode</td>
</tr>
<tr>
<td>4</td>
<td><code>router bgp as-number</code></td>
<td>CONFIGURATION</td>
<td>Enter ROUTER BGP mode.</td>
</tr>
<tr>
<td>5</td>
<td>`neighbor {ip-address</td>
<td>peer-group-name} filter-list as-path-name {in</td>
<td>out}`</td>
</tr>
</tbody>
</table>

**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 9-27** 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.
Table 9-4 lists the Regular Expressions accepted in FTOS.

### Table 9-4. Regular Expressions

<table>
<thead>
<tr>
<th>Regular Expression</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>^ (caret)</td>
<td>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.</td>
</tr>
<tr>
<td>$ (dollar)</td>
<td>Matches the end of the input string.</td>
</tr>
<tr>
<td>. (period)</td>
<td>Matches any single character, including white space.</td>
</tr>
<tr>
<td>* (asterisk)</td>
<td>Matches 0 or more sequences of the immediately previous character or pattern.</td>
</tr>
</tbody>
</table>
As seen in Figure 9-27, 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| `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`, `level-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. |
Enable additional paths

By default, the add-path feature is disabled.

**Note:** Using Multipath and Add Path simultaneously in a Route Reflector is not recommended.

Use the following command in the CONFIGURATION ROUTER BGP mode to allow multiple paths sent to peers.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>bgp add-path {send</td>
<td>both} path-count count bgp add-path receive</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
<tr>
<td>2</td>
<td>neighbor {ipaddress</td>
<td>peergroup name} add-path [send</td>
<td>receive]</td>
</tr>
</tbody>
</table>

**Note:** The path-count parameter controls the number of paths that are advertised, not the number of paths that are received.

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:

- INTERNET
- NO_EXPORT_SUBCONFED
- NO_ADVERTISE
- NO_EXPORT

---

The redistribute ospf command is also available in FTOS for redistributing OSPF routes into BGP. This command does not have a process-id parameter and uses the match command to match external or internal OSPF routes. Here is the syntax and parameters:

**redistribute ospf**

- **process-id**
- **match external**
- **match internal**
- **metric-type**
- **route-map**
- **map-name**

The redistribute ospf command is used in the ROUTER BGP configuration mode. The following parameters are used:

- **process-id**
- **match external**
- **match internal**
- **metric-type**
- **route-map**
- **map-name**

The above command allows you to redistribute OSPF routes into the BGP network. The redistribute ospf command can be used in the ROUTER BGP configuration mode.

The redistribute ospf process-id command is used with the following parameters:

- **match external**
- **match internal**
- **metric-type**
- **route-map**
- **map-name**

The redistribute ospf command is used in the ROUTER BGP configuration mode. The following parameters are used:

- **process-id**
- **match external**
- **match internal**
- **metric-type**
- **route-map**
- **map-name**

The above command allows you to redistribute OSPF processes into the BGP network. The redistribute ospf command can be used in the ROUTER BGP configuration mode.
• All routes with the NO_EXP_ORT_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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ip community-list community-list-name</td>
<td>CONFIGURATION</td>
<td>Create a Community list and enter the COMMUNITY-LIST mode.</td>
</tr>
<tr>
<td>2</td>
<td>{deny</td>
<td>permit} {community-number</td>
<td>local-AS</td>
</tr>
</tbody>
</table>

Use these commands in the following sequence, starting in the CONFIGURATION mode to configure an IP extended community list.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ip extcommunity-list extcommunity-list-name</td>
<td>CONFIGURATION</td>
<td>Create a extended community list and enter the EXTCOMMUNITY-LIST mode.</td>
</tr>
</tbody>
</table>
To set or modify an extended community attribute, use the `set extcommunity {rt | soo} {ASN:NN | IPADDR:NN} | regexp REGEX-LINE` 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 9-28).

**Figure 9-28. 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
  deny 704:666
  deny 705:666
  deny 14551:666
FTOS#
```
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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>route-map map-name [permit</td>
<td>deny] [sequence-number]</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>match {community community-list-name [exact]</td>
<td>extcommunity extcommunity-list-name [exact]}</td>
<td>CONFIG-ROUTE-MAP</td>
</tr>
<tr>
<td>3</td>
<td>exit</td>
<td>CONFIG-ROUTE-MAP</td>
<td>Return to the CONFIGURATION mode.</td>
</tr>
<tr>
<td>4</td>
<td>router bgp as-number</td>
<td>CONFIGURATION</td>
<td>Enter the ROUTER BGP mode. AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte) or 0.1-65535.65535 (Dotted format)</td>
</tr>
<tr>
<td>5</td>
<td>neighbor {ip-address</td>
<td>peer-group-name} route-map map-name {in</td>
<td>out}</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>neighbor {ip-address</td>
<td>peer-group-name} send-community</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>route-map map-name [permit</td>
<td>deny] [sequence-number]</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>set comm-list community-list-name delete</td>
<td>CONFIG-ROUTE-MAP</td>
<td>Configure a set filter to delete all COMMUNITY numbers in the IP Community list.</td>
</tr>
<tr>
<td></td>
<td>set community {community-number</td>
<td>local-as</td>
<td>no-advertise</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• local-AS: routes with the COMMUNITY attribute of NO_EXPORT_SUBCONFED and are not sent to EBGP peers.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• no-advertise: routes with the COMMUNITY attribute of NO_ADVERTISE and are not advertised.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• no-export: routes with the COMMUNITY attribute of NO_EXPORT.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• none: remove the COMMUNITY attribute.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• additive: add the communities to already existing communities.</td>
</tr>
<tr>
<td>3</td>
<td>exit</td>
<td>CONFIG-ROUTE-MAP</td>
<td>Return to the CONFIGURATION mode.</td>
</tr>
<tr>
<td>4</td>
<td>router bgp as-number</td>
<td>CONFIGURATION</td>
<td>Enter the ROUTER BGP mode.</td>
</tr>
<tr>
<td>5</td>
<td>neighbor {ip-address</td>
<td>peer-group-name}</td>
<td>route-map map-name {in</td>
</tr>
</tbody>
</table>

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 9-29) to view BGP routes matching a certain community number or pre-defined BGP community.
Figure 9-29. 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

<table>
<thead>
<tr>
<th>Network</th>
<th>Next Hop</th>
<th>Metric</th>
<th>LocPrf</th>
<th>Weight</th>
<th>Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>* i 3.0.0.0/8</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 701 80 i</td>
<td></td>
</tr>
<tr>
<td>*i 4.2.49.12/30</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 i</td>
<td></td>
</tr>
<tr>
<td>* i 4.21.132.0/23</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 6461 16422 i</td>
<td></td>
</tr>
<tr>
<td>*i 4.24.118.16/30</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 i</td>
<td></td>
</tr>
<tr>
<td>*i 4.24.145.0/30</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 i</td>
<td></td>
</tr>
<tr>
<td>*i 4.24.187.12/30</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 i</td>
<td></td>
</tr>
<tr>
<td>*i 4.24.202.0/30</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 i</td>
<td></td>
</tr>
<tr>
<td>*i 4.25.88.0/30</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 3561 3908 i</td>
<td></td>
</tr>
<tr>
<td>*i 6.1.0.0/16</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 7170 1455 i</td>
<td></td>
</tr>
<tr>
<td>*i 6.2.0.0/22</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 7170 1455 i</td>
<td></td>
</tr>
<tr>
<td>*i 6.3.0.0/18</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 7170 1455 i</td>
<td></td>
</tr>
<tr>
<td>*i 6.4.0.0/16</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 7170 1455 i</td>
<td></td>
</tr>
<tr>
<td>*i 6.5.0.0/19</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 7170 1455 i</td>
<td></td>
</tr>
<tr>
<td>*i 6.8.0.0/20</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 7170 1455 i</td>
<td></td>
</tr>
<tr>
<td>*i 6.9.0.0/20</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 7170 1455 i</td>
<td></td>
</tr>
<tr>
<td>*i 6.10.0.0/15</td>
<td>195.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 7170 1455 i</td>
<td></td>
</tr>
<tr>
<td>*i 6.14.0.0/15</td>
<td>205.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 7170 1455 i</td>
<td></td>
</tr>
<tr>
<td>*i 6.133.0.0/21</td>
<td>205.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 7170 1455 i</td>
<td></td>
</tr>
<tr>
<td>*i 6.151.0.0/16</td>
<td>205.171.0.16</td>
<td>100</td>
<td>0</td>
<td>209 7170 1455 i</td>
<td></td>
</tr>
</tbody>
</table>
```

--More--
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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>bgp always-compare-med</code></td>
<td>CONFIG-ROUTER-BGP</td>
<td>Enable MED comparison in the paths from neighbors with different ASs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>By default, this comparison is not performed.</td>
</tr>
<tr>
<td>`bgp bestpath med {confed</td>
<td>missing-as-best}`</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>confed</code>: Chooses the bestpath MED comparison of paths learned from BGP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>confederations.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>missing-as-best</code>: Treat a path missing an MED as the most preferred</td>
</tr>
<tr>
<td></td>
<td></td>
<td>one</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>bgp default local-preference value</code></td>
<td>CONFIG-ROUTER-BGP</td>
<td>Change the LOCAL_PREF value.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- <code>value</code> range: 0 to 4294967295</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Default is 100.</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>`route-map map-name [permit</td>
<td>deny] [sequence-number]`</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>`neighbor {ip-address</td>
<td>peer-group-name} next-hop-self`</td>
<td>CONFIG-ROUTER-BGP</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>set next-hop ip-address</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Sets the next hop address.</td>
</tr>
</tbody>
</table>

### Change WEIGHT attribute

Use the following command in CONFIGURATION ROUTER BGP mode to change the how the WEIGHT attribute is used.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| `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:

```
set weight weight
```

**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.

**Note:** Using Multipath and Add Path simultaneously in a Route Reflector is not recommended.

Use the following command in the CONFIGURATION ROUTER BGP mode to allow more than one path.

```
maximum-paths {ebgp | ibgp} number
```

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 6, “Access Control Lists (ACLs),” on page 89 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 to ensure the changes are made.

Use these commands in the following sequence, starting in the CONFIGURATION mode to filter routes using prefix lists.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>ip prefix-list prefix-name</code></td>
<td>CONFIGURATION</td>
<td>Create a prefix list and assign it a name.</td>
</tr>
</tbody>
</table>
| 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 be matched  

Refer to Chapter 6, “Access Control Lists (ACLs),” on page 89 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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>`route-map map-name [permit</td>
<td>deny] [sequence-number]`</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>`{match</td>
<td>set}`</td>
<td>CONFIG-ROUTE-MAP</td>
</tr>
<tr>
<td>3</td>
<td><code>exit</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Return to the CONFIGURATION mode.</td>
</tr>
<tr>
<td>4</td>
<td><code>router bgp as-number</code></td>
<td>CONFIGURATION</td>
<td>Enter ROUTER BGP mode.</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>ip as-path access-list as-path-name</code></td>
<td>CONFIGURATION</td>
<td>Create a AS-PATH ACL and assign it a name.</td>
</tr>
</tbody>
</table>
| 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. |
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.

**Note:** Using Multipath and Add Path simultaneously in a Route Reflector is not recommended.

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 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.

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>aggregate-address ip-address</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Assign the IP address and mask of the prefix to be aggregated.</td>
</tr>
<tr>
<td>mask [advertise-map map-name]</td>
<td></td>
<td>Optional parameters are:</td>
</tr>
<tr>
<td>[as-set] [attribute-map map-name]</td>
<td></td>
<td>• <code>advertise-map map-name</code>: set filters for advertising an aggregate route</td>
</tr>
<tr>
<td>[summary-only] [suppress-map map-name]</td>
<td></td>
<td>• <code>as-set</code>: generate path attribute information and include it in the aggregate.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>attribute-map map-name</code>: modify attributes of the aggregate, except for the AS_PATH and NEXT_HOP attributes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>summary-only</code>: advertise only the aggregate address. Specific routes will not be advertised</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>suppress-map map-name</code>: identify which more-specific routes in the aggregate are suppressed</td>
</tr>
</tbody>
</table>

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 9-30. Command Example: show ip bgp**

<table>
<thead>
<tr>
<th>Network</th>
<th>Next Hop</th>
<th>Metric</th>
<th>LocPrf</th>
<th>Weight</th>
<th>Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>*&gt; 7.0.0.0/29</td>
<td>10.114.8.33</td>
<td>0</td>
<td>0</td>
<td>18508</td>
<td>7</td>
</tr>
<tr>
<td>*&gt; 7.0.0.0/30</td>
<td>10.114.8.33</td>
<td>0</td>
<td>0</td>
<td>18508</td>
<td>7</td>
</tr>
<tr>
<td>*a 9.0.0.0/8</td>
<td>192.0.0.0</td>
<td>32768</td>
<td>18508</td>
<td>701</td>
<td>(7018 2686 3786) 7</td>
</tr>
<tr>
<td>*&gt; 9.2.0.0/16</td>
<td>10.114.8.33</td>
<td>0</td>
<td>0</td>
<td>18508</td>
<td>701</td>
</tr>
<tr>
<td>*&gt; 9.141.128.0/24</td>
<td>10.114.8.33</td>
<td>0</td>
<td>0</td>
<td>18508</td>
<td>701 7018 2686 7</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>bgp confederation identifier as-number</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Specifies the confederation ID.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte)</td>
</tr>
<tr>
<td>bgp confederation peers as-number [... as-number]</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Specifies which confederation sub-AS are peers.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AS-number: 0-65535 (2-Byte) or 1-4294967295 (4-Byte)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>All Confederation routers must be either 4-Byte or 2-Byte. You cannot have a mix of router ASN support,</td>
</tr>
</tbody>
</table>

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 9-31. Setting Reuse and Restart Route Values**

Use the following command in the CONFIGURATION ROUTER BGP mode to configure route flap dampening parameters.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>bgp dampening</code></td>
<td>CONFIG-ROUTER-BGP</td>
<td>Enable route dampening. Enter the following optional parameters to configure route dampening parameters:</td>
</tr>
<tr>
<td><code>half-life</code></td>
<td></td>
<td>- <code>half-life</code> 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)</td>
</tr>
<tr>
<td><code>reuse</code></td>
<td></td>
<td>- <code>reuse</code> 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)</td>
</tr>
<tr>
<td><code>suppress</code></td>
<td></td>
<td>- <code>suppress</code> 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.)</td>
</tr>
<tr>
<td><code>max-suppress-time</code></td>
<td></td>
<td>- <code>max-suppress-time</code> 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.)</td>
</tr>
<tr>
<td><code>[route-map map-name]</code></td>
<td></td>
<td>- <code>route-map map-name</code>: 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.</td>
</tr>
</tbody>
</table>
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:

```
set dampening half-life reuse suppress max-suppress-time
```

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>set dampening</code></td>
<td>CONFIG-ROUTE-MAP</td>
<td>Enter the following optional parameters to configure route dampening parameters:</td>
</tr>
<tr>
<td>half-life reuse suppress max-suppress-time</td>
<td></td>
<td>• <code>half-life</code> 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)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>reuse</code> 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)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>suppress</code> 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.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>max-suppress-time</code> 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.)</td>
</tr>
</tbody>
</table>

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 9-32).

**Figure 9-32. 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
```

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>clear ip bgp dampening</td>
<td>EXEC Privilege</td>
<td>Clear all information or only information on a specific route.</td>
</tr>
<tr>
<td>[ip-address mask]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Use the following command in EXEC and EXEC Privilege mode to view statistics on route flapping.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show ip bgp flap-statistics</td>
<td>EXEC</td>
<td>View all flap statistics or for specific routes meeting the</td>
</tr>
<tr>
<td>[ip-address [mask]] [filter-list</td>
<td>EXEC Privilege</td>
<td>following criteria:</td>
</tr>
<tr>
<td>as-path-name] [regexp</td>
<td></td>
<td>• ip-address [mask]: enter the IP address and mask</td>
</tr>
<tr>
<td>regular-expression]</td>
<td></td>
<td>• filter-list as-path-name: enter the name of an</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AS-PATH ACL.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• regexp regular-expression: enter a regular</td>
</tr>
<tr>
<td></td>
<td></td>
<td>express to match on.</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>bgp non-deterministic-med</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Change the best path selection method to non-deterministic.</td>
</tr>
</tbody>
</table>

**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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>`neighbors {ip-address</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Configure timer values for a BGP neighbor or peer group.</td>
</tr>
<tr>
<td>peer-group-name} timers keepalive holdtime</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>timers bgp keepalive holdtime</code></td>
<td>CONFIG-ROUTER-BGP</td>
<td>Configure timer values for all neighbors.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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. Such resets cause undue interruption to traffic due to hard reset of the BGP cache and the time it takes to re-establish the session. BGP soft reconfig allows for policies to be applied to a session without clearing the BGP Session. Soft-reconfig can be done on a per-neighbor basis and can either be inbound or outbound.

BGP Soft Reconfiguration clears the policies without resetting the TCP connection.
Use the `clear ip bgp` command in EXEC Privilege mode at the system prompt to reset a BGP connection using BGP soft reconfiguration.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>clear ip bgp {*}</td>
<td>EXEC Privilege</td>
<td>Clear all information or only specific details.</td>
</tr>
<tr>
<td>neighbor-address</td>
<td></td>
<td>*: Clear all peers</td>
</tr>
<tr>
<td>AS Numbers</td>
<td></td>
<td>neighbor-address: Clear the neighbor with this IP address</td>
</tr>
<tr>
<td>ipv4</td>
<td></td>
<td>AS Numbers: Peers’ AS numbers to be cleared</td>
</tr>
<tr>
<td>peer-group-name</td>
<td></td>
<td>ipv4: Clear information for IPv4 Address family</td>
</tr>
<tr>
<td>[soft in</td>
<td></td>
<td>peer-group-name: Clear all members of the specified peer group</td>
</tr>
<tr>
<td>out]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Entering this command starts the storage of updates, which is required to do inbound soft reconfiguration. Outbound BGP soft reconfiguration does not require inbound soft reconfiguration to be enabled.

When soft-reconfiguration is enabled for a neighbor and the `clear ip bgp soft in` command is executed, the update database stored in the router is replayed and updates are reevaluated. With this command, the replay and update process is triggered only if route-refresh request is not negotiated with the peer. If the request is indeed negotiated (upon execution of `clear ip bgp soft in`), then BGP sends a route-refresh 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 should be 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 (Figure 9-33) enables 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 done later, the stored information is used to generate a new set of inbound updates.

**Figure 9-33. 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** **C**

MBGP for IPv4 Multicast is supported on platforms **C** **E** **T** **S** **Z**

MBGP is *not* supported on the E-Series ExaScale **E** 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>address family ipv4 multicast</td>
<td>CONFIG-ROUTER-BGP</td>
<td>Enables support for the IPv4 Multicast family on the BGP node</td>
</tr>
<tr>
<td>neighbor [ip-address</td>
<td>peer-group-name] activate</td>
<td>CONFIG-ROUTER-BGP-AF (Address Family)</td>
</tr>
</tbody>
</table>

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 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.

Debugging BGP

Use any of the commands in EXEC Privilege mode to enable BGP debugging.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>debug ip bgp [ip-address</td>
<td>peer-group peer-group-name] [in</td>
<td>out]</td>
</tr>
<tr>
<td>debug ip bgp dampening [in</td>
<td>out]</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>debug ip bgp [ip-address</td>
<td>peer-group peer-group-name] events [in</td>
<td>out]</td>
</tr>
<tr>
<td>debug ip bgp [ip-address</td>
<td>peer-group peer-group-name] keepalive [in</td>
<td>out]</td>
</tr>
<tr>
<td>debug ip bgp [ip-address</td>
<td>peer-group peer-group-name] notifications [in</td>
<td>out]</td>
</tr>
<tr>
<td>debug ip bgp [ip-address</td>
<td>peer-group peer-group-name] updates [in</td>
<td>out] [prefix-list name]</td>
</tr>
<tr>
<td>debug ip bgp {ip-address</td>
<td>peer-group-name} soft-reconfiguration</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

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 9-34**.

### Figure 9-34. 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
      ffffffff ffffffff ffffffff ffffffff 00160303 03010000  Last PDUs
    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 0218c0a8
    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 9-35) to change the maximum buffer size. View the captured PDUs using the command `show capture bgp-pdu neighbor`.

**Figure 9-35. 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 0181a1e4 0181a25c 41af92c0 00000000 00000000 00000000
         00000000 00000000 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 0c020a91 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 9-36.

**Figure 9-36. Required Memory for Captured PDUs**

```plaintext
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 entry(s) and 207896 paths using 42364576 bytes of memory
59913 BGP path attribute entry(s) using 2875872 bytes of memory
59910 BGP AS-PATH entry(s) using 2679698 bytes of memory
3 BGP community entry(s) using 81 bytes of memory
```

**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 9-37 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 9-37. Sample Configuration Illustration**

![Sample Configuration Illustration Diagram]
R1(config)#int loop 0
R1(config-if-lo-0)#ip address 192.168.128.1/24
R1(config-if-lo-0)#no shutdown
R1(config-if-lo-0)#show config
!
interface Loopback 0
  ip address 192.168.128.1/24
  no shutdown
R1(config-if-lo-0)#int gig 1/21
R1(config-if-gi-1/21)#ip address 10.0.1.21/24
R1(config-if-gi-1/21)#no shutdown
R1(config-if-gi-1/21)#show config
!
interface GigabitEthernet 1/21
  ip address 10.0.1.21/24
  no shutdown
R1(config-if-gi-1/21)#int gig 1/31
R1(config-if-gi-1/31)#ip address 10.0.3.31/24
R1(config-if-gi-1/31)#no shutdown
R1(config-if-gi-1/31)#show config
!
interface GigabitEthernet 1/31
  ip address 10.0.3.31/24
  no shutdown
R1(config-if-gi-1/31)#router bgp 99
R1(config-router_bgp)#network 192.168.128.0/24
R1(config-router_bgp)#neighbor 192.168.128.2 remote 99
R1(config-router_bgp)#neighbor 192.168.128.2 no shut
R1(config-router_bgp)#neighbor 192.168.128.2 update-source Loopback 0
R1(config-router_bgp)#neighbor 192.168.128.3 remote 100
R1(config-router_bgp)#neighbor 192.168.128.3 no shut
R1(config-router_bgp)#neighbor 192.168.128.3 update-source Loopback 0
R1(config-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(config-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

<table>
<thead>
<tr>
<th>Neighbor</th>
<th>AS</th>
<th>MsgRcvd</th>
<th>MsgSent</th>
<th>TblVer</th>
<th>InQ</th>
<th>OutQ</th>
<th>Up/Down</th>
<th>State/Pfx</th>
</tr>
</thead>
<tbody>
<tr>
<td>192.168.128.2</td>
<td>99</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>00:00:32</td>
<td>1</td>
</tr>
<tr>
<td>192.168.128.3</td>
<td>100</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>00:00:09</td>
<td>4</td>
</tr>
</tbody>
</table>
Figure 9-39. 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)#int gig 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 192.168.128.3 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

<table>
<thead>
<tr>
<th>Neighbor</th>
<th>AS</th>
<th>MsgRcvd</th>
<th>MsgSent</th>
<th>TblVer</th>
<th>InQ</th>
<th>OutQ</th>
<th>Up/Down</th>
<th>State/Pfx</th>
</tr>
</thead>
<tbody>
<tr>
<td>192.168.128.1</td>
<td>99</td>
<td>40</td>
<td>35</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>00:01:05</td>
<td>1</td>
</tr>
<tr>
<td>192.168.128.3</td>
<td>100</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>00:00:16</td>
<td>1</td>
</tr>
</tbody>
</table>

R2#
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-gi-3/11)#int gig 3/21
R3(conf-if-gi-3/21)#ip address 10.0.2.33/24
R3(conf-if-gi-3/21)#no shutdown
R3(conf-if-gi-3/21)#show config
!
interface GigabitEthernet 3/21
  ip address 10.0.2.33/24
  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
R3#
Figure 9-41. Enable Peer Group - Router 1

```
R1#conf
R1(config)#router bgp 99
R1(config-router_bgp)# network 192.168.128.0/24
R1(config-router_bgp)# neighbor AAA peer-group
R1(config-router_bgp)# neighbor AAA no shutdown
R1(config-router_bgp)# neighbor BBB peer-group
R1(config-router_bgp)# neighbor BBB no shutdown
R1(config-router_bgp)# neighbor 192.168.128.2 peer-group AAA
R1(config-router_bgp)# neighbor 192.168.128.3 peer-group BBB

R1(config-router_bgp)#
R1(config-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 entry(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 entry(s) using 74 bytes of memory
2 BGP AS-PATH entry(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 9-42. 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

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 9-43. Enable Peer Groups - Router 2

```bash
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)#show conf

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 entry(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 entry(s) using 128 bytes of memory
2 BGP AS-PATH entry(s) using 90 bytes of memory
2 neighbor(s) using 9216 bytes of memory

<table>
<thead>
<tr>
<th>Neighbor</th>
<th>AS</th>
<th>MsgRcvd</th>
<th>MsgSent</th>
<th>TblVer</th>
<th>InQ</th>
<th>OutQ</th>
<th>Up/Down</th>
<th>State/Pfx</th>
</tr>
</thead>
<tbody>
<tr>
<td>192.168.128.1</td>
<td>99</td>
<td>140</td>
<td>136</td>
<td>2</td>
<td>0</td>
<td>(0)</td>
<td>00:11:24</td>
<td>1</td>
</tr>
<tr>
<td>192.168.128.3</td>
<td>100</td>
<td>138</td>
<td>140</td>
<td>2</td>
<td>0</td>
<td>(0)</td>
<td>00:18:31</td>
<td>1</td>
</tr>
</tbody>
</table>

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 9-44. 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)#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 entries 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 entries using 128 bytes of memory
2 BGP AS-PATH entries 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) 0:00:15 1
192.168.128.2 99 122 120 1 0 (0) 0:00:11 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
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
Bare Metal Provisioning 3.0 (BMP 3.0)

Bare Metal Provisioning 3.0 (BMP 3.0) is included as part of the FTOS image. It is supported on platforms S4810 and Z.

Overview

Bare Metal Provisioning (BMP) is a feature that improves operational efficiency to the system by automatically loading pre-defined configurations and FTOS images using standard protocols such as DHCP and common file transfer mechanisms.

Bare Metal Provisioning:

• reduces the time to install and configure the network device.
• helps eliminate configuration errors and ensure consistent configurations.
• can be used on a single system or on multiple systems.

With BMP 3.0, the system can retrieve a configuration file or a pre-configuration script indicated in the DHCP offer. Using the pre-configuration script, you can:

• verify the integrity of the boot image downloaded from the DHCP offer.
• decide what type of configurations you want to apply based on factors such as your network reachability, port status, and neighbor discovery.
• monitor your CPU and memory utilization, your port traffic status, or perform link and topology checking with LLDP.
• retrieve a configuration from a central repository.

This chapter includes the following topics:

• Prerequisites
• Important Information
• BMP Process Overview
• Preparing BMP
• Reload Modes
  • BMP Mode
  • Normal Mode
• Scripts
• Pre-configuration Scripts
• Post-configuration Scripts
• Auto-execution Scripts
• Configuration Tasks
• Script Examples

Prerequisites

Before you use BMP 3.0 to auto-configure a supported Dell Force10 switch, you must first configure:

• An external Dynamic Host Configuration Protocol (DHCP) server (required) - a network device offering configuration parameters
• File Server (required) - a network device for storing and servicing files.
• Domain Name Server (DNS) (optional) - a server that associates domain names in the network with IP addresses.
• Relay Agent (optional) - an intermediary network device that passes messages between DHCP clients and the DHCP server when the server is not on the same subnet. It can also provide IP addresses for multiple subnets.

For more information, see Domain Name Server and File Server.

Important Information

BMP 3.0 is supported on the user ports and management ports of a switch.

If your network has VLT enabled on aggregator switches and you are configuring the ToR switch to load using BMP, ensure the aggregator switches are configured with the `lacp ungroup member-independent vlt` command if the DHCP and file servers are reachable via the interface configured as part of VLT LAG.

Bare Metal Provisioning works to ease configuration in the following key areas:

• Switch access is allowed through all ports (management and user ports) with or without DHCP-based dynamic IP address configuration of a switch.
• Booting up in Layer 3 mode with interfaces already in no shutdown mode and some basic protocols enabled to protect the system and the network.
• Use the DHCP offer to access the configuration file or a pre-configuration script.

BMP Process Overview

When the system boots up in the default BMP mode, the following items are requested:
2. Configuration file or pre-configuration script (ZSH, TCL, or Expect script).
3. A list of checksums for all these components.

Note: The configuration file is to maintain normal BMP functionality when a pre-configuration script is not sent.

Preparing BMP

DHCP Server

DHCP Configuration

You must first configure a DHCP server before you can use the BMP mode on a switch. Configure the DHCP server with the set of parameters described below for each client switch. Refer to the FTOS Configuration Guide: Dynamic Host Configuration Protocol chapter for detailed information. The DHCP server is configured to assign an IP to the system and other parameters.

Update the following parameters on the appropriate DHCP server.

• Boot File Name: The FTOS image to be loaded on the system. The boot file name is expected to use option 67 or the boot filename in the BOOTP payload of the DHCP offer. If both are specified, option 67 will be used. The system supports TFTP, HTTP, HTTPS, SFTP, SCP and FTP protocols, as well as files stored in Flash.
• Configuration File Name: The configurations to be applied to the system. The configuration file name is expected to use option 209.
• File Server Address: The server where the Image and Configurations file are placed. The address is assumed to be a TFTP address unless it is given as a URL. The system supports TFTP, HTTP, HTTPS, SFTP, SCP and FTP protocols, as well as files stored in Flash.
• Domain Name Server: (Optional) The DNS server to be contacted to resolve the hostname.
• IP Address: Dynamic IP address for the system. This is used only for file transfers.

The DHCP option codes used are:

• 6      Domain Name Server IP
• 60     Vendor class identifier
• 66     TFTP Server name
• 67     Boot filename
• 150    TFTP server IP address
• 209    Configuration File
230 User port stacking

Note: BMP will eventually exit when the timeout occurs.

DHCP Retry Mechanism

BMP requests a different DHCP offer in the following scenarios:

- If the command `reload-type config-scr-download enable` is enabled, the DHCP offer specifies both the boot image and the configuration file. If either download is successful, BMP will not request another DHCP offer.
- If the offer contains only a boot image that cannot be downloaded, BMP will request another DHCP offer.
- If the command `reload-type config-scr-download enable` is enabled and the configuration file in the offer cannot be downloaded, BMP will request another DHCP offer.

DHCP server IP Blacklist

If the process does not complete successfully, the DHCP server IP will be blacklisted and the BMP process will be re-initiated. A DHCP server IP is maintained in the blacklist for ten minutes. If a DHCP offer is received from the blacklisted DHCP server, the offer will be rejected until the IP is alive in the blacklist (10 minutes).

MAC-Based IP assignment

One way to use the BMP mode most efficiently is to configure the DHCP server to assign a fixed IP address, FTOS image, and configuration file based on the system’s MAC address. When this is done, the same IP address is assigned to the system even on repetitive reloads and the same configuration file will be retrieved when using the DNS server or the `network-config` file to determine the hostname.

The assigned IP address is only used to retrieve the files from the file server. It is discarded after the files are retrieved.

Following is an example of a configuration of the DHCP server included on the most popular Linux distribution. The `dhcpd.conf` file shows the MAC based IP and configuration file assignment are fixed.

<table>
<thead>
<tr>
<th>Parameter Example</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>option configfile code 209=text;</td>
<td></td>
</tr>
<tr>
<td>option bootfile-name code 67=text;</td>
<td></td>
</tr>
<tr>
<td>host HOST1 {</td>
<td></td>
</tr>
<tr>
<td>###### Mac to IP mapping</td>
<td></td>
</tr>
<tr>
<td>hardware ethernet 00:01:e8:8c:4d:0e;</td>
<td></td>
</tr>
<tr>
<td>fixed-address 30.0.0.20;</td>
<td></td>
</tr>
<tr>
<td>###### Config file name could be given in the following way</td>
<td></td>
</tr>
</tbody>
</table>
File Server

Set up a file server and ensure connectivity.

The file server that holds the boot and configuration files must be configured to allow file transfers to the switch. The system recognizes FTP, SFTP, TFTP, HTTP, HTTPS, SCP, USB, and Flash URLs.

For example:

- `tftp://server ip or name/filename`
- `ftp://user:passwd@serverip or name//mypath/FTOS-A.B.C.D.bin`
- `flash://filename`
- `http://server ip`  
  - `name/filename`

When loading the FTOS image, if the FTOS image on the server is different from the image on the local Flash, the system downloads the image from the server onto the local Flash and reloads using that image. If it is the same image, the system loads the configuration file, if present, or the startup-config without downloading a new image.

Domain Name Server

Set up a Domain Name Server (DNS) to determine the host name applied in the switch startup configuration when no configuration file is retrieved from the DHCP server. The DNS server is contacted only when no configuration file is contained in a DHCP server response and the host name is not resolved from the `network-config` file on the switch. Refer to the FTOS Configuration Guide, IPv4 Addressing chapter, Resolution of Host Names for information.

Reload Modes

Bare Metal Provisioning supports two types of reload modes: BMP mode and Normal mode.
BMP mode is the default boot mode configured for a new system arriving from Dell Force10. This mode obtains the FTOS image and configuration file from a network source (DHCP and file servers).

Use Normal mode to boot the switch up with the management port in a no shutdown mode. If the management IP address is present in the start-up configuration file, it will be assigned. If the management IP address is not present in the start-up configuration file, no IP address will be assigned to the management interface.

**BMP Mode**

**BMP Mode Prerequisites**

Before you use the BMP mode to auto-configure a switch, you must first configure a DHCP, DNS, and file server in the network.

On a new factory-loaded switch, the switch boots up in the default **BMP** mode to facilitate configuration of the switch. You can reconfigure a switch to reload between **BMP** and **Normal** mode.

- **BMP mode** (default): The switch automatically configures all ports (management and user ports) as Layer 3 physical ports and acts as a DHCP client on the ports for a user-configured time (DHCP timeout). It is set with the `reload-type bmp` command.
- **Normal mode**: The switch loads the FTOS image and startup configuration file stored in the local Flash. New configurations require that the Management IP and Management Interface be configured manually. This mode is set with the `reload-type normal-reload` command.

**Note:** To apply the startup configuration, you must cancel the default BMP setup using the `stop bmp` command from the console. To disable BMP for the next reload, enter the `reload-type normal-reload` command.

**MAC-Based IP Address Assignment**

One way to deploy the BMP mode is to configure the DHCP server to assign a fixed IP address and configuration file based on the system’s MAC address. In this way, the same IP address is assigned and the same configuration file is retrieved when the switch reloads.

Using a dynamic IP address assignment may create a situation in which the desired configuration is not loaded on the system because the IP address changes each time the system is reloaded.

For example, on a DHCP3 server, you can configure the assignment of a fixed MAC-based IP address and configuration file by entering the following lines of configuration parameters in the `dhcpd.conf` file on the server:

```plaintext
host S4810 {
    hardware ethernet 00:01:e8:81:e2:39;
    fixed-address 20.0.0.48;
    option configfile "customer.conf";
}
```

To apply the startup configuration, you must cancel the default BMP setup using the `stop bmp` command from the console. To disable BMP for the next reload, enter the `reload-type normal-reload` command.
Normal Mode

When reloaded in Normal mode, the switch boots up with the management port in a no shutdown mode. If the management IP address is present in the start-up configuration file, it will be assigned. If the management IP address is not present in the start-up configuration file, no IP address will be assigned to the management interface. You can connect to the management port with an IP address on the same network and log in to the system through a telnet or SSH session.

To configure a switch to reload using the Normal mode, enter the `reload-type normal-reload` command.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>reload-type normal-reload</td>
<td>GLOBAL CONFIGURATION</td>
<td>The switch reloads in Normal mode using the FTOS image and startup configuration file stored in the local Flash.</td>
</tr>
</tbody>
</table>

DHCP Configuration

Normal mode does not require a separate DHCP server configuration.

FTOS Image Retrieval

The FTOS image is loaded from the local Flash.

Scripts

With 9.1(0.0), the system supports a scripting environment when a BMP or Normal reload occurs. BMP uses pre- and post-configuration scripts, while a Normal reload uses an auto-execution (auto-exec) script.

Pre-configuration Scripts

In previous BMP 2.0 releases, FTOS accessed the Configuration file or the Image file from the DHCP server. With BMP 3.0, FTOS now accesses the image, followed by the Configuration file or a pre-configuration script from the DHCP offer. Use of the pre-configuration scripts allows you to verify the integrity of the FTOS image downloaded from the DHCP offer. You can also use the pre-configuration script to dynamically decide what types of configurations to apply to your system based on various factors such as network reachability, port status, neighbor discovery and so forth. You can also monitor CPU and memory utilization, port traffic status, and perform link and topology checking with LLDP. The system supports pre-configuration scripts in TCLSH, EXPECT, and ZSH.
Post-configuration Scripts

In BMP 3.0, after the pre-configuration script has completed and the configuration is loaded, you can run a post-configuration script if one is present in the configuration file. Use the post-configuration script to check the status of configured ports or protocols which can then be sent as a status report to a central repository for your network administrators. You may also use the post-configuration script to set the host name of the system or perform additional configuration settings. The system supports post-configuration scripts in TCLSH, EXPECT, and ZSH. If you have SmartScripts installed in FTOS, the system also supports post-configuration scripts in PERL and Python.

Auto-execution Scripts

The function of an auto-execution script is the same as a pre-configuration script except that it is executed on every reboot in Normal mode. Scripts must be stored in a flash://autoexec file. Auto-exec scripts are independent of BMP.

The auto-exec script is only executed when:

- BMP is disabled.
- the script is stored in a flash://autoexec file.
- when the reload-type normal-reload command is used before the chassis is reloaded.

If the auto-exec script fails, the system generates a message indicating the failure and does not load the configuration file. You must update the error in the script before continuing the upgrade.

Process for pre-configuration scripts:

1. Decide what information you want to pre-configure, such as username/password information, verifying the integrity of the boot image downloaded from the DHCP offer, types of configurations to be applied, and so on.
2. Create a pre-configuration script in ZSH, TCL, or Expect. The script cannot be binary.
3. Store the script on the DHCP server in the DHCP offer.
4. Apply the FTOS upgrade. The system will boot in BMP default mode and run the pre-configuration script.
   - The default timer on the script is 10 minutes; the maximum amount of time the script can run is 1 hour. If an error occurs that ends the script abnormally, the system will reboot.
5. The system receives an IP address via the DHCP server which it uses to get an FTOS image to be booted, a configuration file (if supplied) and a pre-configuration script to be run.
   - During the pre-configuration process, a connection to a dedicated instance of the CLI is available.
Configuration Tasks

When the system boots up in BMP mode all ports, including management ports, are placed in L3 mode in a no shut state. The system acts as a DHCP client on these ports for a period of time (dhcp-timeout). This allows the system time to send out a DHCP DISCOVER on all the interface up ports to the DHCP Server in order to obtain its IP address, boot image filename and configuration file from the DHCP server.

- Set up a DHCP server. Refer to the FTOS Configuration Guide: Dynamic Host Configuration Protocol chapter for detailed information. You must configure the DHCP server to assign an IP address to each system along with the following DHCP server parameters:
  - IP address pool: IP address for the system.
  - Image name: FTOS image used to boot up the system.
  - configuration file: Startup configuration parameters to be applied to the system.
  - IP address of TFTP file server: File server from which the image and startup configuration file are downloaded.
  - Domain Name server: DNS server to be contacted to resolve the host name.

Use the following DHCP option codes in the dhcpd.conf file:
- 6 IP address of Domain Name Server, if a DNS server is needed to resolve the IP address of a file server hostname.
- 150 IP address of a TFTP server.
- 209 configuration file.

For example, a sample dhcpd.conf file on a DHCP3 server is as follows:

```plaintext
lease-file-name "/var/lib/dhcpd/dhcpd.leases";
option configfile code 209 = text;
option tftp-server-address code 150 = ip-address;
subnet 20.0.0.0 netmask 255.255.255.0 {
    range 20.0.0.26 20.0.0.28;
    filename "FTOS-SC-1-2-0-363.bin";
    option configfile "customer.conf";
    option tftp-server-address 20.0.0.1;
    option domain-name-servers 20.0.0.1;
}
```

- Set up a TFTP server and ensure connectivity.

Configure a TFTP file server as the network source from which the switch downloads the image file and the configuration file to be applied to the system. On a TFTP server, the required files are commonly stored in the /tftpboot directory.

When loading the FTOS image, if the FTOS image on the file server is different from the image stored in the local Flash memory, the system downloads the image on the file server into the local Flash memory and reboots using that image. If the image is the same, the system reloads the image from the local Flash memory.

- Set up a DNS server. Refer to the FTOS Configuration Guide: IPv4 Addressing chapter: Resolution of Host Names section for information.

Configure the DNS server to determine a system’s host name and the correct configuration file to be downloaded on the system if the DHCP server response does not specify a configuration file and the host name is not resolved using the network-config file.
**System boot and set-up behavior in BMP Mode**

1. System begins boot up process in BMP mode (default mode).

2. The system sends DHCP Discover on all the interface up ports.

   ```
   00:01:31: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Ma 0/0.
   00:01:31: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/0.
   00:01:31: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/5.
   00:01:31: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/6.
   00:01:47: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/8.
   00:01:47: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/35.
   00:01:47: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Fo 0/56.
   00:01:47: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Fo 0/60.
   00:01:47: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Ma 0/0.
   ```

3. The IP address, boot image filename and the configuration file name are reserved for the system and provided in the DHCP reply (one-file read method). The system receives its IP address, subnet mask, DHCP server IP, TFTP server address, DNS server IP, bootfile name and the configuration filename from the DHCP server.

   If a DHCP offer has neither an image path or configuration file path it is considered to be an invalid BMP DHCP offer and the offer is ignored. The first DHCP offer with IP address, FTOS image and configuration file, or the IP address and FTOS image, or the IP address and configuration file is chosen.

4. The DHCP OFFER is selected.

   a. All other ports except the port on which the offer was received and selected are set to shutdown mode.

   ```
   00:01:33: %STKUNIT0-M:CP %JUMPSTART-5-DHCP_OFFER: DHCP acquired IP 30.0.0.20 mask 255.255.0.0
   server IP 30.1.1.1.
   00:01:33: %STKUNIT0-M:CP %JUMPSTART-5-DHCP_OFFER: DHCP tftp IP 30.0.0.1 dns IP 30.0.0.1 router
   IP 30.0.0.14.
   00:01:33: %STKUNIT0-M:CP %JUMPSTART-5-DHCP_OFFER: DHCP image file FTOS-SE-8.3.10.1.bin.
   00:01:33: %STKUNIT0-M:CP %JUMPSTART-5-DHCP_OFFER: DHCP config file pt-s4810-12.
   00:01:33: %STKUNIT0-M:CP %JUMPSTART-5-DHCP_OFFER: DHCP stacking info NIL.
   ```

5. The system sends a unicast message to the server to retrieve the named configuration file and/or boot file from the base directory of the server.

   a. If an optional **bootfile-name** is used, the file name can be 256 bytes. If a **filename** field is specified in the DHCP offer, the filename can be 128 bytes. The name can be a fully qualified URL or it can be a file name only.

   b. When an FTOS build image is found, the system compares that build image to the version currently loaded on the chassis.
• If there is a mismatch between the build images, the system upgrades to the downloaded version and reloads.

00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Downloaded Image Major Version : 1
00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Downloaded Image Minor Version : 0
00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Downloaded Image Main Version : 0
00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Downloaded Image Patch Version : 0
00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Flash A Image Major Version : 9
00:03:07: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Flash A Image Minor Version : 1
00:03:07: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Flash A Image Main Version : 0
00:03:07: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Flash A Image Patch Version : 222
00:03:07: %STKUNIT0-M:CP %JUMPSTART-2-JUMPSTART_DOWNLOAD_START: The FTOS image download has started.
Erasing IOM Primary Image, please wait

• If the versions match, the system downloads the configuration file or pre-configuration script.

00:03:27: %STKUNIT0-M:CP %JUMPSTART-5-CFG_APPLY: The downloaded config from dhcp server is being applied
00:03:27: %STKUNIT0-M:CP %BMP-5-JUMPSTART: DHCP RELEASE sent on Fo 0/56.
00:03:27: %STKUNIT0-M:CP %SYS-5-CONFIG_LOAD: Loading configuration file

c If the configuration file or pre-configuration script is downloaded from the server, any saved startup-configuration on the Flash is ignored. If no configuration file is downloaded from the server, the startup-configuration file on the Flash is loaded as in Normal reload.

6. When the FTOS build image and the configuration file have been downloaded, the IP address is released.

00:04:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART: DHCP RELEASE sent on Fo 0/56.
00:04:06: %STKUNIT0-M:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Fo 0/56

7. The system applies the configuration. The system is now up and running. It can be managed as usual.

**BMP mode: Boot and Set-up Behavior**

When a switch that is configured to reload in BMP mode boots up, one of the following scenarios may occur:

• **Reload without a DHCP Server Offer**
• **Reload with a DHCP Server Offer without an FTOS Image**
• **Reload with a DHCP Server Offer without a Configuration File**
• **Reload with a DHCP Server Offer without a DNS Server**
• **Reload with a Pre-configuration Script (BMP mode only)**
• **Using the Post-Configuration Script (BMP mode only)
Reload without a DHCP Server Offer

A switch configured to reload in BMP mode and if the DHCP server cannot be reached, the system keeps on sending DISCOVER messages.

00:01:44: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/50.
00:01:44: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Te 0/51.
00:01:44: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DISCOVER: DHCP DISCOVER sent on Ma 0/0.

Reload with a DHCP Server Offer without an FTOS Image

If a switch that is configured to reload in BMP mode reaches a DHCP server but does not locate a downloadable FTOS image file on the server, it attempts to download the configuration file.

1. The system boots up with the BMP application.


2. The system downloads the customer.conf configuration file from the file-server address if config-scr-download is enabled.

3. If the configuration download is successful, the following logs are displayed.

00:02:29: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DWNLD_CONFIG_SCRIPT_SUCCESS: The config file download is successful.
00:02:30: %STKUNIT0-M:CP %JUMPSTART-5-DOWNLOAD_INFO: test123 config file successfully downloaded
00:02:31: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_DOWNLOAD: The config file download is successful.
00:02:31: %STKUNIT0-M:CP %JUMPSTART-5-CFG_APPLY: The downloaded config from dhcp server is being applied
00:02:31: %STKUNIT0-M:CP %SYS-5-CONFIG_LOAD: Loading configuration file
00:02:31: %STKUNIT0-M:CP %JUMPSTART-5-PROCESS_COMPLETED: Autoconfig process is complete. The configuration must be saved in order to use it after subsequent reloads

Reload with a DHCP Server Offer without a Configuration File

If a switch that is configured to reload in BMP mode reaches a DHCP server but cannot retrieve a configuration file, the switch looks for a configuration file on the file server only if the config-scr-download command is enabled.

1. The system boots up with the BMP application.
2. The system receives a DHCP offer from a DHCP server with the following parameters:


3. The system downloads the build image from the file server.
4. The system compares the current local build image to the downloaded build image as follows:
   a. If the build image versions match, the system does not try to load any image and comes up with the FTOS prompt.

   00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Downloaded Image Major Version : 1
   00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Downloaded Image Minor Version : 0
   00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Downloaded Image Main Version : 0
   00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Downloaded Image Patch Version : 0
   00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Flash A Image Major Version : 1
   00:03:07: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Flash A Image Minor Version : 0
   00:03:07: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Flash A Image Main Version : 0
   00:03:07: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO: Flash A Image Patch Version : 0
b If the build images versions are different, the system stores the downloaded build image in the local Flash and loads the build image from the Flash. This process is repeated until the build image versions match.

```
00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO:    Downloaded Image Major Version : 1
00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO:    Downloaded Image Minor Version : 0
00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO:    Downloaded Image Main Version : 0
00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO:    Downloaded Image Patch Version : 0
00:03:06: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO:    Flash A Image Major Version : 9
00:03:07: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO:    Flash A Image Minor Version : 1
00:03:07: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO:    Flash A Image Main Version : 0
00:03:07: %STKUNIT0-M:CP %JUMPSTART-5-JUMPSTART_RELEASE_HEADER_INFO:    Flash A Image Patch Version : 222
00:03:07: %STKUNIT0-M:CP %JUMPSTART-2-JUMPSTART_DOWNLOAD_START: The FTOS image download has started.

Erasing Sseries Primary Image, please wait

..................```

The system looks for the configuration file on the file server. If a configuration file is not found, the download fails. If the build image processing is successful, the BMP stops. If the build image download also fails, then the current offer is treated as invalid and the BMP tries to send DISCOVER messages.

**Reload with a DHCP Server Offer without a DNS Server**

Although the DNS server is optional, it allows you to specify the configuration file to be applied to a switch by assigning a hostname.

When the DHCP offer is received and no DNS IP address is specified, if the configuration file cannot be retrieved from a file server the configuration file will not be downloaded.

**Reload with a Pre-configuration Script (BMP mode only)**

1. To reload FTOS on a switch using a pre-configuration script, the following conditions must be true:
   - BMP must be enabled.
   - The download of the script is from an external server. The location will be specified in the DHCP offer.
   - The first line of the script must contain one of the following:

   ```
   #!/usr/bin/expect
   #!/usr/bin/tclsh
   #!/usr/bin/zsh
   ```

2. After the first line, but before the actual start of the script, the script must contain the signature “#!/DELL-FORCE10”.
3. The pre-configuration script is downloaded instead of the configuration file.
4. The pre-configuration script is run before applying the start-up configuration file.
5. The pre-configuration script has the ability to use configuration FTOS CLI commands using the utility name “F10do”.
6. When the pre-configuration script completes, the start-up configuration file will be automatically applied.

Using the Post-Configuration Script (BMP mode only)

To reload FTOS on a switch using a pre-configuration script, the following conditions must be true:

- BMP must be enabled.
- The post-configuration script can be written in Expect, TCLSH, or ZSH. If the SmartScripts package is already installed, the post-configuration script can also be written in PERL or Python.
- No restraints are required for the post-configuration script, such as the signature “#/DELL-FORCE10” that is required for the pre-configuration script.
- The post-configuration script must be configured by the `script post-config` command.
- The post-configuration script must be executed after the start-up configuration process required by BMP has been applied.
- The post-configuration script has the ability to use configuration FTOS CLI commands using the utility name “F10do”.

Reload using the Auto-execution Script (Normal mode only)

To use the auto-execution script, the following conditions must be true:

- BMP must be disabled.
- The auto-execution script must always be stored in `flash://autoexec`
- The auto-execution script can be written in Expect, TCLSH, or ZSH. If the SmartScripts package is already installed, the post-configuration script can also be written in PERL or Python.
- No restraints are required for the auto-execution script, such as the signature “#/DELL-FORCE10” that is required for the pre-configuration script.
- The auto-execution script has the ability to use configuration FTOS CLI commands using the utility name “F10do”.
- When the auto-execution script is complete, the start-up configuration is applied depending on the return status of the script:
  - Success - 0: the start-up configuration is applied.
  - Failure - non-zero: the start-up configuration is not applied.
Script Examples

Auto-execution Script - Normal mode

FTOS#show reload-type
Reload-Type            :   normal-reload [Next boot : normal-reload]
FTOS#show file flash://autoexec
#!/usr/bin/tclsh
puts [ exec f10do "show version" ]
puts [ exec date ]
puts "this is Autoexec script"
FTOS#
FTOS#
FTOS#reload
System configuration has been modified. Save? [yes/no]: no

Proceed with reload [confirm yes/no]: yes
00:32:16: %STKUNIT1-M:CP %CHMGR-5-RELOAD: User request to reload the chassis
syncing disks... done
unmounting file systems...
unmounting /f10/flash (/dev/ld0h)...
unmounting /usr/pkg (/dev/ld0g)...
unmounting /usr (mfs:35)...
unmounting /f10 (mfs:21)...
unmounting /kern (kernfs)...
unmounting / (/dev/md0a) ... done
rebooting
p
..
..
..
Starting Dell Force10 application

00:00:13: %STKUNIT1-M:CP %RAM-6-ELECTION_ROLE: Stack unit 1 is transitioning to Management unit.
00:00:15: %STKUNIT1-M:CP %CHMGR-5-STACKUNITDETECTED: Stack unit 1 present
The following line indicates the start of the auto-execution script.

```
00:00:16: %STKUNIT1-M:CP %JUMPSTART-5-AUTOEXEC_START:
00:00:19: %STKUNIT1-M:CP %CHMGR-5-CHECKIN: Checkin from Stack unit 1 (type S4810, 64 ports)
00:00:20: %STKUNIT1-M:CP %CHMGR-0-PS_UP: Power supply 0 in unit 1 is up
00:00:20: %STKUNIT1-M:CP %CHMGR-5-STACKUNITUP: Stack unit 1 is up
00:00:21: %STKUNIT1-M:CP %CHMGR-5-SYSTEM_READY: System ready
00:00:20: %STKUNIT1-M:CP %CHMGR-5-STACKUNITUP: Stack unit 1 is up
00:00:21: %STKUNIT1-M:CP %CHMGR-5-SYSTEM_READY: System ready
00:00:22: %STKUNIT1-M:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Ma 1/0
00:00:26: %S4810:1 %IFAGT-5-INSERT_OPTICS: Optics SFP inserted in slot 1 port 30
00:00:27: %STKUNIT1-M:CP %CHMGR-5-PSU_FAN_UP: Fan 0 in PSU 0 of Unit 1 is up
00:00:29: %S4810:1 %IFAGT-5-INSERT_OPTICS_PLUS: Optics SFP+ inserted in slot 1 port 11
00:00:36: %STKUNIT1-M:CP %IFMGR-5-IFM_ISCSI_ACL_REGION_NOT_ALLOCATED: iSCSI Session monitoring cannot be enabled without ACL regions allocated to it. To enable iSCSI Session Monitoring allocate cam-blocks to iscsioptacl using cam-acl CLI and then save and reload.
00:00:36: %STKUNIT1-M:CP %IFMGR-5-IFM_ISCSI_ENABLE: iSCSI has been enabled causing flow control to be enabled on all interfaces. EQI detection and enabling iscsi profile-compellent on an interface may cause some automatic configurations to occur like jumbo frames on all ports and no storm control and spanning tree port-fast on the port of detection
00:00:36: %STKUNIT1-M:CP %SEC-5-LOGIN_SUCCESS: Login successful on console
```

The following line indicates the auto-execution script is executing.

```
FTOS>FTOS#terminal length 0
FTOS#show version
```

```
Dell Force10 Operating System Software
Dell Force10 Operating System Version: 2.0
Dell Force10 Application Software Version: 1-0(0-338)
Copyright (c) 1999-2012 by Dell Inc. All Rights Reserved.
Build Path: /sites/sjc/work/build/buildSpaces/build06/FIT-INDUS-1-0-0/SW/SRC
System image file is "dt-maa-s4810-72"
System Type: S4810
Control Processor: Freescale QorIQ P2020 with 2147483648 bytes of memory.
128M bytes of boot flash memory.
1 52-port GE/TE/FG (SE)
40 Ten GigabitEthernet/IEEE 802.3 interface(s)
```

```
4 Forty GigabitEthernet/IEEE 802.3 interface(s)
FTOS#
```

```
Wed Jan 2 22:47:34 GMT 2013
this is Autoexec script
```
The following line indicates the successful completion of the auto-execution script.

```
00:00:49: %STKUNIT1-M:CP %JUMPSTART-5-AUTOEXEC_SUCCESS:
The AutoExec Script execution returned Success.
```

The following line indicates that the Configuration file is loaded into the switch.

```
FTOS#00:00:51: %STKUNIT1-M:CP %SYS-5-CONFIG_LOAD: Loading configuration file
00:00:52: %STKUNIT1-M:CP %IFMGR-5-ASTATE_UP: Changed interface Admin state to up: Te 0/36
00:00:53: %STKUNIT1-M:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Ma 0/0
```
Pre-configuration Script - BMP Mode

#!/usr/bin/expect

#/DELL-FORCE10

# Execute F10do and Print
proc print_f10do {cmd_str} {
    set str [exec f10do "$cmd_str"]
    set tmp_str [string map {
        \r
        \n    } $str ]
    puts $tmp_str
}

set ftp_ip         "20.0.0.1"
set ftp_username   "lab"
set ftp_passwd     "lab"

set config_file    "s4810-10-startup-config"
set post_conf      "s4810-10-post-config.exp"

puts "Executing Pre-Config Script !!!\\n"

exec rm -rf "$config_file"
exec rm -rf "$post_conf"

puts "Downloading Startup Config and Post-Config Script from $ftp_ip ...\\n"

spawn ftp "$ftp_ip"
expect    "Name .*: "
send      "$ftp_username\\n"
expect    "Password:"
send      "$ftp_passwd\\n"
expect    "ftp>"
send      "cd scripts\\n"
expect    "ftp>"
send      "ls\\n"
expect    "ftp>"
send      "get $post_conf\\n"
expect    "ftp>"
send      "get $config_file\\n"
expect    "ftp>"
send      "bye\\n"
expect    eof
after 5000
puts "Download Complete !!!\r\n"

if {
if {file exists $config_file]}
puts "Config File: $config_file downloaded successfully\r\n"
} else {
puts "ERROR: Config File: $config_file - Not Found\r\n"
}

if {
if {file exists $post_conf]}
puts "Post Config Script: $post_conf downloaded successfully\r\n"
} else {
puts "ERROR: Post Config Script: $post_conf - Not Found\r\n"
}

# Copy Config to Startup Config

print_f10do "show version"

print_f10do "copy flash://$config_file startup-config"
print_f10do "yes"

puts "Pre-Config Script Execution Successfull !!!!!!\r\n"

exit 0
Content Addressable Memory (CAM)

Content Addressable Memory (CAM) is supported on platforms: E C S S4810

- Content Addressable Memory
- CAM Profiles
- Microcode
- CAM Profiling for ACLs
- When to Use CAM Profiling
- Differences Between EtherScale and TeraScale
- Important Points to Remember
- Select CAM Profiles
- CAM Allocation
- Test CAM Usage
- View CAM Profiles
- View CAM-ACL settings
- View CAM-ACL settings
- Configure IPv4Flow Sub-partitions
- Configure Ingress Layer 2 ACL Sub-partitions
- Return to the Default CAM Configuration
- CAM Optimization
- Applications for CAM Profiling
- Troubleshoot CAM Profiling

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-inac1`.

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, therefore 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

<table>
<thead>
<tr>
<th>CAM Profile</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>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</td>
</tr>
<tr>
<td>eg-default</td>
<td>For EG-series line cards only. EG series line cards have two CAM modules per Port-pipe. Available Microcodes: default, ipv6-extacl</td>
</tr>
<tr>
<td>ipv4-320k</td>
<td>Provides 320K entries for the IPv4 Forwarding Information Base (FIB) and reduces the IPv4 Flow partition to 12K. Available Microcodes: default, lag-hash-mpls</td>
</tr>
<tr>
<td>ipv4-egacl-16k</td>
<td>Provides 16K entries for egress ACLs Available Microcodes: acl-group</td>
</tr>
<tr>
<td>ipv6-extacl</td>
<td>Provides IPv6 functionality. Available Microcodes: ipv6-extacl</td>
</tr>
<tr>
<td>l2-ipv4-inac1</td>
<td>Provides 32K entries for Layer 2 ingress ACLs and 28K entries for Layer 3 IPv4 ingress ACLs. Available Microcodes: default</td>
</tr>
<tr>
<td>unified-default</td>
<td>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</td>
</tr>
</tbody>
</table>
Table 11-1. CAM Profile Descriptions

<table>
<thead>
<tr>
<th>CAM Profile</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipv4-64k-ipv6</td>
<td>Provides IPv6 functionality; an alternate to ipv6-extacl that redistributes CAM space from the IPv4FIB to IPv4Flow and IPv6FIB.</td>
</tr>
<tr>
<td></td>
<td>Available Microcodes: ipv6-extacl</td>
</tr>
</tbody>
</table>

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, therefore adding entries to one region necessarily decreases the number available to other regions.

Table 11-2. CAM entries per partition

<table>
<thead>
<tr>
<th>Profile</th>
<th>Partition</th>
<th>L2FIB</th>
<th>L2ACL</th>
<th>IPv4FIB</th>
<th>IPv4ACL</th>
<th>IPv4Flow</th>
<th>EgL2ACL</th>
<th>EgIPv4ACL</th>
<th>Reserved</th>
<th>IPv6FIB</th>
<th>IPv6ACL</th>
<th>IPv6Flow</th>
<th>EgIPv6ACL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td></td>
<td>32K</td>
<td>2K</td>
<td>256K</td>
<td>12K</td>
<td>24K</td>
<td>1K</td>
<td>1K</td>
<td>8K</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>eg-default</td>
<td></td>
<td>32K</td>
<td>2K</td>
<td>512K</td>
<td>12K</td>
<td>24K</td>
<td>1K</td>
<td>1K</td>
<td>8K</td>
<td>32K</td>
<td>3K</td>
<td>4K</td>
<td>1K</td>
</tr>
<tr>
<td>ipv4-320k</td>
<td></td>
<td>32K</td>
<td>2K</td>
<td>320K</td>
<td>12K</td>
<td>12K</td>
<td>1K</td>
<td>1K</td>
<td>4K</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ipv4-egacl-16k</td>
<td></td>
<td>32K</td>
<td>2K</td>
<td>192K</td>
<td>8K</td>
<td>24K</td>
<td>0</td>
<td>16K</td>
<td>8K</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ipv6-extacl</td>
<td></td>
<td>32K</td>
<td>2K</td>
<td>192K</td>
<td>12K</td>
<td>8K</td>
<td>1K</td>
<td>1K</td>
<td>2K</td>
<td>6K</td>
<td>3K</td>
<td>4K</td>
<td>2K</td>
</tr>
<tr>
<td>l2-ipv4-inacl</td>
<td></td>
<td>32K</td>
<td>33K</td>
<td>64K</td>
<td>27K</td>
<td>8K</td>
<td>2K</td>
<td>2K</td>
<td>2K</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>unified-default</td>
<td></td>
<td>32K</td>
<td>3K</td>
<td>192K</td>
<td>9K</td>
<td>8K</td>
<td>2K</td>
<td>2K</td>
<td>2K</td>
<td>6K</td>
<td>2K</td>
<td>4K</td>
<td>2K</td>
</tr>
<tr>
<td>ipv4-64k-ipv6</td>
<td></td>
<td>32K</td>
<td>2K</td>
<td>64K</td>
<td>12K</td>
<td>24K</td>
<td>1K</td>
<td>1K</td>
<td>8K</td>
<td>16K</td>
<td>3K</td>
<td>4K</td>
<td>1K</td>
</tr>
</tbody>
</table>

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.
CAM Profiling for ACLs

CAM Profiling for ACLs is supported on platform E only.

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>
<thead>
<tr>
<th>Partition</th>
<th>% Allocated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sysflow</td>
<td>6</td>
</tr>
<tr>
<td>L2ACL</td>
<td>14</td>
</tr>
<tr>
<td>PVST</td>
<td>50</td>
</tr>
<tr>
<td>QoS</td>
<td>12</td>
</tr>
<tr>
<td>L2PT</td>
<td>13</td>
</tr>
<tr>
<td>FRRP</td>
<td>5</td>
</tr>
</tbody>
</table>
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 as shown in the following command output example. 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
```
Example: 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

Example: 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. Refer to LAG Hashing.
- Hash based on bidirectional flow for LAGs. Refer to LAG Hashing based on Bidirectional Flow.
- Optimize the VLAN ACL Group feature, which permits group VLANs for IP egress ACLs. Refer to CAM profile for the VLAN ACL group feature.
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. Refer to Pre-calculating Available QoS CAM Space.
• 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Select a CAM profile.</td>
<td>cam-profile profile microcode</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td></td>
<td></td>
<td>microcode</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Save the running-configuration.</td>
<td>copy running-config startup-config</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>
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. Note that there are 16 FP blocks, but the System Flow requires 3 blocks that cannot be reallocated. The default CAM Allocation settings are:

- L3 ACL (ipv4acl): 6
- L2 ACL (l2acl): 5
- IPv6 L3 ACL (ipv6acl): 0
- L3 QoS (ipv4qos): 1
- L2 QoS (l2qos): 1
- L2PT (l2pt): 1
- MAC ACLs (ipmacacl): 2
- ECFMACL (ecfmacl): 0
- VMAN QoS (vman-qos): 0
- VMAN Dual QoS (vman-dual-qos): 0

The following additional CAM Allocation settings are supported on the S4810:

- FCoE ACL (fcoeacl): 0
- ISCSI Opt ACL (isesioptacl): 2

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.

Note: On the S4810, there can be only one odd number of Blocks in the CLI configuration; the other Blocks must be in factors of 2. For example, a CLI configuration of 5+4+2+1+1 Blocks is not supported; a configuration of 6+4+2+1 Blocks is supported.

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Verify that the new CAM profile will be written to the CAM on the next boot.</td>
<td>show cam-profile summary</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>4</td>
<td>Reload the system.</td>
<td>reload</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>
To configure the IPv4 and IPv6 ACLs and QoS regions on the entire system:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Select a cam-acl action</td>
<td>cam-acl [default</td>
<td>I2acl]</td>
</tr>
</tbody>
</table>

**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 |

**Note:** If allocation values are not entered for the CAM regions, the value is 0.

| 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. The following example gives a sample of the output shown when executing the command. The status column indicates whether or not the policy can be enabled.

```
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 the following example. This command also shows the profile that will be loaded upon the next chassis or component reload.
FTOS#show cam-profile

-- Chassis CAM Profile --

CamSize : 18-Meg
Profile Name : Default
L2FIB : 32K entries
L2ACL : 1K entries
IPv4FIB : 256K entries
IPv4ACL : 12K entries
IPv4Flow : 24K entries
EgL2ACL : 1K entries
EgIPv4ACL : 1K entries
Reserved : 8K entries
FIB : 0 entries
ACL : 0 entries
Flow : 0 entries
EqACL : 0 entries
MicroCode Name : 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 as shown in the following example.

![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.]

FTOS# show running-config cam-profile
!
cam-profile default microcode default

FTOS#

View CAM-ACL settings

The show cam-acl command is supported on platforms C S  S

View the current cam-acl settings for the C-Series, S-Series and systems chassis and each component using the command show cam-acl, as shown in as shown in the following examples.

The default values for the show cam-acl command for C-Series and S-Series are:

FTOS# show cam-acl

-- Chassis Cam ACL --

Current Settings(in block sizes)
L2Acl : 2
Ipv4Acl : 2
Ipv6Acl : 2
The default values for the show cam-acl command for the S4810 are:

FTOS#show cam-acl

-- Chassis Cam ACL --
  Current Settings (in block sizes)
  L2Acl       :     4
  Ipv4Acl     :     4
  Ipv6Acl     :     0
  Ipv4Qos     :     2
  L2Qos       :     1
  L2PT        :     0
  IpMacAcl    :     0
  VmanQos     :     0
  VmanDualQos :     0
  EcfmAcl     :     0
  FcoeAcl     :     0
  iscsiOptAcl :     2

-- Stack unit 0 --
  Current Settings (in block sizes)
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 the following example.

```
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       |       30   | 0      |       30
|        | 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-partition are supported on platform E
The IPv4Flow CAM partitions have 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>
<thead>
<tr>
<th>Partition</th>
<th>Space Allocated (EtherScale)</th>
<th>Space Allocated (TeraScale)</th>
<th>Space Allocated (ExaScale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACL</td>
<td>8K</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Multicast FIB/ACL</td>
<td>9K</td>
<td>3K</td>
<td>3K</td>
</tr>
<tr>
<td>PBR</td>
<td>1K</td>
<td>1K</td>
<td>1K</td>
</tr>
<tr>
<td>QoS</td>
<td>8K</td>
<td>2K</td>
<td>2K</td>
</tr>
<tr>
<td>System Flow</td>
<td>5K</td>
<td>5K</td>
<td>5K</td>
</tr>
<tr>
<td>Trace Lists</td>
<td>1</td>
<td>1K</td>
<td>1K</td>
</tr>
</tbody>
</table>

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 (refer to 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Re-allocate CAM space within the IPv4Flow partition.</td>
<td>cam-ipv4flow</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Save the running-configuration.</td>
<td>copy running-config startup-config</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>3</td>
<td>Verify that the new CAM configuration will be written to the CAM on the next boot.</td>
<td>show cam-ipv4flow</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>4</td>
<td>Reload the system.</td>
<td>reload</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>
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 --

<table>
<thead>
<tr>
<th>Current Settings</th>
<th>Next Boot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multicast Fib/Acl</td>
<td>8K</td>
</tr>
<tr>
<td>Pbr</td>
<td>2K</td>
</tr>
<tr>
<td>Qos</td>
<td>7K</td>
</tr>
<tr>
<td>System Flow</td>
<td>6K</td>
</tr>
<tr>
<td>Trace Lists</td>
<td>1K</td>
</tr>
</tbody>
</table>

-- Line card 0 --

<table>
<thead>
<tr>
<th>Current Settings</th>
<th>Next Boot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multicast Fib/Acl</td>
<td>8K</td>
</tr>
<tr>
<td>Pbr</td>
<td>2K</td>
</tr>
<tr>
<td>Qos</td>
<td>7K</td>
</tr>
<tr>
<td>System Flow</td>
<td>6K</td>
</tr>
<tr>
<td>Trace Lists</td>
<td>1K</td>
</tr>
</tbody>
</table>

-- Line card 1 --

<table>
<thead>
<tr>
<th>Current Settings</th>
<th>Next Boot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multicast Fib/Acl</td>
<td>8K</td>
</tr>
<tr>
<td>Pbr</td>
<td>2K</td>
</tr>
<tr>
<td>Qos</td>
<td>7K</td>
</tr>
<tr>
<td>System Flow</td>
<td>6K</td>
</tr>
<tr>
<td>Trace Lists</td>
<td>1K</td>
</tr>
</tbody>
</table>
Configure Ingress Layer 2 ACL Sub-partitions

IPv4Flow sub-partitions are supported on platform.

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

<table>
<thead>
<tr>
<th>Partition</th>
<th>% Allocated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sysflow</td>
<td>6</td>
</tr>
<tr>
<td>L2ACL</td>
<td>14</td>
</tr>
<tr>
<td>*PVST</td>
<td>50</td>
</tr>
<tr>
<td>QoS</td>
<td>12</td>
</tr>
<tr>
<td>L2PT</td>
<td>13</td>
</tr>
<tr>
<td>FRRP</td>
<td>5</td>
</tr>
</tbody>
</table>

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 (refer to 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 as shown in the following example:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Re-allocate CAM space within the Ingress Layer 2 ACL partition.</td>
<td>cam-l2acl</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Save the running-configuration.</td>
<td>copy running-config startup-config</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>3</td>
<td>Verify that FTOS will write the new CAM configuration to the CAM on the next boot.</td>
<td>show cam-l2acl</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>4</td>
<td>Reload the system.</td>
<td>reload</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

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

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

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 the following example.

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 **E1** 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Verify that you have configured a CAM profile that allocates 24K entries to the IPv4 system flow region. Refer to View CAM Profiles.</td>
</tr>
<tr>
<td>2</td>
<td>Allocate more entries in the IPv4Flow region to QoS. Refer to Configure IPv4Flow Sub-partitions.</td>
</tr>
</tbody>
</table>

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, refer to Pre-calculating Available QoS CAM Space.

> **Note:** For troubleshooting other CAM issues, refer to the *E-Series Network Operations Guide*. 
Control Plane Policing (CoPP)

Overview

Control Plane Policing (CoPP) uses ACL rules and QoS policies to create filters for a system’s control plane. That filter prevents traffic not specifically identified as legitimate from reaching the system control plane, rate-limits, traffic to an acceptable level.

Control Plane Policing (CoPP) increases security on the system by protecting the Routing Processor from unnecessary or DoS traffic, giving priority to important control plane and management traffic. CoPP uses a dedicated control plane configuration through the ACL and QoS CLIs to provide filtering and rate-limiting capabilities for the control plane packets.

The following illustration shows an example of the difference between having CoPP implemented and not having CoPP implemented.
Configure Control Plane Policing

The S4810 can process a maximum of 4200 PPS (packets per second). Protocols that share a single queue may experience flaps if one of the protocols receives a high rate of control traffic even though Per Protocol CoPP is applied. This happens because Queue-Based Rate Limiting is applied first.

For example, BGP and ICMP share same queue (Q6); Q6 has 400 PPS of bandwidth by default. The desired rate of ICMP is 100 pps and the remaining 300 pps is assigned to BGP. If ICMP packets come at 400 pps, BGP packets may be dropped though ICMP packets are rate-limited to 100 PPS. This may be solved by increasing Q6 bandwidth to 700 pps to allow both ICMP and BGP packets and then applying per-flow CoPP for ICMP and BGP packets. The setting of this Q6 bandwidth is purely dependent on the incoming traffic for the set of protocols sharing the same queue. If the user is not aware of the incoming protocol traffic rate then the required Queue Rate Limit value cannot be set. Such queue bandwidth tuning must be done carefully because the system cannot open up to handle any rate, including traffic coming at the line rate.

CoPP policies are assigned on a per-protocol or a per-queue basis, and are assigned in CONTROL-PLANE mode to each port-pipe.
The CoPP policies are configured by creating extended ACL rules and specifying rate-limits through QoS policies. The ACLs and QoS policies are assigned as service-policies.

**Configure CoPP for protocols**

This section lists the commands necessary to create and enable the service-policies for CoPP. Refer to *Access Control Lists (ACLs)* and *Quality of Service (QoS)* for complete information about creating ACLs and QoS rules.

The basics for creating a CoPP service policy is to create a Layer 2, Layer 3, and/or an IPv6 ACL rule for the desired protocol type. Then, create a QoS input policy to rate-limit the protocol traffics according to the ACL. The ACL and QoS policies are finally assigned to a control-plane service policy for each port-pipe.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create a Layer 2 extended ACL for control-plane traffic policing for a</td>
<td>`mac access-list extended name cpu-qos permit {arp</td>
<td>frrp</td>
</tr>
<tr>
<td></td>
<td>particular protocol.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Create a Layer 3 extended ACL for control-plane traffic policing for a</td>
<td>`ip access-list extended name cpu-qos permit {bgp</td>
<td>dhcp</td>
</tr>
<tr>
<td></td>
<td>particular protocol.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Create an IPv6 ACL for control-plane traffic policing for a particular</td>
<td>`ipv6 aqaccess-list name cpu-qos permit {bgp</td>
<td>icmp</td>
</tr>
<tr>
<td></td>
<td>protocol.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Create a QoS input policy for the router and assign the policing.</td>
<td><code>qos-policy-input name cpu-qos rate-police</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>5</td>
<td>Create a QoS class map to differentiate the control-plane traffic and</td>
<td>`class-map match-any name cpu-qos match {ip</td>
<td>mac</td>
</tr>
<tr>
<td></td>
<td>assign to an ACL.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Create a QoS input policy map to match to the class-map and qos-policy</td>
<td><code>policy-map-input name cpu-qos class-map name qos-policy name</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td></td>
<td>for each desired protocol.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Enter Control Plane mode.</td>
<td><code>control-plane-cpuqos</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>8</td>
<td>Assign the protocol based service policy on the control plane. Enabling</td>
<td><code>service-policy rate-limit-protocols</code></td>
<td>CONTROL-PLANE</td>
</tr>
<tr>
<td></td>
<td>this command on a port-pipe automatically enables the ACL and QoS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rules creates with the cpu-qos keyword.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sample Config for CoPP protocol configuration

Create IP/IPv6/MAC Extended ACL

FTOS(conf)#ip access-list extended ospf cpu-qos
FTOS(conf-ip-aclo-cpuqos)#permit ospf
FTOS(conf-ip-aclo-cpuqos)#exit

FTOS(conf)#ip access-list extended bgp cpu-qos
FTOS(conf-ip-aclo-cpuqos)#permit bgp
FTOS(conf-ip-aclo-cpuqos)#exit

FTOS(conf)#mac access-list extended lacp cpu-qos
FTOS(conf-mac-aclo-cpuqos)#permit lacp
FTOS(conf-mac-aclo-cpuqos)#exit

FTOS(conf)#ipv6 access-list ipv6-icmp cpu-qos
FTOS(conf-ipv6-aclo-cpuqos)#permit icmp
FTOS(conf-ipv6-aclo-cpuqos)#exit

FTOS(conf)#ipv6 access-list ipv6-vrrp cpu-qos
FTOS(conf-ipv6-aclo-cpuqos)#permit vrrp
FTOS(conf-ipv6-aclo-cpuqos)#exit

Create QoS Input Policy

FTOS(conf)#qos-policy-in rate_limit_200k cpu-qos
FTOS(conf-in-qos-policy-cpuqos)#rate-police 200 40 peak 500 40
FTOS(conf-in-qos-policy-cpuqos)#exit

FTOS(conf)#qos-policy-in rate_limit_400k cpu-qos
FTOS(conf-in-qos-policy-cpuqos)#rate-police 400 50 peak 600 50
FTOS(conf-in-qos-policy-cpuqos)#exit

FTOS(conf)#qos-policy-in rate_limit_500k cpu-qos
FTOS(conf-in-qos-policy-cpuqos)#rate-police 500 50 peak 1000 50
FTOS(conf-in-qos-policy-cpuqos)#exit

Create QoS Class Map

FTOS(conf)#class-map match-any class_ospf cpu-qos
FTOS(conf-class-map-cpuqos)#match ip access-group ospf
FTOS(conf-class-map-cpuqos)#exit

FTOS(conf)#class-map match-any class_bgp cpu-qos
FTOS(conf-class-map-cpuqos)#match ip access-group bgp
FTOS(conf-class-map-cpuqos)#exit

FTOS(conf)#class-map match-any class_lacp cpu-qos
FTOS(conf-class-map-cpuqos)#match mac access-group lacp
FTOS(conf-class-map-cpuqos)#exit

FTOS(conf)#class-map match-any class-ipv6-icmp cpu-qos
FTOS(conf-class-map-cpuqos)#match ipv6 access-group ipv6-icmp
FTOS(conf-class-map-cpuqos)#exit
Match QoS Class Map to QoS Policy

FTOS(conf)#policy-map-input egressFP_rate_policy cpu-qos
FTOS(conf-policy-map-in-cpuqos)#class-map class ospf qos-policy rate_limit_500k
FTOS(conf-policy-map-in-cpuqos)#class-map class bgp qos-policy rate_limit_400k
FTOS(conf-policy-map-in-cpuqos)#class-map class lacp qos-policy rate_limit_200k
FTOS(conf-policy-map-in-cpuqos)#class-map class-ipv6 qos-policy rate_limit_200k
FTOS(conf-policy-map-in-cpuqos)#exit

Create Control Plane Service Policy

FTOS(conf)#control-plane-cpuqos
FTOS(conf-control-cpuqos)#service-policy rate-limit-protocols egressFP_rate_policy
FTOS(conf-control-cpuqos)#exit

Configure CoPP for CPU queues

Controlling traffic on the CPU queues does not require ACL rules, but does require QoS policies.

CoPP for CPU queues converts the input rate from kbps to pps, assuming 64 bytes is the average packet size, and applies that rate to the corresponding queue. Consequently, 1 kbps is roughly equivalent to 2 pps.

The basics for creating a CoPP service policy is to create QoS policies for the desired CPU bound queue and associate it with a particular rate-limit. The QoS policies are assigned to a control-plane service policy for each port-pipe.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create a QoS input policy for the router and assign the policing.</td>
<td>qos-policy-input name cpu-qos</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Create an input policy-map to assign the QoS policy to the desired service queues.</td>
<td>policy-map--input name cpu-qos service-queue 0 qos-policy name</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>3</td>
<td>Enter Control Plane mode.</td>
<td>control-plane-cpuqos</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>4</td>
<td>Assign a CPU queue-based service policy on the control plane in cpu-qos mode. Enabling this command sets the queue rates according to those configured.</td>
<td>service-policy rate-limit-cpu-queues name</td>
<td>CONTROL-PLANE</td>
</tr>
</tbody>
</table>

Sample Config for CoPP CPU queue configuration

Create QoS Policy

FTOS#conf
FTOS(conf)#qos-policy-input cpuq_1
FTOS(conf-qos-policy-in)#rate-police 3000 40 peak 500 40
FTOS(conf-qos-policy-in)#exit
Assign QoS Policy to Queues

FTOS(conf)#policy-map-input cpuq_rate_policy cpu-qos
FTOS(conf-qos-policy-in)#service-queue 5 qos-policy cpuq_1
FTOS(conf-qos-policy-in)#service-queue 6 qos-policy cpuq_2
FTOS(conf-qos-policy-in)#service-queue 7 qos-policy cpuq_1

Create Control Plane Service Policy

FTOS#conf
FTOS(conf)#control-plane
FTOS(conf-control-plane)#service-policy rate-limit-cpu-queues cpuq_rate_policy

Show commands

Use the `show cpu-queue rate cp` command to view the rates for each queue.

```
FTOS#show cpu-queue rate cp
Service-Queue       Rate (PPS)
---------------------
Q0                   1300
Q1                   300
Q2                   300
Q3                   300
Q4                   2000
Q5                   400
Q6                   400
Q7                   1100
FTOS#
```

Use the `show ip protocol-queue-mapping` command to view the queue mapping for each configured protocol.

```
FTOS#show ip protocol-queue-mapping
Protocol     Src-Port     Dst-Port     TcpFlag  Queue  EgPort  Rate (kbps)
------------ ---------- ---------- ------ ------- ------ ------------
TCP (BGP)    any/179     179/any     _       Q6      CP    100
UDP (DHCP)   67/68      68/67       _       Q6/Q5   CP    _
UDP (DHCP-R) 67         67          _       Q6      CP    _
TCP (FTP)    any         21          _       Q6      CP    _
ICMP         any         any         _       Q6      CP    _
IGMP         any         any         _       Q7      CP    _
TCP (MSDP)   any/639     639/any     _       Q6      CP    _
UDP (NTP)    any         123         _       Q6      CP    _
OSPF         any         any         _       Q7      CP    _
PIM          any         any         _       Q7      CP    _
UDP (RIP)    any         520         _       Q7      CP    _
TCP (SSH)    any         22          _       Q6      CP    _
TCP (TELNET) any         23          _       Q6      CP    _
VRRP         any         any         _       Q7      CP    _
```
Use the `show mac protocol-queue-mapping` command to view the queue mapping for the MAC protocols.

```
FTOS# show mac protocol-queue-mapping

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Destination Mac</th>
<th>EtherType</th>
<th>Queue</th>
<th>EgPort</th>
<th>Rate (kbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARP</td>
<td>any</td>
<td>0x0806</td>
<td>Q5/Q6</td>
<td>CP</td>
<td></td>
</tr>
<tr>
<td>FRRP</td>
<td>01:01:e8:00:00:10/11, any, 0x88cc</td>
<td>Q7, CP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LACP</td>
<td>01:80:c2:00:00:02, any, 0x8809</td>
<td>Q7, CP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLDP</td>
<td>any</td>
<td>0x88cc</td>
<td>Q7, CP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GVRP</td>
<td>01:80:c2:00:00:21</td>
<td>any, 0x88cc</td>
<td>Q7, CP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STP</td>
<td>01:80:c2:00:00:00</td>
<td>any, 0x88cc</td>
<td>Q7, CP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISIS</td>
<td>01:80:c2:00:00:14/15</td>
<td>any, 0x88cc</td>
<td>Q7, CP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>09:00:2b:00:00:04/05</td>
<td>any, 0x88cc</td>
<td>Q7, CP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FTOS#
```

Use the `show ipv6 protocol-queue-mapping` command to view the queue mapping for IPv6 protocols.

```
FTOS# show ipv6 protocol-queue-mapping

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Src-Port</th>
<th>Dst-Port</th>
<th>TcpFlag</th>
<th>Queue</th>
<th>EgPort</th>
<th>Rate (kbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCP (BGP)</td>
<td>any/179</td>
<td>179/any</td>
<td>_</td>
<td>Q6</td>
<td>CP</td>
<td></td>
</tr>
<tr>
<td>ICMP</td>
<td>any</td>
<td>any</td>
<td>_</td>
<td>Q6</td>
<td>CP</td>
<td></td>
</tr>
<tr>
<td>VRRP</td>
<td>any</td>
<td>any</td>
<td>_</td>
<td>Q7</td>
<td>CP</td>
<td></td>
</tr>
</tbody>
</table>

FTOS#
```
Data Center Bridging (DCB)

The data center bridging (DCB) features are supported on the S4810 and S4820T.

This chapter describes the following data center bridging topics:

- Ethernet Enhancements in Data Center Bridging
- Enabling Data Center Bridging
- Configuring Priority-Based Flow Control
- Configuring Enhanced Transmission Selection
- Applying DCB Policies in a Switch Stack
- Configuring DCBx Operation
- Verifying DCB Configuration
- PFC and ETS Configuration Examples

Ethernet Enhancements in Data Center Bridging

The S4810 and S4820T can support loading two DCB_Config files:

- FCoE_DCB_Config
- iSCSI_DCB_Config

These files are located in the root directory C:/CONFIG_TEMPLATE. After copying the config files to the startup config and reloading the system.

The S4810 and S4820T support the following DCB features:

- Data center bridging exchange protocol (DCBx)
- Priority-based Flow Control (PFC)
- Enhanced Transmission Selection (ETS)

Data center bridging (DCB) refers to a set of IEEE Ethernet enhancements that provide data centers with a single, robust, converged network to support multiple traffic types, including local area network (LAN), server, and storage traffic. Through network consolidation, DCB results in reduced operational cost, simplified management, and easy scalability by avoiding the need to deploy separate application-specific networks.
For example, instead of deploying an Ethernet network for LAN traffic, additional storage area networks (SANs) to ensure lossless fiber-channel traffic, and a separate InfiniBand network for high-performance inter-processor computing within server clusters, only one DCB-enabled network is required in a data center. The Dell Force10 switches that support a unified fabric and consolidate multiple network infrastructures use a single input/output (I/O) device called a converged network adapter (CNA).

A CNA is a computer input/output device that combines the functionality of a host bus adapter (HBA) with a network interface controller (NIC). Multiple adapters on different devices for several traffic types are no longer required.

Data center bridging satisfies the needs of the following types of data center traffic in a unified fabric:

- LAN traffic consists of a large number of flows that are generally insensitive to latency requirements, while certain applications, such as streaming video, are more sensitive to latency. Ethernet functions as a best-effort network that may drop packets in case of network congestion. IP networks rely on transport protocols (for example, TCP) for reliable data transmission with the associated cost of greater processing overhead and performance impact.
- Storage traffic based on Fibre Channel media uses the SCSI protocol for data transfer. This traffic typically consists of large data packets with a payload of 2K bytes that cannot recover from frame loss. To successfully transport storage traffic, data center Ethernet must provide no-drop service with lossless links.
- Servers use InterProcess Communication (IPC) traffic within high-performance computing clusters to share information. Server traffic is extremely sensitive to latency requirements.

To ensure lossless delivery and latency-sensitive scheduling of storage and service traffic and I/O convergence of LAN, storage, and server traffic over a unified fabric, IEEE data center bridging adds the following extensions to a classical Ethernet network:

- 802.1Qbb - Priority-based Flow Control (PFC)
- 802.1Qaz - Enhanced Transmission Selection (ETS)
- 802.1Qau - Congestion Notification
- Data Center Bridging Exchange (DCBx) protocol

Note: In FTOS version 8.3.12.0, only the PFC, ETS, and DCBx features are supported in data center bridging.

**Priority-Based Flow Control**

In a data center network, priority-based flow control (PFC) manages large bursts of one traffic type in multiprotocol links so that it does not affect other traffic types and no frames are lost due to congestion. When PFC detects congestion on a queue for a specified priority, it sends a pause frame for the 802.1p priority traffic to the transmitting device. In this way, PFC ensures that large amounts of queued LAN traffic do not cause storage traffic to be dropped, and that storage traffic does not result in high latency for high-performance computing (HPC) traffic between servers.
PFC enhances the existing 802.3x pause and 802.1p priority capabilities to enable flow control based on 802.1p priorities (classes of service). Instead of stopping all traffic on a link (as performed by the traditional Ethernet pause mechanism), PFC pauses traffic on a link according to the 802.1p priority set on a traffic type. You can create lossless flows for storage and server traffic while allowing for loss in case of LAN traffic congestion on the same physical interface.

Figure 13-1 shows how PFC handles traffic congestion by pausing the transmission of incoming traffic with dot1p priority 4.

**Figure 13-1. Priority-Based Flow Control**

PFC is implemented as follows in the Dell Force10 operating software (FTOS):

- PFC is supported on specified 802.1p priority traffic (dot1p 0 to 7) and is configured per interface. However, only two lossless queues are supported on an interface: one for FCoE converged traffic and one for iSCSI storage traffic. You must configure the same lossless queues on all ports.
- PFC delay constraints place an upper limit on the transmit time of a queue after receiving a message to pause a specified priority.
- By default, PFC is enabled when DCB is enabled. If FCoE_DCB_Config and iSCSI_DCB_Config have not been loaded and DCB is disabled. When DCB is enabled globally, TX and RX cannot be enabled simultaneously on the interface for flow control and link-level flow control is disabled.
- During DCBx negotiation with a remote peer:
  - DCBx communicates with the remote peer by LLDP TLV to determine current policies, such as PFC support and ETS BW allocation.
  - If the negotiation fails and PFC is enabled on the port, any user-configured PFC input policies are applied. If no PFC input policy has been previously applied, the PFC default setting is used (no priorities configured). If you do not enable PFC on an interface, you can enable the 802.3x link-level pause function. By default, the link-level flow pause is disabled when DCBx and PFC are enabled. If no PFC input policy has been applied on the interface, the default PFC settings are used.
  - If DCBx negotiation is not successful (due to, for example, a version or TLV mismatch), DCBx will be disabled and PFC or ETS cannot be enabled.
  - PFC supports buffering to receive data that continues to arrive on an interface while the remote system reacts to the PFC operation.
  - PFC uses the DCB MIB IEEE 802.1azd2.5 and the PFC MIB IEEE 802.1bb-d2.2.
Enhanced Transmission Selection

Enhanced transmission selection (ETS) supports optimized bandwidth allocation between traffic types in multiprotocol (Ethernet, FCoE, SCSI) links. ETS allows you to divide traffic according to its 802.1p priority into different priority groups (traffic classes) and configure bandwidth allocation and queue scheduling for each group to ensure that each traffic type is correctly prioritized and receives its required bandwidth. For example, you can prioritize low-latency storage or server cluster traffic in a traffic class to receive more bandwidth and restrict best-effort LAN traffic assigned to a different traffic class.

Although you can configure strict-priority queue scheduling for a priority group, ETS introduces flexibility that allows the bandwidth allocated to each priority group to be dynamically managed according to the amount of LAN, storage, and server traffic in a flow. Unused bandwidth is dynamically allocated to prioritized priority groups. Traffic is queued according to its 802.1p priority assignment, while flexible bandwidth allocation and the configured queue-scheduling for a priority group is supported.

Figure 13-2 shows how ETS allows you to allocate bandwidth when different traffic types are classed according to 802.1p priority and mapped to priority groups.

Figure 13-2. Enhanced Transmission Selection

ETS uses the following traffic groupings to select multiprotocol traffic for transmission:

- Priority group: A group of 802.1p priorities used for bandwidth allocation and queue scheduling. All 802.1p priority traffic in a group should have the same traffic handling requirements for latency and frame loss.
- Group ID: A 4-bit identifier assigned to each priority group. Range is from 0 to 7.
- Group bandwidth: Percentage of available bandwidth allocated to a priority group.
- Group transmission selection algorithm (TSA): Type of queue scheduling used by a priority group.

ETS is implemented as follows in FTOS:

- ETS supports groups of 802.1p priorities that have:
  - PFC enabled or disabled
  - No bandwidth limit or no ETS processing
• Bandwidth allocated by the ETS algorithm is made available after strict-priority groups are serviced. If a priority group does not use its allocated bandwidth, the unused bandwidth is made available to other priority groups.
• For ETS traffic selection, an algorithm is applied to priority groups using:
  • Strict-priority shaping
  • ETS shaping
    Credit-based shaping is not supported.
• ETS uses the DCB MIB IEEE 802.1azd2.5.

Data Center Bridging Exchange Protocol (DCBx)

The data center bridging exchange (DCBx) protocol is disabled by default on the S4810 and S4820T; ETS is also disabled. DCBx allows a switch to automatically discover DCB-enabled peers and exchange configuration information. PFC and ETS use DCBx to exchange and negotiate parameters with peer devices. DCBx capabilities include:

• Discovery of DCB capabilities on peer-device connections
• Determination of possible mismatch in DCB configuration on a peer link
• Configuration of a peer device over a DCB link

DCBx requires the link layer discovery protocol (LLDP) to provide the path to exchange DCB parameters with peer devices. Exchanged parameters are sent in organizationally specific type, length, values (TLVs) in LLDP data units. For more information, refer to Chapter 29, Link Layer Discovery Protocol (LLDP). The following LLDP TLVs are supported for DCB parameter exchange:

• PFC parameters: PFC Configuration TLV and Application Priority Configuration TLV.
• ETS parameters: ETS Configuration TLV and ETS Recommendation TLV.

Data Center Bridging in a Traffic Flow

The following figure shows how DCB handles a traffic flow on an interface.
Enabling Data Center Bridging

Data center bridging (DCB) is automatically configured when FCoE or iSCSI Optimization are configured. Data center bridging supports converged enhanced Ethernet (CEE) in a data center network. DCB is disabled by default. It must be enabled to support CEE.

- Priority-based flow control
- Enhanced transmission selection
- Data center bridging exchange protocol
- FCoE initialization protocol (FIP) snooping

DCB processes virtual local area network (VLAN)-tagged packets and dot1p priority values. Untagged packets are treated with a dot1p priority of 0.

For DCB to operate effectively, you can classify ingress traffic according to its dot1p priority so that it maps to different data queues. The dot1p-queue assignments used are shown in Table 13-1.

To enable DCB, you must enable either the iSCSI Optimization configuration or the FCoE configuration. For information to configure iSCSI Optimization, refer to Enabling and Disabling iSCSI Optimization. For information to configure FCoE, refer to Step 1 in Configuring FIP Snooping.

To enable DCB with PFC buffers on a switch, enter the following commands, save the configuration, and reboot the system to allow the changes to take effect:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable DCB.</td>
<td>dcb enable</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
QoS dot1p Traffic Classification and Queue Assignment

DCB supports PFC, ETS, and DCBx to handle converged Ethernet traffic that is assigned to an egress queue according to the following quality of service (QoS) methods:

- **Honor dot1p**: Using the `service-class dynamic dot1p` command in INTERFACE configuration mode, you can honor dot1p priorities in ingress traffic at the port or global switch level (refer to Default dot1p to Queue Mapping).
- **Layer 2 class maps**: You can use dot1p priorities to classify traffic in a class map and apply a service policy to an ingress port to map traffic to egress queues (refer to Policy-based QoS Configurations).

**Note**: Dell Force10 does not recommend mapping all ingress traffic to a single queue when using PFC and ETS. Ingress traffic classification using the `service-class dynamic dot1p` command (honor dot1p) is recommended on all DCB-enabled interfaces. If you use L2 class maps to map dot1p priority traffic to egress queues, take into account the default dot1p-queue assignments in Table 13-1 and the maximum number of two lossless queues supported on a port (refer to Configuring Lossless Queues).

Although FTOS allows you to change the default dot1p priority-queue assignments (refer to Set dot1p Priorities for Incoming Traffic), DCB policies applied to an interface may become invalid if dot1p-queue mapping is reconfigured. If the configured DCB policy remains valid, the change in the dot1p-queue assignment is allowed.

### FTOS Behavior:

DCB is not supported if you enable link-level flow control on one or more interfaces (refer to Ethernet Pause Frames on page 469).

### Task Command Command Mode

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set PFC buffering on the DCB stack unit.</td>
<td>dcb stack-unit all pfc-buffering pfc-ports 64 pfc-queues 2</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**Note**: Save the configuration and reboot the system to save the pfc buffering configuration changes.
### Configuring Priority-Based Flow Control

Priority-based flow control (PFC) provides a flow control mechanism based on the 802.1p priorities in converged Ethernet traffic received on an interface and is enabled by default when DCB is enabled. As an enhancement to the existing Ethernet pause mechanism, PFC stops traffic transmission for specified priorities (CoS values) without impacting other priority classes. Different traffic types are assigned to different priority classes.

When traffic congestion occurs, PFC sends a pause frame to a peer device with the CoS priority values of the traffic that needs to be stopped. DCBx provides the link-level exchange of PFC parameters between peer devices. PFC allows network administrators to create zero-loss links for SAN traffic that requires no-drop service, while at the same time retaining packet-drop congestion management for LAN traffic.

To ensure complete no-drop service, you must apply the same DCB input policy with the same pause time and dot1p priorities on all PFC-enabled peer interfaces.

To configure PFC and apply a PFC input policy to an interface, follow these steps:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create a DCB input policy to apply pause or flow control for specified priorities using a configured delay time. Maximum: 32 alphanumeric characters.</td>
<td><code>dcb-input policy-name</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Configure the link delay used to pause specified priority traffic. One quantum is equal to a 512-bit transmission. Range (in quanta): 712-65535. Default: 45556 quantum in link delay.</td>
<td><code>pfc link-delay value</code></td>
<td>DCB INPUT POLICY</td>
</tr>
</tbody>
</table>

---

#### Table 13-1. dot1p Priority-Queue Assignment

<table>
<thead>
<tr>
<th>dot1p Value in Incoming Frame</th>
<th>Egress Queue Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Step</td>
<td>Task</td>
</tr>
<tr>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>3</td>
<td>Configure the CoS traffic to be stopped for the specified delay. Enter the 802.1p values of the frames to be paused. Range: 0-7. Default: None. Maximum number of loss less queues supported on the switch: 2. Separate priority values with a comma. Specify a priority range with a dash, for example: <code>pfc priority 1,3,5-7.</code></td>
</tr>
<tr>
<td>4</td>
<td>Enable the PFC configuration on the port so that the priorities are included in DCBx negotiation with peer PFC devices. Default: PFC mode is on.</td>
</tr>
<tr>
<td>5</td>
<td>(Optional) Enter a text description of the input policy. Maximum: 32 characters.</td>
</tr>
<tr>
<td>6</td>
<td>Exit DCB input policy configuration mode.</td>
</tr>
<tr>
<td>7</td>
<td>Enter interface configuration mode.</td>
</tr>
<tr>
<td>8</td>
<td>Apply the input policy with the PFC configuration to an ingress interface.</td>
</tr>
<tr>
<td>9</td>
<td>Repeat Steps 1 to 8 on all PFC-enabled peer interfaces to ensure lossless traffic service.</td>
</tr>
</tbody>
</table>
FTOS Behavior:

As soon as you apply a DCB policy with PFC enabled on an interface, DCBx starts exchanging information with PFC-enabled peers. The IEEE802.1Qbb, CEE and CIN versions of PFC TLV are supported. DCBx also validates PFC configurations that are received in TLVs from peer devices.

By applying a DCB input policy with PFC enabled, you enable PFC operation on ingress port traffic. To achieve complete lossless handling of traffic, you must also enable PFC on all DCB egress ports or configure the dot1p priority-queue assignment of PFC priorities to lossless queues (refer to Configuring Lossless Queues).

To remove a DCB input policy, including the PFC configuration it contains, use the `no dcb-input policy-name` command in INTERFACE Configuration mode. To disable PFC operation on an interface, use the `no pfc mode on` command in DCB Input Policy Configuration mode. PFC is enabled and disabled as the global DCB operation is enabled (`dcb enable`) or disabled (`no dcb enable`).

You can enable any number of 802.1p priorities for PFC. Queues to which PFC priority traffic is mapped are lossless by default. Traffic may be interrupted due to an interface flap (going down and coming up) when you reconfigure the lossless queues for no-drop priorities in a PFC input policy and re-apply the policy to an interface.

For PFC to be applied, the configured priority traffic must be supported by a PFC peer (as detected by DCBx).

To honor a PFC pause frame multiplied by the number of PFC-enabled ingress ports, the minimum link delay should be greater than the round-trip transmission time required by a peer.

If you apply an input policy with PFC disabled (`no pfc mode on`):
- Link-level flow control can be enabled on the interface (refer to Ethernet Pause Frames on page 469).
- PFC still allows you to configure lossless queues on a port to ensure no-drop handling of lossless traffic (refer to Configuring Lossless Queues).

PFC and link-level flow control cannot be enabled at the same time on an interface.

When you apply an input policy to an interface, an error message is displayed if:
- The PFC dot1p priorities result in more than two lossless port queues globally on the switch.
- Link-level flow control is already enabled. PFC and link-level flow control cannot be enabled at the same time on an interface.
- In a switch stack, you must configure all stacked ports with the same PFC configuration.
A DCB input policy for PFC applied to an interface may become invalid if dot1p-queue mapping is reconfigured (refer to Create Input Policy Maps in Chapter 38, Quality of Service (QoS)). This situation occurs when the new dot1p-queue assignment exceeds the maximum number (2) of lossless queues supported globally on the switch. In this case, all PFC configurations received from PFC-enabled peers are removed and re-synchronized with the peer devices.

Traffic may be interrupted when you reconfigure PFC no-drop priorities in an input policy or re-apply the policy to an interface.

**Configuring Lossless Queues**

DCB also supports the manual configuration of lossless queues on an interface when PFC mode is turned off and priority classes are disabled in a DCB input policy applied to the interface. The configuration of no-drop queues provides flexibility for ports on which PFC is not needed but lossless traffic should egress from the interface.

Lossless traffic egresses out the no-drop queues. Ingress dot1p traffic from PFC-enabled interfaces is automatically mapped to the no-drop egress queues.

**Prerequisite**: A DCB input policy with PFC configuration is applied to the interface with the following conditions:

- PFC mode is off (`no pfc mode on`).
- No PFC priority classes are configured (`no pfc priority priority-range`).

To configure lossless queues on a port interface, follow these steps:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter INTERFACE Configuration mode.</td>
<td><code>interface type slot/port</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Configure the port queues that will still function as no-drop queues for lossless traffic. For the dot1p-queue assignments, refer to Table 13-1. Maximum number of lossless queues globally supported on the switch: 2. Range: 0-3. Separate queue values with a comma; specify a priority range with a dash; for example: <code>pfc no-drop queues 1,3</code> or <code>pfc no-drop queues 2-3</code> Default: No lossless queues are configured.</td>
<td><code>pfc no-drop queues queue-range</code></td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>
Configuring the PFC Buffer in a Switch Stack

In a switch stack, you must configure all stacked ports with the same PFC configuration. In addition, you must configure a separate buffer of memory allocated exclusively to a service pool accessed by queues on which priority-based control flows are mapped.

These PFC-enabled queues ensure the lossless transmission of storage and server traffic. The buffer required for the PFC service pool is calculated based on the number of ports and port queues used by PFC traffic.

You can configure the size of the PFC buffer for all switches in a stack or all port pipes on a specified stack unit by entering the following commands on the master switch:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure the PFC buffer for all switches in the stack. Default: The PFC buffer is enabled on all ports on the stack unit.</td>
<td>[no] dcb stack-unit all pfc-buffering pfc-port {1-64} pfc-queues {1-2}</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Configure the PFC buffer for all port pipes in a specified stack unit by specifying the port-pipe number, number of PFC-enabled ports, and number of configured lossless queues. Valid stack-unit IDs are 0 to 5. The only valid port-set ID (port-pipe number) is 0.</td>
<td>[no] dcb stack-unit stack-unit-id [port-set port-set-id] pfc-buffering pfc-ports {1-64} pfc-queues {1-2}</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

FTOS Behavior:
If you configure PFC on a 40GbE port, count the 40GbE port as four PFC-enabled ports in the pfc-port number you enter in the command syntax.

To achieve lossless PFC operation, the PFC port count and queue number used for the reserved buffer size that is created must be greater than or equal to the buffer size required for PFC-enabled ports and lossless queues on the switch.

For the PFC buffer configuration to take effect, you must reload the stack or a specified stack unit (reload command at the EXEC Privilege level).
DCB and Switch Stacking Caveats for the S4820T

The following is a list of behaviors and limitations regarding the use of DCB over S4820T ports involved in switch stacking:

- You can enable DCB only on 40 Gig (QSPF+) ports.
- DCB is not supported over any of the 48 RJ-45 10 Gig ports while they are configured in stacking mode.
- You cannot configure stacking on any of the 48 RJ-45 10 Gig ports, if DCB is enabled on any of the 40 Gig stacking ports.
- If DCB is enabled, it will disable the ability to use any of the 48 10 Gig ports for stacking.

Configuring Enhanced Transmission Selection

Enhanced transmission selection (ETS) provides a way to optimize bandwidth allocation to outbound 802.1p classes of converged Ethernet traffic. Different traffic types have different service needs. Using ETS, you can create groups within an 802.1p priority class to configure different treatment for traffic with different bandwidth, latency, and best-effort needs.

For example, storage traffic is sensitive to frame loss; interprocess communication (IPC) traffic is latency-sensitive. ETS allows different traffic types to coexist without interruption in the same converged link by:

- Allocating a guaranteed share of bandwidth to each priority group
- Allowing each group to exceed its minimum guaranteed bandwidth if another group is not fully using its allotted bandwidth

To configure ETS and apply an ETS output policy to an interface, you must:

1. Create a QoS output policy with ETS scheduling and bandwidth allocation settings.
2. Create a priority group of 802.1p traffic classes.
3. Configure a DCB output policy in which you associate a priority group with a QoS ETS output policy.
4. Apply the DCB output policy to an interface.

ETS Prerequisites and Restrictions

The following prerequisites and restrictions apply when you configure ETS bandwidth allocation or queue scheduling and apply a QoS ETS output policy on an interface:

- Configuring ETS bandwidth allocation or a queue scheduler for dot1p priorities in a priority group is applicable if the DCBx version used on a port is CIN (refer to Configuring DCBx on an Interface) or CEE as a port version where CNA supports CEE and DUT port versions in AUTO or CEE mode.
• When allocating bandwidth or configuring a queue scheduler for dot1p priorities in a priority group on a DCBx CIN interface, take into account the CIN bandwidth allocation (Configuring Bandwidth Allocation for DCBx CIN) and dot1p-queue mapping (Table 13-1).

• Although an ETS output policy does not support WRED, ECN, rate shaping, and rate limiting because these parameters are not negotiated by DCBx with peer devices, you can apply a QoS output policy with WRED and/or rate shaping on a DCBx CIN-enabled interface (refer to Configure Port-based Rate Shaping on page 763 and Weighted Random Early Detection). In this case, the WRED or rate shaping configuration in the QoS output policy should take into account the bandwidth allocation or queue scheduler configured in the ETS output policy.

• You can only use a QoS ETS output policy in association with a priority group in a DCB output policy and cannot be applied to an interface as a normal QoS output policy (refer to Applying an ETS Output Policy for a Priority Group to an Interface and Create an output QoS policy in Chapter 38, Quality of Service (QoS)).

Note: The IEEE 802.1Qaz, CEE, and CIN versions of ETS are supported.

Creating a QoS ETS Output Policy

A QoS output policy that you create to optimize bandwidth on an output interface for specified priority traffic consists of the ETS settings used in DCBx negotiations with peer devices:

• Bandwidth percentage
• Queue scheduling

To create a QoS output policy with ETS settings, follow these steps:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create a QoS output policy to configure the ETS bandwidth allocation and scheduling for priority traffic. Maximum: 32 characters.</td>
<td>qos-policy-output policy-name ets</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
| 2    | (Optional) Configure the method used to schedule priority traffic in port queues.  
• strict - Strict priority traffic is serviced before any other queued traffic (refer to Strict-priority Queueing in Chapter 38, Quality of Service (QoS)).  
Note: If you configure a scheduling method, you cannot configure bandwidth allocation in Step 3. | scheduler value | POLICY-MAP-OUT-ETS |
<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>(Optional) Configure the bandwidth percentage allocated to priority traffic in port queues. Percentage range: 1 to 100% in units of 1%. The sum of bandwidth percentage assigned to dot1p priorities/queues in a priority group should be 100%. Default: None. <strong>Note:</strong> If you configure bandwidth allocation, you cannot configure a scheduling method in Step 2.</td>
<td>bandwidth-percentage percentage</td>
<td>POLICY-MAP-OUT-ETS</td>
</tr>
<tr>
<td>4</td>
<td>Exit ETS Output Policy Configuration mode.</td>
<td>exit</td>
<td>POLICY-MAP-OUT-ETS</td>
</tr>
</tbody>
</table>
**FTOS Behavior:**
Traffic in priority groups is assigned to strict-queue or WERR scheduling in an ETS output policy and is managed using the ETS bandwidth-assignment algorithm. FTOS de-queues all frames of strict-priority traffic before servicing any other queues. A queue with strict-priority traffic can starve other queues in the same port.

ETS-assigned bandwidth allocation and scheduling apply only to data queues, not to control queues.

FTOS supports hierarchical scheduling on an interface. FTOS control traffic is redirected to control queues as higher priority traffic with strict priority scheduling. After control queues drain out, the remaining data traffic is scheduled to queues according to the bandwidth and scheduler configuration in the ETS output policy. The available bandwidth calculated by the ETS algorithm is equal to the link bandwidth after scheduling non-ETS higher-priority traffic.

The configuration of bandwidth allocation and strict-queue scheduling is not supported at the same time for a priority group. If both are configured, the configured bandwidth allocation is ignored for priority-group traffic when you apply the output policy on an interface (refer to Applying an ETS Output Policy for a Priority Group to an Interface).

**Bandwidth assignment in a dot.1p priority-queue:** By default, equal bandwidth is assigned to each port queue and each dot1p priority in a priority group. Use the `bandwidth-percentage` command to configure bandwidth amounts in associated dot1p queues. When specified bandwidth is assigned to some port queues and not to others, the remaining bandwidth (100% minus assigned bandwidth amount) is equally distributed to unassigned non-strict priority queues in the priority group. The sum of the allocated bandwidth to all queues in a priority group should be 100% of the bandwidth on the link.

**Bandwidth assignment in a priority group:** By default, equal bandwidth is assigned to each priority group in the ETS output policy applied to an egress port if you did not configure bandwidth allocation. The sum of configured bandwidth allocation to dot1p priority traffic in all ETS priority groups must be 100%. You must allocate at least 1% of the total bandwidth to each priority group and queue. If you assign bandwidth to some priority groups but not to others, the remaining bandwidth (100% minus assigned bandwidth amount) is equally distributed to non-strict-priority groups which have no configured scheduler.

**Scheduling of priority traffic:** dot1p priority traffic on the switch is scheduled to the current queue mapping. dot1p priorities within the same queue should have the same traffic properties and scheduling method.

**ETS output-policy error:** If an error occurs in an ETS output-policy configuration, the configuration is ignored and the scheduler and bandwidth allocation settings are reset to the ETS default values (all priorities are in the same ETS priority group and bandwidth is allocated equally to each priority). If an error occurs when a port receives a peer's ETS configuration, the port's configuration is reset to the previously configured ETS output policy. If no ETS output policy was previously applied, the port is reset to the default ETS parameters.
Creating an ETS Priority Group

An ETS priority group specifies the range of 802.1p priority traffic to which a QoS output policy with ETS settings is applied on an egress interface. You can associate a priority group to more than one ETS output policy on different interfaces.

To create a priority group for ETS, follow these steps:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create an ETS priority group to use with an ETS output policy. Maximum: 32 characters.</td>
<td>priority-group group-name</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Configure the priority-group identifier. Range: 0 to 7. Default: None.</td>
<td>set-pgid value</td>
<td>PRIORITY-GROUP</td>
</tr>
<tr>
<td>3</td>
<td>Configure the 802.1p priorities for the traffic on which you want to apply an ETS output policy. Range: 0 to 7. Default: None. Separate priority values with a comma. Specify a priority range with a dash. For example: priority-list 3,5-7.</td>
<td>priority-list value</td>
<td>PRIORITY-GROUP</td>
</tr>
<tr>
<td>4</td>
<td>Exit priority-group configuration mode.</td>
<td>exit</td>
<td>PRIORITY-GROUP</td>
</tr>
<tr>
<td>5</td>
<td>Repeat Steps 1 to 4 to configure all remaining dot1p priorities in an ETS priority group.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FTOS Behavior:**

A priority group consists of 802.1p priority values that are grouped together for similar bandwidth allocation and scheduling, and that share the same latency and loss requirements. All 802.1p priorities mapped to the same queue should be in the same priority group.

All 802.1p priorities should be configured in priority groups associated with an ETS output policy (refer to Applying an ETS Output Policy for a Priority Group to an Interface). You can assign each dot1p priority to only one priority group.

By default:
- All 802.1p priorities are grouped in priority group 0.
- 100% of the port bandwidth is assigned to priority group 0. The complete bandwidth is equally assigned to each priority class so that each class has 12 to 13%.

The maximum number of priority groups supported in ETS output policies on an interface is equal to the number of data queues (4) on the port. The 802.1p priorities in a priority group can map to multiple queues.

If you configure more than one priority queue as strict priority or more than one priority group as strict priority, the higher numbered priority queue is given preference when scheduling data traffic.
## Applying an ETS Output Policy for a Priority Group to an Interface

To apply ETS on egress port traffic, you must associate a priority group with an ETS output policy which has scheduling and bandwidth configuration in a DCB output policy, and then apply the output policy to an interface.

To apply ETS on egress port traffic, follow these steps:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create a DCB output policy to associate an ETS configuration with priority traffic. Maximum: 32 alphanumeric characters.</td>
<td>dcb-output policy-name</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Enable the ETS configuration so that scheduling and bandwidth allocation configured in an ETS output policy or received in a DCBx TLV from a peer can take effect on an interface. Default: ETS mode is on.</td>
<td>ets mode on</td>
<td>DCB OUTPUT POLICY</td>
</tr>
<tr>
<td>3</td>
<td>Associate the 802.1p priority traffic in a priority group with the ETS configuration in a QoS output policy.</td>
<td>priority-group group-name qos-policy ets-policy-name</td>
<td>DCB OUTPUT POLICY</td>
</tr>
<tr>
<td>4</td>
<td>(Optional) Enter a text description of the output policy. Maximum: 32 characters.</td>
<td>description text</td>
<td>DCB OUTPUT POLICY</td>
</tr>
<tr>
<td>5</td>
<td>Repeat Steps 1 to 4 to configure all remaining ETS priority groups with an ETS output policy.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Exit DCB Output Policy Configuration mode.</td>
<td>exit</td>
<td>DCB OUTPUT POLICY</td>
</tr>
<tr>
<td>7</td>
<td>Enter INTERFACE Configuration mode.</td>
<td>interface type slot/port</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>8</td>
<td>Apply the output policy with the ETS configuration to an egress interface.</td>
<td>dcb-policy output policy-name</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>
FTOS Behavior:
Create a DCB output policy to associate a priority group with an ETS output policy with scheduling and bandwidth configuration. You can apply a DCB output policy on multiple egress ports.

The ETS configuration associated with 802.1p priority traffic in a DCB output policy is used in DCBx negotiation with ETS peers.

When you apply an ETS output policy to an interface, ETS-configured scheduling and bandwidth allocation take precedence over any configured settings in the QoS output policies.

To remove an ETS output policy from an interface, use the no dcb-policy output policy-name command. DCB and ETS are both disabled by default. When DCB is enabled, ETS is enabled on all interfaces that have the default ETS configuration applied.

If you disable ETS in an output policy applied to an interface (the no ets mode on command), any previously configured QoS parameters at the interface or global level take effect. If QoS settings are configured at the interface or global level and in an output policy map (the service-policy output command), the QoS configuration in the output policy take precedence.

ETS Operation with DCBx

In DCBx negotiation with peer ETS devices, ETS configuration is handled as follows:

- ETS TLVs are supported in DCBx versions CIN, CEE, and IEEE2.5.
- ETS operational parameters are determined by the DCBx port-role configurations (see Configuring DCBx Operation).
- ETS configurations received from TLVs from a peer are validated.
- In case of a hardware limitation or TLV error:
  - DCBx operation on an ETS port goes down.
  - New ETS configurations are ignored and existing ETS configurations are reset to the previously configured ETS output policy on the port or to the default ETS settings if no ETS output policy was previously applied.
- ETS operates with legacy DCBx versions as follows:
  - In the CEE version, the priority group/traffic class group (TCG) ID 15 represents a non-ETS priority group. Any priority group configured with a scheduler type is treated as a strict-priority group and is given the priority-group (TCG) ID 15.
  - The CIN version supports two types of strict-priority scheduling:
    - Group strict priority: Allows a single priority flow in a priority group to increase its bandwidth usage to the bandwidth total of the priority group. A single flow in a group can use all the bandwidth allocated to the group.
    - Link strict priority: Allows a flow in any priority group to increase to the maximum link bandwidth.

CIN supports only the dot1p priority-queue assignment in a priority group. To configure a dot1p priority flow in a priority group to operate with link strict priority, you must configure:
- The dot1p priority for strict-priority scheduling (strict-priority command; Strict-priority Queuing)
- The priority group for strict-priority scheduling (scheduler strict command; Creating a QoS ETS Output Policy)

If you configure only the priority group in an ETS output policy or only the dot1p priority for strict-priority scheduling, the flow is handled with group strict priority.

Configuring Bandwidth Allocation for DCBx CIN

After you apply an ETS output policy to an interface, if the DCBx version used in your data center network is CIN, you may need to configure a QoS output policy to overwrite the default CIN bandwidth allocation. This default setting divides the bandwidth allocated to each port queue equally between the dot1p priority traffic assigned to the queue.

For more information, refer to Allocate bandwidth to queue.

To create a QoS output policy that allocates different amounts of bandwidth to the different traffic types/dot1p priorities assigned to a queue and apply the output policy to the interface, follow these steps.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create a QoS output policy. Maximum: 32 alphanumeric characters.</td>
<td>qos-policy-output output-policy-name</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Configure the percentage of bandwidth to be allocated to the dot1p priority/queue traffic in the associated L2 class map. Default: None.</td>
<td>bandwidth-percentage percentage</td>
<td>QoS OUTPUT POLICY</td>
</tr>
<tr>
<td>3</td>
<td>Repeat Step 2 to configure bandwidth percentages for other priority queues on the port.</td>
<td>bandwidth-percentage percentage</td>
<td>QoS OUTPUT POLICY</td>
</tr>
<tr>
<td>4</td>
<td>Exit QoS Output Policy Configuration mode.</td>
<td>exit</td>
<td>QoS OUTPUT POLICY</td>
</tr>
<tr>
<td>5</td>
<td>Enter INTERFACE Configuration mode.</td>
<td>interface type slot/port</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>6</td>
<td>Apply the QoS output policy with the bandwidth percentage for specified priority queues to an egress interface.</td>
<td>service-policy output output-policy-name</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>
Applying DCB Policies in a Switch Stack

**Note:** The S4820T does not support DCB on any of the 48 RJ-45 10 Gigabit stacking links.

You can apply a DCB input policy with PFC configuration to all stacked ports in a switch stack or on a stacked switch. You can apply different DCB input policies to different stacked switches.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apply the specified DCB input policy on all ports of the switch stack or a single stacked switch.</td>
<td>`dcb-policy input stack-unit {all</td>
<td>stack-unit-id} stack-ports all dcbb-input-policy-name`</td>
</tr>
</tbody>
</table>

**FTOS Behavior:**

Entering the command removes all DCB input policies applied to stacked ports.

A `dcb-policy input stack-unit all` command overwrites any previous `dcb-policy input stack-unit stack-unit-id` configurations. Similarly, a `dcb-policy input stack-unit stack-unit-id` command overwrites any previous `dcb-policy input stack-unit all` configuration.

Entering the **no dcb-policy input stack-unit all** command removes all DCB input policies applied to stacked ports and resets PFC to its default settings. The **no dcb-policy input stack-unit stack-unit-id** command removes only the DCB input policy applied to the specified switch.

You can apply a DCB output policy with ETS configuration to all stacked ports in a switch stack or an individual stacked switch. In addition, you can apply different DCB output policies to different stack units.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apply the specified DCB output policy on all ports of the switch stack or a stacked switch.</td>
<td>`dcb-policy output stack-unit {all</td>
<td>stack-unit-id} stack-ports all dcbb-output-policy-name`</td>
</tr>
</tbody>
</table>

**FTOS Behavior:**

Entering the command removes all DCB input policies applied to stacked ports.

A `dcb-policy output stack-unit all` command overwrites any previous `dcb-policy output stack-unit stack-unit-id` configurations. Similarly, a `dcb-policy output stack-unit stack-unit-id` command overwrites any previous `dcb-policy output stack-unit all` configuration.

Entering the **no dcb-policy output stack-unit all** command removes all DCB output policies applied to stacked ports. The **no dcb-policy output stack-unit stack-unit-id** command removes only the DCB output policy applied to the specified switch.
Configuring DCBx Operation

The data center bridging exchange protocol (DCBx) is used by DCB devices to exchange configuration information with directly connected peers using the link layer discovery protocol (LLDP) protocol. DCBx can detect the mis-configuration of a peer DCB device, and optionally, configure peer DCB devices with DCB feature settings to ensure consistent operation in a data center network.

DCBx is a prerequisite for using DCB features, such as priority-based flow control (PFC) and enhanced traffic selection (ETS), to exchange link-level configurations in a converged Ethernet environment. DCBx is also deployed in topologies that support lossless operation for FCoE or iSCSI traffic. In these scenarios, all network devices are DCBx-enabled (DCBx is enabled end-to-end). For more information about how these features are implemented and used, refer to:

- Configuring Priority-Based Flow Control
- Configuring Enhanced Transmission Selection
- FIP Snooping
- Chapter 13, Data Center Bridging (DCB)

The following versions of DCBx are supported CIN, CEE, and IEEE2.5.

**Prerequisite:** DCBx requires the LLDP to be enabled on all DCB devices.

**DCBx Operation**

DCBx performs the following operations:

- Discovers DCB configuration (such as PFC and ETS) in a peer device.
- Detects DCB mis-configuration in a peer device; that is, when DCB features are not compatibly configured on a peer device and the local switch. Mis-configuration detection is feature-specific because some DCB features support asymmetric configuration.
- Reconfigures a peer device with the DCB configuration from its configuration source if the peer device is willing to accept configuration.
- Accepts the DCB configuration from a peer if a DCBx port is in “willing” mode to accept a peer’s DCB settings and then internally propagates the received DCB configuration to its peer ports.

**DCBx Port Roles**

Use the following DCBx port roles to enable the auto-configuration of DCBx-enabled ports and propagate DCB configurations learned from peer DCBx devices internally to other switch ports:

- Auto-upstream: The port advertises its own configuration to DCBx peers and is willing to receive peer configuration. The port also propagates its configuration to other ports on the switch.

  The first auto-upstream that is capable of receiving a peer configuration is elected as the *configuration source*. The elected configuration source then internally propagates the configuration to other auto-upstream and auto-downstream ports. A port that receives an internally propagated configuration overwrites its local configuration with the new parameter values.
When an auto-upstream port (besides the configuration source) receives and overwrites its configuration with internally propagated information, one of the following actions is taken:

- If the peer configuration received is compatible with the internally propagated port configuration, the link with the DCBx peer is enabled.
- If the received peer configuration is not compatible with the currently configured port configuration, the link with the DCBx peer port is disabled and a syslog message for an incompatible configuration is generated. The network administrator must then reconfigure the peer device so that it advertises a compatible DCB configuration.

The configuration received from a DCBx peer or from an internally propagated configuration is not stored in the switch’s running configuration.

On a DCBx port in an auto-upstream role, the PFC and application priority TLVs are enabled. ETS recommend TLVs are disabled and ETS configuration TLVs are enabled.

- Auto-downstream: The port advertises its own configuration to DCBx peers but is not willing to receive remote peer configuration. The port always accepts internally propagated configurations from a configuration source. An auto-downstream port that receives an internally propagated configuration overwrites its local configuration with the new parameter values.

When an auto-downstream port receives and overwrites its configuration with internally propagated information, one of the following actions is taken:

- If the peer configuration received is compatible with the internally propagated port configuration, the link with the DCBx peer is enabled.
- If the received peer configuration is not compatible with the currently configured port configuration, the link with the DCBx peer port is disabled and a syslog message for an incompatible configuration is generated. The network administrator must then reconfigure the peer device so that it advertises a compatible DCB configuration.

The internally propagated configuration is not stored in the switch's running configuration.

On a DCBx port in an auto-downstream role, all PFC, application priority, ETS recommend, and ETS configuration TLVs are enabled.

- Configuration source: The port is configured to serve as a source of configuration information on the switch. Peer DCB configurations received on the port are propagated to other DCBx auto-configured ports. If the peer configuration is compatible with a port configuration, DCBx is enabled on the port.

On a configuration-source port, the link with a DCBx peer is enabled when the port receives a DCB configuration that can be internally propagated to other auto-configured ports.

The configuration received from a DCBx peer is not stored in the switch’s running configuration.

On a DCBx port that is the configuration source, all PFC and application priority TLVs are enabled. ETS recommend TLVs are disabled and ETS configuration TLVs are enabled.

- Manual: The port is configured to operate only with administrator-configured settings and does not auto-configure with DCB settings received from a DCBx peer or from an internally propagated configuration from the configuration source. If you enable DCBx, ports in Manual mode advertise their configurations to peer devices but do not accept or propagate internal or external configurations.

Unlike other user-configured ports, the configuration of DCBx ports in Manual mode is saved in the running configuration.

On a DCBx port in a manual role, all PFC, application priority, ETS recommend, and ETS configuration TLVs are enabled.

When making a configuration change to a DCBx port in a Manual role, Dell Force10 recommends that you shut down the interface using the `shutdown` command, change the configuration, then re-activate the interface using the `no shutdown` command.
Default DCBx port role: Manual.

**Note:** On a DCBx port, application priority TLV advertisements are handled as follows:
- The application priority TLV is transmitted only if the priorities in the advertisement match the configured PFC priorities on the port.
- On auto-upstream and auto-downstream ports:
  - If a configuration source is elected, the ports send an application priority TLV based on the application priority TLV received on the configuration-source port. When an application priority TLV is received on the configuration-source port, the auto-upstream and auto-downstream ports use the internally propagated PFC priorities to match against the received application priority. Otherwise, these ports use their locally configured PFC priorities in application priority TLVs.
  - If no configuration source is configured, auto-upstream and auto-downstream ports check to see that the locally configured PFC priorities match the priorities in a received application priority TLV.
- On manual ports: An application priority TLV is advertised only if the priorities in the TLV match the PFC priorities configured on the port.

**DCB Configuration Exchange**

The DCBx protocol supports the exchange and propagation of configuration information for the following DCB features.

- Enhanced transmission selection (ETS)
- Priority-based flow control (PFC)

DCBx uses the following methods to exchange DCB configuration parameters:

- Asymmetric: DCB parameters are exchanged between a DCBx-enabled port and a peer port without requiring that a peer port and the local port use the same configured values for the configurations to be compatible. For example, ETS uses an asymmetric exchange of parameters between DCBx peers.
- Symmetric: DCB parameters are exchanged between a DCBx-enabled port and a peer port with the requirement that each configured parameter value is the same for the configurations to be compatible. For example, PFC uses an symmetric exchange of parameters between DCBx peers.

**Configuration Source Election**

When an auto-upstream or auto-downstream port receives a DCB configuration from a peer, the port first checks to see if there is an active configuration source on the switch.

- If a configuration source already exists, the received peer configuration is checked against the local port configuration. If the received configuration is compatible, the DCBx marks the port as DCBx-enabled. If the configuration received from the peer is not compatible, a warning message is logged and the DCBx frame error counter is incremented. Although DCBx is operationally disabled, the port keeps the peer link up and continues to exchange DCBx packets. If a compatible peer configuration is later received, DCBx is enabled on the port.
- If there is no configuration source, a port may elect itself as the configuration source. A port may become the configuration source if the following conditions exist:
  - No other port is the configuration source.
  - The port role is auto-upstream.
• The port is enabled with link up and DCBx enabled.
• The port has performed a DCBx exchange with a DCBx peer.
• The switch is capable of supporting the received DCB configuration values through either a symmetric or asymmetric parameter exchange.

A newly elected configuration source propagates configuration changes received from a peer to the other auto-configuration ports. Ports receiving auto-configuration information from the configuration source ignore their current settings and use the configuration source information.

**Propagation of DCB Information**

When an auto-upstream or auto-downstream port receives a DCB configuration from a peer, the port acts as a DCBx client and checks if a DCBx configuration source exists on the switch.

• If a configuration source is found, the received configuration is checked against the currently configured values that are internally propagated by the configuration source. If the local configuration is compatible with the received configuration, the port is enabled for DCBx operation and synchronization.
• If the configuration received from the peer is not compatible with the internally propagated configuration used by the configuration source, the port is disabled as a client for DCBx operation and synchronization and a syslog error message is generated. The port keeps the peer link up and continues to exchange DCBx packets. If a compatible configuration is later received from the peer, the port is enabled for DCBx operation.

**Note:** DCB configurations internally propagated from a configuration source do not overwrite the configuration on a DCBx port in a manual role.

When a configuration source is elected, all auto-upstream ports other than the configuration source are marked as willing disabled. The internally propagated DCB configuration is refreshed on all auto-configuration ports and each port may begin configuration negotiation with a DCBx peer again.

**Auto-Detection and Manual Configuration of the DCBx Version**

When operating in Auto-Detection mode ([DCBx version auto](#) command in the [DCBx Configuration Procedure](#)), a DCBx port automatically detects the DCBx version on a peer port. Legacy CIN and CEE versions are supported in addition to the standard IEEE version 2.5 DCBx.

A DCBx port detects a peer version after receiving a valid frame for that version. The local DCBx port reconfigures to operate with the peer version and maintains the peer version on the link until one of the following conditions occurs:

• The switch reboots.
• The link is reset (goes down and up).
• User-configured CLI commands require the version negotiation to restart.
• The peer times out.
• Multiple peers are detected on the link.
If you configure a DCBx port to operate with a specific version (\texttt{DCBx version \{cee | cin | ieee-v2.5\}} command in the \texttt{DCBx Configuration Procedure}), DCBx operations are performed according to the configured version, including fast and slow transmit timers and message formats. If a DCBx frame with a different version is received, a syslog message is generated and the peer version is recorded in the peer status table. If the frame cannot be processed, it is discarded and the discard counter is incremented.

\textbf{Note:} Because DCBx TLV processing is best effort, it is possible that CIN frames may be processed when DCBx is configured to operate in CEE mode and vice versa. In this case, the unrecognized TLVs cause the unrecognized TLV counter to be incremented, but the frame is processed and is not discarded.

Legacy DCBx (CIN and CEE) supports the DCBx control state machine that is defined to maintain the sequence number and acknowledge number sent in the DCBx control TLVs.

\section*{DCBx Example}

\textit{Figure 13-4} shows how DCBx is used. The external 40GbE ports on the base module (ports 33 and 37) of two switches are used for uplinks configured as DCBx auto-upstream ports. The S4810 or S4820T is connected to third-party, top-of-rack (ToR) switches through 40GbE uplinks. The ToR switches are part of a Fibre Channel storage network.

- The internal ports (ports 1-32) connected to the 10GbE backplane are configured as auto-downstream ports.
  On the S4810 and S4820T, PFC and ETS use DCBx to exchange link-level configuration with DCBx peer devices.
Figure 13-4. DCBx Sample Topology
### DCBx Prerequisites and Restrictions

The following prerequisites and restrictions apply when you configure DCBx operation on a port:

- DCBx requires LLDP in both send (TX) and receive (RX) mode to be enabled on a port interface (`protocol lldp mode` command; refer to the example in CONFIGURATION versus INTERFACE Configurations in the Link Layer Discovery Protocol (LLDP) chapter). If a multiple DCBx peer ports are detected on a local DCBx interface, LLDP is shut down.

- The CIN version of DCBx supports only PFC, ETS, and FCOE; it does not support iSCSI, backward congestion management (BCN), logical link down (LLDF), and network interface virtualization (NIV).

### DCBx Configuration Procedure

To configure an S4810 or S4820T for DCBx operation in a data center network, you must:

1. Configure ToR- and FCF-facing interfaces as auto-upstream ports.
2. Configure server-facing interfaces as auto-downstream ports.
3. Configure a port to operate in a configuration-source role.

To verify the DCBx configuration on a port, use the `show interface DCBx detail` command (Figure 13-16).

Configure DCBx operation at the interface level on a switch or globally on the switch.

**Prerequisite:** DCBx requires LLDP to be enabled to advertise DCBx TLVs to peers. For more information, refer to Chapter 29, Link Layer Discovery Protocol (LLDP).

### Configuring DCBx on an Interface

To configure DCBx operation on an interface, follow these steps:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter INTERFACE Configuration mode.</td>
<td><code>interface type slot/port</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Enter LLDP Configuration mode to enable DCBx operation.</td>
<td><code>[no] protocol lldp</code></td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>
### Step 3
Configure the DCBx version used on the interface, where:
- `auto` configures the port to operate using the DCBx version received from a peer.
- `cee` configures the port to use CEE (Intel 1.01).
- `cin` configures the port to use Cisco-Intel-Nuova (DCBx 1.0).
- `ieee-v2.5` configures the port to use IEEE 802.1Qaz (Draft 2.5).
Default: Auto.

```plaintext
[no] DCBx version {auto | cee | cin | ieee-v2.5}
```

**Command Mode**: PROTOCOL LLDP

### Step 4
Configure the DCBx port role used by the interface to exchange DCB information, where:
- `auto-upstream` configures the port to receive a peer configuration. The configuration source is elected from auto-upstream ports.
- `auto-downstream` configures the port to accept the internally propagated DCB configuration from a configuration source.
- `config-source` configures the port to serve as the configuration source on the switch.
- `manual` configures the port to operate only on administer-configured DCB parameters. The port does not accept a DCB configuration received from a peer or a local configuration source.

```plaintext
[no] DCBx port-role {config-source | auto-downstream | auto-upstream | manual}
```

**Command Mode**: PROTOCOL LLDP

### Step 5
**On manual ports only:**
Configure the PFC and ETS TLVs advertised to DCBx peers, where:
- `ets-conf` enables the advertisement of ETS Configuration TLVs.
- `ets-reco` enables the advertisement of ETS Recommend TLVs.
- `pfc` enables the advertisement of PFC TLVs.
Default: All PFC and ETS TLVs are advertised.

**Note:** You can configure the transmission of more than one TLV type at a time; for example: advertise DCBx-tlv ets-conf ets-reco. You can enable ETS recommend TLVs (ets-reco) only if ETS configuration TLVs (ets-conf) are enabled. To disable TLV transmission, use the `no` form of the command; for example, no advertise DCBx-tlv pfc ets-reco.

```plaintext
[no] advertise DCBx-tlv {ets-conf | ets-reco | pfc}
```

**Command Mode**: PROTOCOL LLDP
To globally configure DCBx operation on a switch, follow these steps:

### Configuring DCBx Globally on the Switch

To globally configure DCBx operation on a switch, follow these steps:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter Global Configuration mode.</td>
<td>configure</td>
<td>EXEC PRIVILEGE</td>
</tr>
<tr>
<td>2</td>
<td>Enter LLDP Configuration mode to enable DCBx operation.</td>
<td>[no] protocol lldp</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>3</td>
<td>Configure the DCBx version used on all interfaces not already configured to exchange DCB information, where:</td>
<td>[no] DCBx version {auto</td>
<td>cee</td>
</tr>
</tbody>
</table>

Note: You can configure the DCBx port role used by interfaces to exchange DCB information by using the `DCBx port-role` command in INTERFACE Configuration mode (see Step 3 in Configuring DCBx Globally on the Switch).
<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
</table>
| 4    | Configure the PFC and ETS TLVs to be advertised on un-configured interfaces with a manual port-role, where:  
• ets-conf enables transmission of ETS Configuration TLVs.  
• ets-reco enables transmission of ETS Recommend TLVs.  
• pfc enables transmission of PFC TLVs.  
**Note:** You can configure the transmission of more than one TLV type at a time. ETS recommend TLVs (ets-reco) can be enabled only if ETS configuration TLVs (ets-conf) are enabled. To disable TLV transmission, use the no form of the command; for example, no advertise DCBx-tlv pfc ets-reco.  
Default: All TLV types are enabled. | [no] advertise DCBx-tlv  
{ets-conf | ets-reco | pfc}  
[ets-conf | ets-reco | pfc]  
[ets-conf | ets-reco | pfc] | PROTOCOL LLDP |
| 5    | Configure the Application Priority TLVs to be advertised on un-configured interfaces with a manual port-role, where:  
• fcoe enables the advertisement of FCoE in Application Priority TLVs.  
• iscsi enables the advertisement of iSCSI in Application Priority TLVs.  
Default: Application Priority TLVs are enabled and advertise FCoE and iSCSI.  
**Note:** To disable TLV transmission, use the no form of the command; for example, no advertise DCBx-appln-tlv iscsi. | [no] advertise DCBx-appln-tlv  
{fcoe | iscsi} | PROTOCOL LLDP |
| 6    | Configure the FCoE priority advertised for the FCoE protocol in Application Priority TLVs.  
The priority-bitmap range is from 1 to FF.  
Default: 0x8. | [no] fcoe priority-bits  
priority-bitmap | PROTOCOL LLDP |
| 7    | Configure the iSCSI priority advertised for the iSCSI protocol in Application Priority TLVs.  
The priority-bitmap range is from 1 to FF.  
Default: 0x10. | [no] iscsi priority-bits  
priority-bitmap | PROTOCOL LLDP |
DCBx Error Messages

An error in DCBx operation is displayed using the syslog messages:

- **LLDP_MULTIPLE_PEER_DETECTED**: DCBx is operationally disabled after detecting more than one DCBx peer on the port interface.
- **LLDP_PEER_AGE_OUT**: DCBx is disabled as a result of LLDP timing out on a DCBx peer interface.
- **DSM_DCBx_PEER_VERSION_CONFLICT**: A local port expected to receive the IEEE, CIN, or CEE version in a DCBx TLV from a remote peer but received a different, conflicting DCBx version.
- **DSM_DCBx_PFC_PARAMETERS_MATCH** and **DSM_DCBx_PFC_PARAMETERS_MISMATCH**: A local DCBx port received a compatible (match) or incompatible (mismatch) PFC configuration from a peer.
- **DSM_DCBx_ETS_PARAMETERS_MATCH** and **DSM_DCBx_ETS_PARAMETERS_MISMATCH**: A local DCBx port received a compatible (match) or incompatible (mismatch) ETS configuration from a peer.
- **LLDP_UNRECOGNISED_DCBx_TLV_RECEIVED**: A local DCBx port received an unrecognized DCBx TLV from a peer.

Debugging DCBx on an Interface

To enable DCBx debug traces for all or a specific control path, use the following command:

```
<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable DCBx debugging, where:</td>
<td>debug DCBx {all</td>
<td>auto-detect-timer</td>
</tr>
<tr>
<td>all: Enables all DCBx debugging operations.</td>
<td>config-exchng</td>
<td>fail</td>
</tr>
<tr>
<td>auto-detect-timer: Enables traces for DCBx auto-detect timers.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>config-exchng: Enables traces for DCBx configuration exchanges.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fail: Enables traces for DCBx failures.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mgmt: Enables traces for DCBx management frames.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>resource: Enables traces for DCBx system resource frames.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sem: Enables traces for the DCBx state machine.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tlv: Enables traces for DCBx TLVs.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Verifying DCB Configuration

Use the `show` commands in Table 13-2 to display DCB configurations.

Table 13-2. Displaying DCB Configurations

<table>
<thead>
<tr>
<th>Command</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>show dot1p-queue mapping</code></td>
<td>Displays the current 802.1p priority-queue mapping.</td>
</tr>
<tr>
<td><code>show dcb [stack-unit unit-number]</code></td>
<td>Displays data center bridging status, number of PFC-enabled ports, and number of PFC-enabled queues. On the master switch in a stack, you can specify a stack-unit number. Range is : 0 to 5.</td>
</tr>
<tr>
<td><code>show qos dcb-input [pfc-profile]</code></td>
<td>Displays the PFC configuration in a DCB input policy.</td>
</tr>
<tr>
<td><code>show qos dcb-output [ets-profile]</code></td>
<td>Displays the ETS configuration in a DCB output policy.</td>
</tr>
<tr>
<td><code>show qos priority-groups</code></td>
<td>Displays the ETS priority groups configured on the switch, including the 802.1p priority classes and ID of each group.</td>
</tr>
<tr>
<td><code>show interface port-type slot/port pfc</code></td>
<td>Displays the PFC configuration applied to ingress traffic on an interface, including priorities and link delay.</td>
</tr>
<tr>
<td>`{summary</td>
<td>detail}`</td>
</tr>
<tr>
<td><code>show interface port-type slot/port pfc statistics</code></td>
<td>Displays counters for the PFC frames received and transmitted (by dot1p priority class) on an interface.</td>
</tr>
<tr>
<td><code>show interface port-type slot/port ets</code></td>
<td>Displays the ETS configuration applied to egress traffic on an interface, including priority groups with priorities and bandwidth allocation.</td>
</tr>
<tr>
<td>`{summary</td>
<td>detail}`</td>
</tr>
<tr>
<td><code>show interface port-type slot/port DCBx detail</code></td>
<td>Displays the DCBx configuration on an interface.</td>
</tr>
<tr>
<td>`show stack-unit {0-11</td>
<td>all} stack ports all pfc details`</td>
</tr>
<tr>
<td>`show stack-unit {0-11</td>
<td>all} stack ports all ets details`</td>
</tr>
</tbody>
</table>

Figure 13-5. `show dot1p-queue mapping` Command Example

```
FTOS(conf)# show dot1p-queue-mapping
Dot1p Priority: 0 1 2 3 4 5 6 7
Queue : 0 0 0 1 2 3 3 3
```

Figure 13-6. `show dcb` Command Example

```
FTOS# show dcb
stack-unit 0 port-set 0
DCB Status : Enabled
PFC Port Count : 56 (current), 56 (configured)
PFC Queue Count : 2 (current), 2 (configured)
```
Figure 13-7. show qos dcb-input Command Example

FTOS(conf)# show qos dcb-input
dcb-input pfc-profile
  pfc link-delay 32
  pfc priority 0-1
dcb-input pfc-profile1
  no pfc mode on
  pfc priority 6-7

Figure 13-8. show qos dcb-output Command Example

FTOS# show qos dcb-output
dcb-output ets
  priority-group san qos-policy san
  priority-group ipc qos-policy ipc
  priority-group lan qos-policy lan

Figure 13-9. show qos priority-groups Command Example

FTOS#show qos priority-groups
priority-group ipc
  priority-list 4
  set-pgid 2
Figure 13-10. show interfaces pfc summary Command Example

FTOS# show interfaces tengigabitethernet 0/49 pfc summary
Interface TenGigabitEthernet 0/49
  Admin mode is on
  Admin is enabled
  Remote is enabled, Priority list is 4
  Remote Willing Status is enabled
  Local is enabled
  Oper status is Recommended
  PFC DCBx Oper status is Up
  State Machine Type is Feature
  TLV Tx Status is enabled
  PFC Link Delay 45556 pause quanta
  Application Priority TLV Parameters:
  FCOE TLV Tx Status is disabled
  ISCSI TLV Tx Status is disabled
  Local FCOE PriorityMap is 0x8
  Local ISCSI PriorityMap is 0x10
  Remote FCOE PriorityMap is 0x8
  Remote ISCSI PriorityMap is 0x8

FTOS# show interfaces tengigabitethernet 0/49 pfc detail
Interface TenGigabitEthernet 0/49
  Admin mode is on
  Admin is enabled
  Remote is enabled
  Remote Willing Status is enabled
  Local is enabled
  Oper status is recommended
  PFC DCBx Oper status is Up
  State Machine Type is Feature
  TLV Tx Status is enabled
  PFC Link Delay 45556 pause quanta
  Application Priority TLV Parameters:
  FCOE TLV Tx Status is disabled
  ISCSI TLV Tx Status is disabled
  Local FCOE PriorityMap is 0x8
  Local ISCSI PriorityMap is 0x10
  Remote FCOE PriorityMap is 0x8
  Remote ISCSI PriorityMap is 0x8

0 Input TLV pkts, 1 Output TLV pkts, 0 Error pkts, 0 Pause Tx pkts, 0 Pause Rx pkts

Table 13-3. show interface pfc summary Command Description

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface</td>
<td>Interface type with stack-unit and port number.</td>
</tr>
<tr>
<td>Admin mode is on</td>
<td>PFC Admin mode is on or off with a list of the configured PFC priorities. When PFC admin mode is on, PFC advertisements are enabled to be sent and</td>
</tr>
<tr>
<td>Admin is enabled</td>
<td>received from peers; received PFC configuration takes effect. The admin operational status for a DCBx exchange of PFC configuration is enabled or disabled.</td>
</tr>
</tbody>
</table>
Table 13-3. show interface pfc summary Command Description

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remote is enabled, Priority list</td>
<td>Operational status (enabled or disabled) of peer device for DCBx exchange of PFC configuration with a list of the configured PFC priorities. Willing status of peer device for DCBx exchange (Willing bit received in PFC TLV): enabled or disabled.</td>
</tr>
<tr>
<td>Remote Willing Status is enabled</td>
<td></td>
</tr>
<tr>
<td>Local is enabled</td>
<td>DCBx operational status (enabled or disabled) with a list of the configured PFC priorities.</td>
</tr>
</tbody>
</table>
| Operational status (local port)                           | Port state for current operational PFC configuration:  
Init: Local PFC configuration parameters were exchanged with peer.  
Recommend: Remote PFC configuration parameters were received from peer.  
Internally propagated: PFC configuration parameters were received from configuration source. |
| PFC DCBx Oper status                                       | Operational status for exchange of PFC configuration on local port: match (up) or mismatch (down).                                                                                                          |
| State Machine Type                                         | Type of state machine used for DCBx exchanges of PFC parameters:  
Feature - for legacy DCBx versions;  
Symmetric - for an IEEE version.                                                                                           |
| TLV Tx Status                                              | Status of PFC TLV advertisements: enabled or disabled.                                                                                                                                            |
| PFC Link Delay                                             | Link delay (in quanta) used to pause specified priority traffic.                                                                                                                                   |
| Application Priority TLV: FCoE TLV Tx Status               | Status of FCoE advertisements in application priority TLVs from local DCBx port: enabled or disabled.                                                                                                  |
| Application Priority TLV: iSCSI TLV Tx Status              | Status of iSCSI advertisements in application priority TLVs from local DCBx port: enabled or disabled.                                                                                               |
| Application Priority TLV: Local FCoE Priority Map          | Priority bitmap used by local DCBx port in FCoE advertisements in application priority TLVs.                                                                                                           |
| Application Priority TLV: Local iSCSI Priority Map         | Priority bitmap used by local DCBx port in iSCSI advertisements in application priority TLVs.                                                                                                        |
| Application Priority TLV: Remote FCoE Priority Map         | Status of FCoE advertisements in application priority TLVs from remote peer port: enabled or disabled.                                                                                               |
| Application Priority TLV: Remote iSCSI Priority Map        | Status of iSCSI advertisements in application priority TLVs from remote peer port: enabled or disabled.                                                                                               |
| PFC TLV Statistics: Input TLV pkts                        | Number of PFC TLVs received.                                                                                                                                                                           |
| PFC TLV Statistics: Output TLV pkts                       | Number of PFC TLVs transmitted.                                                                                                                                                                        |
| PFC TLV Statistics: Error pkts                            | Number of PFC error packets received.                                                                                                                                                    |
| PFC TLV Statistics: Pause Tx pkts                         | Number of PFC pause frames transmitted.                                                                                                                                                             |
Table 13-3.  show interface pfc summary Command Description

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFC TLV Statistics:</td>
<td></td>
</tr>
<tr>
<td>Pause Rx pkts</td>
<td>Number of PFC pause frames received</td>
</tr>
</tbody>
</table>

Figure 13-11.  show interface pfc statistics Command Example

```
FTOS#show interfaces te 0/0 pfc statistics
Interface TenGigabitEthernet 0/0
Priority  Received  PFC  Frames  Transmitted  PFC  Frames
        --------                  ---------------  ---------------
        0                      0                  0
        1                      0                  0
        2                      0                  0
        3                      0                  0
        4                      0                  0
        5                      0                  0
        6                      0                  0
        7                      0                  0
```
Figure 13-12. show interface ets summary Command Example

FTOS(conf-qos-policy-out-ets)#do sho int te 0/3 ets de

Interface TenGigabitEthernet 0/3
Max Supported TC Groups is 4
Number of Traffic Classes is 8
Admin mode is on

Admin Parameters :
---------------------
Admin is enabled

<table>
<thead>
<tr>
<th>TC-grp</th>
<th>Priority#</th>
<th>Bandwidth</th>
<th>TSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0,1,2</td>
<td>100%</td>
<td>ETS</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0%</td>
<td>SP</td>
</tr>
<tr>
<td>3</td>
<td>4,5,6,7</td>
<td>0%</td>
<td>SP</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Remote Parameters :
-------------------
Remote is disabled

Local Parameters :
------------------
Local is enabled

<table>
<thead>
<tr>
<th>TC-grp</th>
<th>Priority#</th>
<th>Bandwidth</th>
<th>TSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0,1,2</td>
<td>100%</td>
<td>ETS</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0%</td>
<td>SP</td>
</tr>
<tr>
<td>3</td>
<td>4,5,6,7</td>
<td>0%</td>
<td>SP</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Oper status is init
ETS DCBx Oper status is Down
State Machine Type is Asymmetric
Conf TLV Tx Status is enabled
Reco TLV Tx Status is enabled

0 Input Conf TLV Pkts, 1955 Output Conf TLV Pkts, 0 Error Conf TLV Pkts
0 Input Reco TLV Pkts, 1955 Output Reco TLV Pkts, 0 Error Reco TLV Pkts
FTOS(conf)# show interfaces tengigabitethernet 0/0 ets detail
Interface TenGigabitEthernet 0/0
Max Supported TC Groups is 4
Number of Traffic Classes is 8
Admin mode is on
Admin Parameters:
--------------------------------------------------------
Admin is enabled
TC-grp Priority# Bandwidth TSA
0 0,1,2,3,4,5,6,7 100% ETS
1 0% ETS
2 0% ETS
3 0% ETS
4 0% ETS
5 0% ETS
6 0% ETS
7 0% ETS
Priority# Bandwidth TSA
0 13% ETS
1 13% ETS
2 13% ETS
3 13% ETS
4 12% ETS
5 12% ETS
6 12% ETS
7 12% ETS
Remote Parameters:
-------------------
Remote is disabled
Local Parameters:
------------------
Local is enabled
TC-grp Priority# Bandwidth TSA
0 0,1,2,3,4,5,6,7 100% ETS
1 0% ETS
2 0% ETS
3 0% ETS
4 0% ETS
5 0% ETS
6 0% ETS
7 0% ETS
Priority# Bandwidth TSA
0 13% ETS
1 13% ETS
2 13% ETS
3 13% ETS
4 12% ETS
5 12% ETS
6 12% ETS
7 12% ETS
Oper status is init
Conf TLV Tx Status is disabled
Traffic Class TLV Tx Status is disabled
0 Input Conf TLV Pkts, 0 Output Conf TLV Pkts, 0 Error Conf TLV Pkts
0 TLV Input Traffic Class TLV Pkts, 0 Output Traffic Class TLV Pkts, 0 Error Traffic Class Pkts
Figure 13-13. show interface ets detail Command Example

FTOS(conf)# show interfaces tengigabitethernet 0/0 ets detail
Interface TenGigabitEthernet 0/0
Max Supported TC Groups is 4
Number of Traffic Classes is 8
Admin mode is on
Admin Parameters:
------------------
Admin is enabled
TC-grp  Priority#  Bandwidth  TSA
0  0,1,2,3,4,5,6,7  100%  ETS
1  0%  ETS
2  0%  ETS
3  0%  ETS
4  0%  ETS
5  0%  ETS
6  0%  ETS
7  0%  ETS

Remote Parameters:
-------------------
Remote is disabled
Local Parameters:
------------------
Local is enabled
TC-grp  Priority#  Bandwidth  TSA
0  0,1,2,3,4,5,6,7  100%  ETS
1  0%  ETS
2  0%  ETS
3  0%  ETS
4  0%  ETS
5  0%  ETS
6  0%  ETS
7  0%  ETS

Oper status is init
Conf TLV Tx Status is disabled
Traffic Class TLV Tx Status is disabled
0 Input Conf TLV Pkts, 0 Output Conf TLV Pkts, 0 Error Conf TLV Pkts
0 Input Traffic Class TLV Pkts, 0 Output Traffic Class TLV Pkts, 0 Error Traffic Class TLV Pkts
<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface</td>
<td>Interface type with stack-unit and port number.</td>
</tr>
<tr>
<td>Max Supported TC Group</td>
<td>Maximum number of priority groups supported.</td>
</tr>
<tr>
<td>Number of Traffic Classes</td>
<td>Number of 802.1p priorities currently configured.</td>
</tr>
<tr>
<td>Admin mode</td>
<td>ETS mode: on or off. When on, the scheduling and bandwidth allocation configured in an ETS output policy or received in a DCBx TLV from a peer can take effect on an interface.</td>
</tr>
<tr>
<td>Admin Parameters</td>
<td>ETS configuration on local port, including priority groups, assigned dot1p priorities, and bandwidth allocation.</td>
</tr>
<tr>
<td>Remote Parameters</td>
<td>ETS configuration on remote peer port, including Admin mode (enabled if a valid TLV was received or disabled), priority groups, assigned dot1p priorities, and bandwidth allocation. If the ETS Admin mode is enabled on the remote port for DCBx exchange, the Willing bit received in ETS TLVs from the remote peer is included.</td>
</tr>
<tr>
<td>Local Parameters</td>
<td>ETS configuration on local port, including Admin mode (enabled when a valid TLV is received from a peer), priority groups, assigned dot1p priorities, and bandwidth allocation.</td>
</tr>
<tr>
<td>Operational status (local port)</td>
<td>Port state for current operational ETS configuration: Init: Local ETS configuration parameters were exchanged with peer. Recommend: Remote ETS configuration parameters were received from peer. Internally propagated: ETS configuration parameters were received from configuration source.</td>
</tr>
<tr>
<td>ETS DCBx Oper status</td>
<td>Operational status of ETS configuration on local port: match or mismatch.</td>
</tr>
<tr>
<td>State Machine Type</td>
<td>Type of state machine used for DCBx exchanges of ETS parameters: Feature - for legacy DCBx versions; Asymmetric - for an IEEE version.</td>
</tr>
<tr>
<td>Conf TLV Tx Status</td>
<td>Status of ETS Configuration TLV advertisements: enabled or disabled.</td>
</tr>
<tr>
<td>ETS TLV Statistic: Input Conf TLV pkts</td>
<td>Number of ETS Configuration TLVs received.</td>
</tr>
<tr>
<td>ETS TLV Statistic: Output Conf TLV pkts</td>
<td>Number of ETS Configuration TLVs transmitted.</td>
</tr>
<tr>
<td>ETS TLV Statistic: Error Conf TLV pkts</td>
<td>Number of ETS Error Configuration TLVs received.</td>
</tr>
</tbody>
</table>
Figure 13-14. show stack-unit all stack-ports all pfc details Command Example

FTOS(conf)# show stack-unit all stack-ports all pfc details

stack unit 0 stack-port all
  Admin mode is On
  Admin is enabled, Priority list is 4-5
  Local is enabled, Priority list is 4-5
  Link Delay 45556 pause quantum
  0 Pause Tx pkts, 0 Pause Rx pkts

stack unit 1 stack-port all
  Admin mode is On
  Admin is enabled, Priority list is 4-5
  Local is enabled, Priority list is 4-5
  Link Delay 45556 pause quantum
  0 Pause Tx pkts, 0 Pause Rx pkts

Figure 13-15. show stack-unit all stack-ports all ets details Command Example

FTOS(conf)# show stack-unit all stack-ports all ets details

Stack unit 0 stack port all
Max Supported TC Groups is 4
Number of Traffic Classes is 1
Admin mode is on

Admin Parameters:
-------------------
Admin is enabled

<table>
<thead>
<tr>
<th>TC-grp</th>
<th>Priority#</th>
<th>Bandwidth</th>
<th>TSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0,1,2,3,4,5,6,7</td>
<td>100%</td>
<td>ETS</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Stack unit 1 stack port all
Max Supported TC Groups is 4
Number of Traffic Classes is 1
Admin mode is on

Admin Parameters:
-------------------
Admin is enabled

<table>
<thead>
<tr>
<th>TC-grp</th>
<th>Priority#</th>
<th>Bandwidth</th>
<th>TSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0,1,2,3,4,5,6,7</td>
<td>100%</td>
<td>ETS</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 13-16. show interface DCBx detail for ieee Command Example

FTOS(conf-if-te-0/17-lldp)#do sho int te 2/12 dc d

E-ETS Configuration TLV enabled     e-ETS Configuration TLV disabled
R-ETS Recommendation TLV enabled   r-ETS Recommendation TLV disabled
P-PFC Configuration TLV enabled     p-PFC Configuration TLV disabled
F-Application priority for FCOE enabled       f-Application Priority for FCOE disabled
I-Application priority for iSCSI enabled       i-Application Priority for iSCSI disabled

Interface TenGigabitEthernet 2/12
  Remote Mac Address 00:01:e8:8a:df:a0
  Port Role is Manual
  DCBx Operational Status is Enabled
  Is Configuration Source? FALSE
  Local DCBx Compatibility mode is IEEEv2.5
  Local DCBx Configured mode is IEEEv2.5
  Peer Operating version is IEEEv2.5
  Local DCBx TLVs Transmitted: ERPFi
  1 Input PFC TLV pkts, 2 Output PFC TLV pkts, 0 Error PFC pkts
  0 PFC Pause Tx pkts, 0 Pause Rx pkts
  1 Input ETS Conf TLV Pkts, 1 Output ETS Conf TLV Pkts, 0 Error ETS Conf TLV Pkts
  1 Input ETS Reco TLV pkts, 1 Output ETS Reco TLV pkts, 0 Error ETS Reco TLV Pkts
Figure 13-17. `show interface DCBx detail` for legacy cee Command Example

```bash
FTOS(conf-if-te-0/17-lldp)#do sho int te 1/14 dc d
```

```
E-ETS Configuration TLV enabled
e-ETS Configuration TLV disabled
R-ETS Recommendation TLV enabled
e-ETS Recommendation TLV disabled
P-PFC Configuration TLV enabled
e-PFC Configuration TLV disabled
F-Application priority for FCOE enabled
f-Application Priority for FCOE disabled
I-Application priority for iSCSI enabled
i-Application Priority for iSCSI disabled
```

```
Interface TenGigabitEthernet 1/14
Remote Mac Address 00:01:e8:8a:df:a0
Port Role is Auto-Upstream
DCBx Operational Status is Enabled
Is Configuration Source? FALSE
Local DCBx Compatibility mode is CEE
Local DCBx Configured mode is CEE
Peer Operating version is CEE
Local DCBx TLVs Transmitted: ErPFi

Local DCBx Status
------------------
DCBx Operational Version is 0
DCBx Max Version Supported is 0
Sequence Number: 1
Acknowledgment Number: 1
Protocol State: In-Sync

Peer DCBx Status:
------------------
DCBx Operational Version is 0
DCBx Max Version Supported is 0
Sequence Number: 1
Acknowledgment Number: 1
Total DCBx Frames transmitted 994
Total DCBx Frames received 646
Total DCBx Frame errors 0
Total DCBx Frames unrecognized 0
```

Table 13-5. `show interface DCBx detail` Command Description

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface</td>
<td>Interface type with chassis slot and port number.</td>
</tr>
<tr>
<td>Port-Role</td>
<td>Configured DCBx port role: auto-upstream, auto-downstream, config-source, or manual.</td>
</tr>
<tr>
<td>DCBx Operational Status</td>
<td>Operational status (enabled or disabled) used to elect a configuration source and internally propagate a DCB configuration. The DCBx operational status is the combination of PFC and ETS operational status.</td>
</tr>
<tr>
<td>Configuration Source</td>
<td>Specifies whether the port serves as the DCBx configuration source on the switch: true (yes) or false (no).</td>
</tr>
</tbody>
</table>
Table 13-5.  show interface DCBx detail Command Description

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local DCBx Compatibility mode</td>
<td>DCBx version accepted in a DCB configuration as compatible. In auto-upstream mode, a port can only received a DCBx version supported on the remote peer.</td>
</tr>
<tr>
<td>Local DCBx Configured mode</td>
<td>DCBx version configured on the port: CEE, CIN, IEEE v2.5, or Auto (port auto-configures to use the DCBx version received from a peer).</td>
</tr>
<tr>
<td>Peer Operating version</td>
<td>DCBx version that the peer uses to exchange DCB parameters.</td>
</tr>
<tr>
<td>Local DCBx TLVs Transmitted</td>
<td>Transmission status (enabled or disabled) of advertised DCB TLVs (see TLV code at the top of the show command output).</td>
</tr>
<tr>
<td>Local DCBx Status: DCBx Operational Version</td>
<td>DCBx version advertised in Control TLVs.</td>
</tr>
<tr>
<td>Local DCBx Status: DCBx Max Version Supported</td>
<td>Highest DCBx version supported in Control TLVs.</td>
</tr>
<tr>
<td>Local DCBx Status: Sequence Number</td>
<td>Sequence number transmitted in Control TLVs.</td>
</tr>
<tr>
<td>Local DCBx Status: Acknowledgment Number</td>
<td>Acknowledgement number transmitted in Control TLVs.</td>
</tr>
<tr>
<td>Local DCBx Status: Protocol State</td>
<td>Current operational state of DCBx protocol: ACK or IN-SYNC.</td>
</tr>
<tr>
<td>Peer DCBx Status: DCBx Operational Version</td>
<td>DCBx version advertised in Control TLVs received from peer device.</td>
</tr>
<tr>
<td>Peer DCBx Status: DCBx Max Version Supported</td>
<td>Highest DCBx version supported in Control TLVs received from peer device.</td>
</tr>
<tr>
<td>Peer DCBx Status: Sequence Number</td>
<td>Sequence number transmitted in Control TLVs received from peer device.</td>
</tr>
<tr>
<td>Peer DCBx Status: Acknowledgment Number</td>
<td>Acknowledgement number transmitted in Control TLVs received from peer device.</td>
</tr>
<tr>
<td>Total DCBx Frames transmitted</td>
<td>Number of DCBx frames sent from local port.</td>
</tr>
<tr>
<td>Total DCBx Frames received</td>
<td>Number of DCBx frames received from remote peer port.</td>
</tr>
<tr>
<td>Total DCBx Frame errors</td>
<td>Number of DCBx frames with errors received.</td>
</tr>
<tr>
<td>Total DCBx Frames unrecognized</td>
<td>Number of unrecognizable DCBx frames received.</td>
</tr>
</tbody>
</table>
PFC and ETS Configuration Examples

This section contains examples of how to configure and apply DCB input and output policies on an interface.

Using PFC and ETS to Manage Data Center Traffic

In the following example:

- Incoming SAN traffic is configured for priority-based flow control.
- Outbound LAN, IPC, and SAN traffic is mapped into three ETS priority groups and configured for enhanced traffic selection (bandwidth allocation and scheduling).
- One lossless queue is used.
QoS Traffic Classification: The service-class dynamic dot1p command has been used in Global Configuration mode to map ingress dot1p frames to the queues shown in Table 13-6. For more information, refer to QoS dot1p Traffic Classification and Queue Assignment.
Lossless SAN traffic with dot1p priority 3 is assigned to queue 1. Other traffic types are assigned the 802.1p priorities shown in Table 13-7 and the bandwidth allocations shown in Table 13-8.

### Table 13-6. Example: dot1p-Queue Assignment

<table>
<thead>
<tr>
<th>dot1p Value in Incoming Frame</th>
<th>Queue Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

### Table 13-7. Example: dot1p-priority class group Assignment

<table>
<thead>
<tr>
<th>dot1p Value in Incoming Frame</th>
<th>Priority Group Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>LAN</td>
</tr>
<tr>
<td>1</td>
<td>LAN</td>
</tr>
<tr>
<td>2</td>
<td>LAN</td>
</tr>
<tr>
<td>3</td>
<td>SAN</td>
</tr>
<tr>
<td>4</td>
<td>IPC</td>
</tr>
<tr>
<td>5</td>
<td>LAN</td>
</tr>
<tr>
<td>6</td>
<td>LAN</td>
</tr>
<tr>
<td>7</td>
<td>LAN</td>
</tr>
</tbody>
</table>

### Table 13-8. Example: priority group-bandwidth Assignment

<table>
<thead>
<tr>
<th>Priority Group</th>
<th>Bandwidth Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPC</td>
<td>5%</td>
</tr>
<tr>
<td>SAN</td>
<td>50%</td>
</tr>
<tr>
<td>LAN</td>
<td>45%</td>
</tr>
</tbody>
</table>
Figure 13-19. PFC and ETS Configuration Command Example

Configure QoS priority-queue assignment to honor dot1p priorities or use L2 class maps to mark and map ingress traffic to output queues; for example:

FTOS(conf)# service-class dynamic dot1p

Or

FTOS(conf)# interface tengigabitethernet 0/1
FTOS(conf-if-te-0/1)# service-class dynamic dot1p

Configure a DCB input policy for applying PFC to lossless SAN priority traffic:
FTOS(conf)# dcb-input ipc_san_lan
FTOS(conf-qos-policy-in)# pfc mode on
FTOS(conf-qos-policy-in)# pfc priority 3

Configure an ETS priority group:
FTOS(conf)# priority-group san
FTOS(conf-pg)# priority-list 3
FTOS(conf-pg)# set-pgid 1
FTOS(conf-pg)# exit
FTOS(conf)# priority-group ipc
FTOS(conf-pg)# priority-list 4
FTOS(conf-pg)# set-pgid 2
FTOS(conf-pg)# exit
FTOS(conf)# priority-group lan
FTOS(conf-pg)# priority-list 0-2,5-7
FTOS(conf-pg)# set-pgid 3
FTOS(conf-pg)# exit

Configure an ETS output policy for egress traffic:
FTOS(conf)# qos-policy-output san ets
FTOS(conf-qos-policy-out)# bandwidth-percentage 50
FTOS(conf-qos-policy-out)# exit
FTOS(conf)# qos-policy-output lan ets
FTOS(conf-qos-policy-out)# bandwidth-percentage 45
FTOS(conf-qos-policy-out)# exit
FTOS(conf)# qos-policy-output ipc ets
FTOS(conf-qos-policy-out)# bandwidth-percentage 5
FTOS(conf-qos-policy-out)# exit
Using PFC and ETS to Manage Converged Ethernet Traffic in a Switch Stack

Figure 13-21 shows how to apply the DCB PFC input policy (ipc_san_lan) and ETS output policy (ets) configured in Figure 13-19 and Figure 13-20 on all ports in a switch stack.

Figure 13-21. Apply DCB PFC Input Policy and ETS Output Policy in a Switch Stack Example

On the stack master, apply DCB PFC input and ETS output policies to all port interfaces on stacked switches:

```
FTOS(conf)# dcb-policy output stack-unit all stack-ports all ets
FTOS(conf)# dcb-policy input stack-unit all stack-ports all pfc
```
Hierarchical Scheduling in ETS Output Policies

ETS supports up to three levels of hierarchical scheduling. For example, you can apply ETS output policies with the following configurations:

- Priority group 1 assigns traffic to one priority queue with 20% of the link bandwidth and strict-priority scheduling.
- Priority group 2 assigns traffic to one priority queue with 30% of the link bandwidth.
- Priority group 3 assigns traffic to two priority queues with 50% of the link bandwidth and strict-priority scheduling.

In this example, the configured ETS bandwidth allocation and scheduler behavior is as follows:

- Unused bandwidth usage: Normally, if there is no traffic or unused bandwidth for a priority group, the bandwidth allocated to the group is distributed to the other priority groups according to the bandwidth percentage allocated to each group. However, when three priority groups with different bandwidth allocations are used on an interface:
  - If priority group 3 has free bandwidth, it is distributed as follows: 20% of the free bandwidth to priority group 1 and 30% of the free bandwidth to priority group 2.
  - If priority group 1 or 2 has free bandwidth, (20 + 30)% of the free bandwidth is distributed to priority group 3. Priority groups 1 and 2 retain whatever free bandwidth remains up to the (20 + 30)%.

- Strict-priority groups: If two priority groups have strict-priority scheduling, traffic assigned from the priority group with the higher priority-queue number is scheduled first. However, when three priority groups are used and two groups have strict-priority scheduling (such as groups 1 and 3 in the example), the strict priority group whose traffic is mapped to one queue takes precedence over the strict priority group whose traffic is mapped to two queues.

Therefore, in the example, scheduling traffic to priority group 1 (mapped to one strict-priority queue) takes precedence over scheduling traffic to priority group 3 (mapped to two strict-priority queues).
S-Series Debugging and Diagnostics

The chapter contains the following major sections:

- Offline diagnostics
- Trace logs
- Last restart reason
- show hardware commands
- Environmental monitoring
- Buffer tuning
- Troubleshooting packet loss
- Application core dumps
- Mini core dumps
- TCP dumps

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) 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 14-1. 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 14-1. 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 14-2.
3. Start diagnostics on the unit using the command `diag`, as shown in Figure 14-3. 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 14-3 and Figure 14-4, log messages differ somewhat when diagnostics are done on a standalone unit and on a stack member.
Figure 14-3. 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 drwx 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 14-4 shows the output of the master and member units when you run offline diagnostics on a member unit.

Figure 14-4. 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 14-5.

**Figure 14-5. Viewing the Results of Offline Diagnostics on a Standalone Unit**

```plaintext
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 - 3.30000 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: 759003660
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.

Last restart reason

If an S4810 system restarted for some reason (automatically or manually), the show system command output includes the reason for the restart. The following table shows the reasons displayed in the output and their corresponding causes.

Table 14-1. Line card restart causes and reasons

<table>
<thead>
<tr>
<th>Causes</th>
<th>Displayed Reasons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remote power cycle of the chassis</td>
<td>push button reset</td>
</tr>
<tr>
<td>reload</td>
<td>soft reset</td>
</tr>
<tr>
<td>reboot after a crash</td>
<td>soft reset</td>
</tr>
</tbody>
</table>

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.

Table 14-2. Hardware Watchdog Command

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hardware watchdog</td>
<td>Enable the hardware watchdog mechanism.</td>
</tr>
</tbody>
</table>
show hardware commands

**Note:** The show hardware command tree is supported on the S4810 and Z9000.

The show hardware command tree consists of EXEC Privilege commands used with the S4810 system. These commands display information from a hardware sub-component and from hardware-based feature tables.

Table 14-3 lists the show hardware commands available as of the latest FTOS version on the S4810.

**Note:** The show hardware commands should only be used under the guidance of Dell Force10 Technical Assistance Center.

### Table 14-3. show hardware Commands

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>show hardware stack-unit {0-11} cpu management statistics</td>
<td>View internal interface status of the stack-unit CPU port which connects to the external management interface.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} cpu data-plane statistics</td>
<td>View driver-level statistics for the data-plane port on the CPU for the specified stack-unit. It provides insight into the packet types entering the CPU to see whether CPU-bound traffic is internal (IPC traffic) or network control traffic, which the CPU must process.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} buffer total-buffer</td>
<td>View the modular packet buffers details per stack unit and the mode of allocation.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} buffer unit {0-1} total-buffer</td>
<td>View the modular packet buffers details per unit and the mode of allocation.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} buffer unit {0-1} port {1-64</td>
<td>all} buffer-info</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} buffer unit {0-1} port {1-64} queue {0-14</td>
<td>all} buffer-info</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} cpu party.bus statistics</td>
<td>View input and output statistics on the party bus, which carries inter-process communication traffic between CPUs.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} drops unit {0-1} port {1-64}</td>
<td>View the ingress and egress internal packet-drop counters, MAC counters drop, and FP packet drops for the stack unit on per port basis. It assists in identifying the stack unit/port pipe/port that may experience internal drops.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} stack-port {0-64}</td>
<td>View the input and output statistics for a stack-port interface.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} unit {0-1} counters</td>
<td>View the counters in the field processors of the stack unit.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} unit {0-1} details</td>
<td>View the details of the FP Devices, and Hi gig ports on the stack-unit.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} unit {0-1} execute-shell-cmd {command}</td>
<td>Execute a specified bShell commands from the CLI without going into the bShell.</td>
</tr>
</tbody>
</table>
Environmental monitoring

The S4810 components use environmental monitoring hardware to detect transmit power readings, receive power readings, and temperature updates. To receive periodic power updates, the `enable optic-info-update interval` command must be enabled. The output in **Figure 14-6** displays the environment status of the RPM.

**Figure 14-6. show interfaces transceiver Command Example**

```plaintext
FTOS#show interfaces

-- RPM Environment Status --
Slot  Status        Temp   Voltage
--------------------------------------
  0  active        33C    ok
  1  not present
```

Recognize an over-temperature condition

An over-temperature 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 3**.

**Message 3 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

---

**Table 14-3. show hardware Commands**

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>show hardware stack-unit {0-11} unit {0-1} ipmc-replication</td>
<td>View the Multicast IPMC replication table from the bShell.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} unit {0-1} port-stats [detail]</td>
<td>View the internal statistics for each port-pipe (unit) on per port basis.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} unit {0-1} register</td>
<td>View the stack-unit internal registers for each port-pipe.</td>
</tr>
<tr>
<td>show hardware stack-unit {0-11} unit {0-1} table-dump {table name}</td>
<td>View the tables from the bShell through the CLI without going into the bShell.</td>
</tr>
</tbody>
</table>
To view the programmed alarm thresholds levels, including the shutdown value, execute the `show alarms threshold` command shown in Figure 14-7.

**Figure 14-7.  show alarms threshold Command Example**

```
FTOS#show alarms threshold
-- Temperature Limits (deg C) --
------------------------------------
Linecard    Minor  Minor Off  Major  Major Off  Shutdown
------------------------------------
RPM        65     60       75       70       80
FTOS#
```

**Troubleshoot an over-temperature 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 4**.

**Message 4  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 4** 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.

The SNMP traps and OIDs in Table 14-4 provide information on S-Series environmental monitoring hardware and hardware components.

Table 14-4. SNMP Traps and OIDs

<table>
<thead>
<tr>
<th>OID String</th>
<th>OID Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receiving power</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.1.3.6.1.4.1.6027.3.10.1.2.5.1.6</td>
<td>chSysPortXfpRecvPower</td>
<td>OID to display the receiving power of the connected optics.</td>
</tr>
<tr>
<td>Transmitting power</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.1.3.6.1.4.1.6027.3.10.1.2.5.1.8</td>
<td>chSysPortXfpTxPower</td>
<td>OID to display the transmitting power of the connected optics.</td>
</tr>
<tr>
<td>Temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.1.3.6.1.4.1.6027.3.10.1.2.5.1.7</td>
<td>chSysPortXfpRecvTemp</td>
<td>OID to display the Temperature of the connected optics.</td>
</tr>
</tbody>
</table>

Note: These OIDs will only be generated if the CLI command enable optic-info-update-interval is enabled.

Hardware MIB Buffer Statistics

<table>
<thead>
<tr>
<th>OID String</th>
<th>OID Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>.1.3.6.1.4.1.6027.3.16.1.1.4</td>
<td>fpPacketBufferTable</td>
<td>View the modular packet buffers details per stack unit and the mode of allocation.</td>
</tr>
<tr>
<td>.1.3.6.1.4.1.6027.3.16.1.1.5</td>
<td>fpStatsPerPortTable</td>
<td>View the forwarding plane statistics containing the packet buffer usage per port per stack unit.</td>
</tr>
<tr>
<td>.1.3.6.1.4.1.6027.3.16.1.1.6</td>
<td>fpStatsPerCOSTable</td>
<td>View the forwarding plane statistics containing the packet buffer statistics per COS per port.</td>
</tr>
</tbody>
</table>

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 14-5 describes the type and number of ASICs per platform.

**Table 14-5. ASICS by Platform**

<table>
<thead>
<tr>
<th>Hardware</th>
<th>FP</th>
<th>CSF</th>
</tr>
</thead>
<tbody>
<tr>
<td>S50N, S50V</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>S25V, S25P, S25N</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

You can tune buffers at three locations, as shown in Figure 14-8.

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
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

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Define a buffer profile for the FP queues.</td>
<td>buffer-profile fp fsqueue</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Define a buffer profile for the CSF queues.</td>
<td>buffer-profile csf csqueue</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Change the dedicated buffers on a physical 1G interface.</td>
<td>buffer dedicated</td>
<td>BUFFER PROFILE</td>
</tr>
<tr>
<td>Change the maximum amount of dynamic buffers an interface can request.</td>
<td>buffer dynamic</td>
<td>BUFFER PROFILE</td>
</tr>
<tr>
<td>Change the number of packet-pointers per queue.</td>
<td>buffer packet-pointers</td>
<td>BUFFER PROFILE</td>
</tr>
<tr>
<td>Apply the buffer profile to a line card.</td>
<td>buffer fp-uplink linecard</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Apply the buffer profile to a CSF to FP link.</td>
<td>buffer csf linecard</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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.

```
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
```

Table 14-6. Buffer Allocation Error

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 14-10. Display the default buffer profile using the command show buffer-profile (summary | detail) from EXEC Privilege mode, as shown in Figure 14-9.
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

FTOS#show running-config interface tengigabitethernet 2/0
interface TenGigabitEthernet 2/0
no ip address
mtu 9252
switchport
no shutdown
buffer-policy myfsbufferprofile

FTOS#show 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#show 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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apply one of two pre-defined buffer profiles for all port pipes in the system.</td>
<td>buffer-profile global [1Q</td>
<td>4Q]</td>
</tr>
</tbody>
</table>

You must reload the system for the global buffer profile to take effect (Message 5).

**Message 5  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 5 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 6  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 7  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 14-11.
Troubleshooting packet loss

The `show hardware stack-unit` command is intended primarily to troubleshoot packet loss.

- `show hardware stack-unit cpu data-plane statistics`
- `show hardware stack-unit cpu party-bus statistics`
- `show hardware stack-unit 0-11 drops unit 0-1 port 0-63`
- `show hardware stack-unit 0-11 stack-port 48-51`
- `show hardware stack-unit 0-11 drops unit 0-1 {counters | details | port-stats [detail] | register | execute-shell-cmd | ipmc-replication | table-dump}:
- `show hardware layer2|layer3 {e.g. acl |in acl} stack-unit 0-11 port-set 0-1`
- `show hardware layer3 qos stack-unit 0-7 port-set 0-1`
- `show hardware ipv6 {e.g.-acl |in-acl} stack-unit 0-11 port-set 0-1`
- `show hardware system-flow layer2 stack-unit 0-11 port-set 0-1 [counters]`
- `clear hardware stack-unit 0-11 counters`
- `clear hardware stack-unit 0-11 unit 0-1 counters`
- `clear hardware stack-unit 0-11 cpu data-plane statistics`
- `clear hardware stack-unit 0-11 cpu party-bus statistics`
- `clear hardware stack-unit 0-11 stack-port 48-51`

Displaying Drop Counters

The `show hardware stack-unit 0–11 drops [unit 0 [port 0–63]]` command assists in identifying which stack unit, port pipe, and port is experiencing internal drops, as shown in Figure 14-12 and Figure 14-13.
Display drop counters with the `show hardware stack-unit drops unit port` command:

```
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
```
Table 14-13. Displaying Buffer Statistics, 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 14-14, 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 14-14. Displaying Buffer Statistics, 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
recv :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

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 14-15.

Figure 14-15. 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 14-16.

Figure 14-16. 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 0–11 {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 14-17.

Figure 14-17. Displaying Stack Unit Counters

```
RIPC4.ge0 :     1,202                                  +1,202
RUC.ge0  :     1,224                                  +1,217
RDBG0.ge0 :     34                                    +24
RDBG1.ge0 :     366                                   +235
RDBG5.ge0 :     16                                    +12
RDBG7.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
TDBG06.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:

Enable RPM core dumps and specify the shutdown mode.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable RPM core dumps and specify the shutdown mode.</td>
<td>logging coredump server</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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.

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 14-18 and sample file text is shown in Figure 14-19.

Figure 14-18.  Mini application core file naming example

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.
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.

TCP dumps

TCP dump captures CPU bound control plane traffic to improve troubleshooting and system manageability. When enabled, a TCP dump captures all the packets on the local CPU, as specified in the CLI.

The traffic capture files can be saved to flash, to FTP, SCP, or TFTP. The files saved on the flash are located in the `flash:///TCP_DUMP_DIR/Tcpdump_<time_stamp_dir>/` directory, and labeled `tcpdump_*.*pcap`. There can be up to 20 `Tcpdump_<time_stamp_dir>` directories. The file after 20 overwrites the oldest saved file. The maximum file size for a TCP dump capture is 1MB. When a file reaches 1MB, a new file is created, up to the specified total number of files.

Maximize the number of packets recorded in a file by specifying the snap-length to capture the file headers only.

The `tcpdump` CLI has a finite run process. When the command is enabled, it will run until the capture-duration timer and/or the packet-count counter threshold is met. If no threshold is set, the system uses a default of 5 minute capture-duration and/or a single 1k file as the stopping point for the dump.

The capture-duration timer and the packet-count counter can be used at the same time. The TCP dump stops when the first of the thresholds is met. That means that even if the duration timer is 9000 seconds, if the maximum file count parameter is met first, the dumps stop.

---

**Figure 14-19. 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
```
<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable a TCP dump for CPU bound traffic.</td>
<td>tcpdump cp [capture-duration <em>time</em></td>
<td>filter <em>expression</em></td>
</tr>
</tbody>
</table>
Dynamic Host Configuration Protocol (DHCP)

Dynamic Host Configuration Protocol (DHCP) is available on platforms: **E C S Z S4810**

This chapter contains the following sections:

- Protocol Overview
- Implementation Information
- Configuration Tasks
- Configure the System to be a DHCP Server
- Configure the System to be a Relay Agent
- Configure the System for User Port Stacking
- Configure Secure DHCP

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 on 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 dhcp snoopingtransmits 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 of 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 15-1.

Figure 15-1. DHCP Packet Format

![DHCP Packet Format](image-url)

Table 15-1. Common DHCP Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subnet Mask</td>
<td>1</td>
<td>Specifies the client’s subnet mask.</td>
</tr>
<tr>
<td>Router</td>
<td>3</td>
<td>Specifies the router IP addresses that may serve as the client’s default gateway.</td>
</tr>
<tr>
<td>Domain Name Server</td>
<td>6</td>
<td>Specifies the DNS servers that are available to the client.</td>
</tr>
<tr>
<td>Domain Name</td>
<td>15</td>
<td>Specifies the domain name that clients should use when resolving hostnames via DNS.</td>
</tr>
<tr>
<td>IP Address Lease Time</td>
<td>51</td>
<td>Specifies the amount of time that the client is allowed to use an assigned IP address.</td>
</tr>
</tbody>
</table>
| 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 octets 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 the illustration below.

- **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.
Implementation Information

- The Dell Force10 implementation of DHCP is based on RFC 2131 and RFC 3046.
- 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 on 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. |

**Note:** If DHCP Snooping is enabled globally and any L2 port is configured, any IP ACL,MAC ACL, or DHCP Source-Address validation ACL won't block DHCP packets.

- 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. For example, if all pools were configured for a /24 mask, the total would be 40000/253 (approximately 158). If the subnet is increased, more pools can be configured. The maximum subnet that can be configured for a single pool is /17. FTOS displays an error message for configurations that exceed the allocated memory.
- E-Series supports 16K DHCP Snooping entries across 500 VLANs.
- C-Series, S-Series (S25/S50), S55, S60 and S4810 support 4K DHCP Snooping entries.
- All platforms support Dynamic ARP Inspection on 16 VLANs per system. Refer to Dynamic ARP Inspection.

**Note:** If the DHCP server is located on the ToR and the VLTi (ICL) is down due to a failed link when a VLT node is rebooted in JumpStart mode, it will not be able to reach the DHCP server, resulting in BMP failure.

Configuration Tasks

- Configure the System to be a DHCP Server
- Configure the System to be a Relay Agent
- Configure Secure DHCP

Configure the System to be a DHCP Server

*Configure the System to be a DHCP Server* is supported only on platforms: C S S4810
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
- Create Manual Binding Entries
- Debug DHCP server
- DHCP Clear Commands

**Configure the Server for Automatic Address Allocation**

This feature is available on C and S (S25/S50) platforms only.

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Access the DHCP server CLI context.</td>
<td>ip dhcp server</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Create an address pool and give it a name.</td>
<td>pool name</td>
<td>DHCP</td>
</tr>
<tr>
<td>3</td>
<td>Specify the range of IP addresses from which the DHCP server may assign addresses.</td>
<td>network network /prefix-length</td>
<td>DHCP &lt;POOL&gt;</td>
</tr>
<tr>
<td></td>
<td>• network is the subnet address.</td>
<td>Prefix-length Range: 17 to 31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• prefix-length specifies the number of bits used for the network portion of the address you specify.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Display the current pool configuration.</td>
<td>show config</td>
<td>DHCP &lt;POOL&gt;</td>
</tr>
</tbody>
</table>

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.

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclude an address range from DHCP assignment. The exclusion applies to all configured pools.</td>
<td>excluded-address</td>
<td>DHCP</td>
</tr>
</tbody>
</table>
Specify an Address Lease Time

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify an address lease time for the addresses in a pool.</td>
<td>lease {days [hours] [minutes]</td>
<td>infinite}</td>
</tr>
<tr>
<td>Default: 24 hours</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Specify a Default Gateway

The IP address of the default router should be on the same subnet as the client.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify default gateway(s) for the clients on the subnet, in order of preference.</td>
<td>default-router address</td>
<td>DHCP &lt;POOL&gt;</td>
</tr>
</tbody>
</table>

Enable DHCP Server

This feature is available on C and S (S25/S50) platforms only.

The DHCP server is disabled by default.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter the DHCP command-line context.</td>
<td>ip dhcp server</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Enable DHCP server.</td>
<td>no disable</td>
<td>DHCP</td>
</tr>
<tr>
<td></td>
<td>Default: Disabled</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Display the current DHCP configuration.</td>
<td>show config</td>
<td>DHCP</td>
</tr>
</tbody>
</table>

In the illustration below, 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.
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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create a domain</td>
<td>domain-name name</td>
<td>DHCP &lt;POOL&gt;</td>
</tr>
<tr>
<td>2</td>
<td>Specify in order of preference the DNS servers that are available</td>
<td>dns-server address</td>
<td>DHCP &lt;POOL&gt;</td>
</tr>
<tr>
<td></td>
<td>to a DHCP client</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Specify the NetBIOS Windows Internet Naming Service (WINS) name</td>
<td>netbios-name-server</td>
<td>DHCP &lt;POOL&gt;</td>
</tr>
<tr>
<td></td>
<td>servers, in order of preference, that are available to Microsoft</td>
<td>address</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dynamic Host Configuration Protocol (DHCP) clients.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Specify the NetBIOS node type for a Microsoft DHCP client. Dell</td>
<td>netbios-node-type</td>
<td>DHCP &lt;POOL&gt;</td>
</tr>
<tr>
<td></td>
<td>Force10 recommends specifying clients as hybrid.</td>
<td>type</td>
<td></td>
</tr>
</tbody>
</table>

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 assigns the client an available IP address automatically, and then creates an 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 a 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create an address pool</td>
<td>pool name</td>
<td>DHCP</td>
</tr>
<tr>
<td>2</td>
<td>Specify the client IP address.</td>
<td>host address</td>
<td>DHCP &lt;POOL&gt;</td>
</tr>
<tr>
<td>3</td>
<td>Specify the client hardware address.</td>
<td>hardware-address hardware-address type</td>
<td>DHCP &lt;POOL&gt;</td>
</tr>
</tbody>
</table>

- hardware-address is the client MAC address.
- type is the protocol of the hardware platform. The default protocol is Ethernet.

### Debug DHCP server

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display debug information for DHCP server.</td>
<td>debug ip dhcp server [events</td>
<td>packets]</td>
</tr>
</tbody>
</table>

### DHCP Clear Commands

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear DHCP binding entries for the entire binding table.</td>
<td>clear ip dhcp binding</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Clear a DHCP binding entry for an individual IP address.</td>
<td>clear ip dhcp binding ip address</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Clear a DHCP address conflict. (This feature is available on C and S (S25/S50 platforms only.)</td>
<td>clear ip dhcp conflict</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Clear DHCP server counters. (This feature is available on C and S (S25/S50 platforms only.)</td>
<td>clear ip dhcp server statistics</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

### Configure the System to be a Relay Agent

The following feature is available on platforms: C, E, S and S4810

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 the illustration below. 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 and VLANs.

To view the `ip helper-address` configuration for an interface, use the command `show ip interface` from EXEC privilege mode, as shown in the following example.

```
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 unreachable is not sent
```

```
R1(conf-if-gi-1/3)#show config
!
interface GigabitEthernet 1/3
 ip address 10.11.0.3/24
 ip helper-address 10.11.1.5
 ip helper-address 10.11.2.5
 no shutdown
```
Configure the System for User Port Stacking

When you set the DHCP offer on the DHCP server, you can set the stacking-option variable to provide the stack-port detail so a stack can be formed when the units are connected.

Configure Secure DHCP

The following feature is available on platforms: C E S 54810 Z (except where noted).

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
- DHCP Snooping
- Dynamic ARP Inspection
- Source Address Validation

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insert Option 82 into DHCP packets. For routers between the relay agent and the DHCP server, enter the trust-downstream option.</td>
<td>ip dhcp relay information-option</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td></td>
<td>[trust-downstream]</td>
<td></td>
</tr>
<tr>
<td>Manually reset the remote ID for Option 82.</td>
<td>ip dhcp relay information-option</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td></td>
<td>remote-id</td>
<td></td>
</tr>
</tbody>
</table>

**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 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 (DHCOFFER, DHCPACK, DHCPNACK) that arrive on an untrusted port are also dropped. This check-point prevents an attacker from imposter 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.

**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 DHCP Snooping

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enable DHCP Snooping globally.</td>
<td>ip dhcp snooping</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Specify ports connected to DHCP servers as trusted.</td>
<td>ip dhcp snooping trust</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>3</td>
<td>Enable DHCP Snooping on a VLAN.</td>
<td>ip dhcp snooping vlan</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

Add a static entry in the binding table

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Add a static entry in the binding table.</td>
<td>ip dhcp snooping binding mac</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

Clear the binding table

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delete all of the entries in the binding table</td>
<td>clear ip dhcp snooping binding</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

Display the contents of the binding table

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display the contents of the binding table.</td>
<td>show ip dhcp snooping</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>
View the DHCP Snooping statistics with the `show ip dhcp snooping` command.

```
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.

```
FTOS#show ip dhcp snooping binding
Codes :  S - Static  D - Dynamic

<table>
<thead>
<tr>
<th>IP Address</th>
<th>MAC Address</th>
<th>Expires (Sec)</th>
<th>Type</th>
<th>VLAN</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.1.1.251</td>
<td>00:00:4d:57:52:58</td>
<td>172800</td>
<td>D</td>
<td>Vl 10</td>
<td>Gi 0/2</td>
</tr>
<tr>
<td>10.1.1.252</td>
<td>00:00:4d:57:66:76</td>
<td>172800</td>
<td>D</td>
<td>Vl 10</td>
<td>Gi 0/1</td>
</tr>
<tr>
<td>10.1.1.253</td>
<td>00:00:4d:57:88:e8</td>
<td>172740</td>
<td>D</td>
<td>Vl 10</td>
<td>Gi 0/3</td>
</tr>
<tr>
<td>10.1.1.254</td>
<td>00:00:4d:69:e8:f2</td>
<td>172740</td>
<td>D</td>
<td>Vl 10</td>
<td>Te 0/50</td>
</tr>
</tbody>
</table>
```

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 accept ARP requests 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 the 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 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 L2SystemFlow CAM region, a sub-region of SystemFlow. One CAM entry is required for every DAI-enabled VLAN. You can enable DAI on up to 16 VLANs on a system. However, the ExaScale default CAM profile allocates only 9 entries to the L2SystemFlow region for DAI. You can configure 10 to 16 DAI-enabled VLANs by allocating more CAM space to the L2SystemFlow 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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enable DHCP Snooping.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Validate ARP frames against the DHCP Snooping binding table.</td>
<td>arp inspection</td>
<td>INTERFACE VLAN</td>
</tr>
</tbody>
</table>

View the number of entries in the ARP database with the `show arp inspection database` command.

```
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     V1 10  CP
Internet     10.1.1.252  -         00:00:4d:57:e6:f6  Gi 0/1     V1 10  CP
Internet     10.1.1.253  -         00:00:4d:57:f8:e8  Gi 0/3     V1 10  CP
Internet     10.1.1.254  -         00:00:4d:69:e8:f2  Te 0/50     V1 10  CP
FTOS#
```
Use **show arp inspection statistics** command to see how many valid and invalid ARP packets have been processed.

```
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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify an interface as trusted so that ARPs are not validated against the binding table.</td>
<td><code>arp inspection-trust</code></td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

**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** prevents IP spoofing by forwarding only IP packets that have been validated against the DHCP binding table.
- **DHCP MAC Source Address Validation** verifies a DHCP packet’s source hardware address matches the client hardware address field (CHADDR) in the payload.
- **IP+MAC Source Address Validation** 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.

### 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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable DHCP MAC Source Address Validation</td>
<td>ip dhcp snooping verify mac-address</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**Note:** If IP Source Guard is enabled using the `ip dhcp source-address-validation` command and there are 187 entries or more in the current DHCP snooping binding table, Source Address Validation (SAV) may not be applied to all entries.

To ensure that SAV is applied correctly to all entries, enable the `ip dhcp source-address-validation` command before adding entries to the binding table.

### IP+MAC Source Address Validation

The following feature is available on platforms: C, S and 5410.

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Allocate at least one FP block to the ipmacacl CAM region.</td>
<td>cam-acl l2acl</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Save the running-config to the startup-config.</td>
<td>copy running-config startup-config</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>3</td>
<td>Reload the system.</td>
<td>reload</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>
FTOS creates an ACL entry for each IP+MAC address pair in the binding table and applies it to the interface.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Enable IP+MAC Source Address Validation.</td>
<td>ip dhcp source-address-validation ipmac</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display the IP+MAC ACL for an interface for the entire system.</td>
<td>show ip dhcp snooping source-address-validation [interface]</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>
Equal Cost Multi-Path (ECMP)

Equal Cost Multi-Path (ECMP) is supported on platforms: E C S S4810

ECMP for Flow-based Affinity

ECMP for Flow-based Affinity is available on platforms E and S4810

The hashing algorithm on E-Series TeraScale and E-Series ExaScale are different. Hashing on ExaScale is based on CRC, checksum, or XOR, and the algorithm on TeraScale 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.

Note: IPv6 /128 routes having multiple paths do not form ECMPs. The /128 route is treated as a host entry and finds its place in the host table.

Note: Using XOR algorithms will result in imbalanced loads across an ECMP/LAG when the number of members in said ECMP/LAG is a multiple of 4.

- Configurable Hash Algorithm
- Configurable Hash Algorithm Seed
- Deterministic ECMP Next Hop
- Link Bundle Monitoring

Configurable Hash Algorithm

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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change the ExaScale hash-algorithm for LAG, ECMP, and NH-ECMP to match TeraScale.</td>
<td><code>hash-algorithm ecmp checksum 0</code> \ <code>lag checksum 0 nh-ecmp checksum 0</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
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.

### 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.

### 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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable IPv4 Deterministic ECMP Next Hop.</td>
<td><code>ip ecmp-deterministic</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Enable IPv6 Deterministic ECMP Next Hop.</td>
<td><code>ipv6 ecmp-deterministic</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**Note:** Packet loss might occur when you enable `ip/ipv6 ecmp-deterministic` for the first-time only.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify the hash algorithm seed.</td>
<td><code>hash-algorithm seed value [linecard number] [port-set number]</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Range: 0 to 4095.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In the illustration below, 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.

**Link Bundle Monitoring**

**Link Bundle Monitoring** is supported only on platform S4810

Monitoring linked ECMP bundles allows traffic distribution amounts in a link to be monitored for unfair distribution at any given time. A default threshold of 60% is defined as an acceptable amount of traffic on a member link. Links are monitored in 15-second intervals for three consecutive instances. Any deviation within that time causes a syslog to be sent and an alarm event to be generated. When the deviation clears, another syslog is sent and a clear alarm event is generated.

**Message 1** Link bundle monitoring percent threshold

```
%STKUNIT0-M:CP %IFMGR-5-BUNDLE_UNEVEN_DISTRIBUTION: Found uneven distribution in LAG bundle 11.
```

The link bundle utilization is calculated as the total bandwidth of all links divided by the total bytes-per-second of all links. Within each ECMP group, interfaces can be specified. If monitoring is enabled for the ECMP group, the utilization calculation is performed when the utilization of the link-bundle (not a link within a bundle) exceeds 60%.
Enable link bundle monitoring using the `ecmp-group` command.

**Note:** An ecmp-group index is generated automatically for each unique ecmp-group when the user configures multipath routes to the same network. The system can generate a maximum of 512 unique ecmp-groups. The ecmp-group indexes are generated in even numbers (0, 2, 4, 6... 1022) and are for information only.

For link bundle monitoring with ECMP, the `ecmp-group` command is used to enable the link bundle monitoring feature. The ecmp-group with `id 2`, enabled for link bundle monitoring is *user configured*. This is different from the ecmp-group `index 2` that is created by configuring routes and is *automatically generated*.

These two ecmp-groups are not related in any way.

### Managing ECMP Group Paths

Managing ECMP Group Paths is supported only on platform: S4810

Configure the maximum number of paths for an ECMP route that the L3 CAM can hold to avoid path degeneration. When the maximum number of routes is not configured, the CAM can hold a maximum ECMP per route.

Use the `ip ecmp-group path-fallback` command to enable or disable the feature.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure the maximum number of paths per ECMP group</td>
<td><code>ip ecmp-group maximum-paths {2-64}</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Enable ECMP group path management</td>
<td><code>ip ecmp-group path-fallback</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**Note:** You must save the new ECMP settings to the startup-config (`write-mem`) then reload the system for the new settings to take effect.

```bash
pt-s4801-temp1(conf)#ip ecmp-group maximum-paths 3
User configuration has been changed. Save the configuration and reload to take effect
pt-s4801-temp1(conf)#
```
Enabling FIPS Cryptography

FIPS Cryptography is supported on the following platforms: S4810

This chapter describes how to enable FIPS cryptography requirements on the Dell Force10 S4810 platform. This feature provides cryptographic algorithms conforming to various FIPS standards published by the National Institute of Standards and Technology (NIST), a non-regulatory agency of the US Department of Commerce. The FIPS mode is also validated for numerous platforms to meet the FIPS-140-2 standard for a software-based cryptographic module.

Note: The FIPS mode included in this release is the OpenSSL FIPS Object Module v2.0, which has been validated to meet FIPS-140-2 requirements, per certificate #1747. The S4810 platform is not one of the validated platforms. Dell Force10 has contracted with the OpenSSL Foundation to complete a Change Letter validation of the S4810 platform for this FIPS mode. A patch release will be available once that Change Letter validation has been completed.

Note: For release 8.3.12.0 only the SSH and SCP copy features use the FIPS cryptographic mode to secure management interface user sessions and file transfers. Other features that use cryptographic algorithms do not, or cannot, use the FIPS mode. The administrator must configure the management interfaces to limit access to/from the system to SSH alone.

This chapter describes the FIPS configuration procedure:

- Preparing the System
- Enabling FIPS Mode
- Generating Host-Keys
- Monitoring FIPS Mode Status
- Disabling the FIPS Mode

Preparing the System

Before you enable FIPS mode on the S4810, Dell Force10 recommends making the following steps to your system:

- disable the Telnet server (only SSH (Secure Shell) should be used to access the system).
- disable the FTP server (only SCP (Secure Copy) should be used to transfer files to and from the system).
- Attach a secure, standalone host to the console port to be used for FIPS configuration.

Enabling FIPS Mode

You must use the console port to enable or disable FIPS mode. The host attached to the console port must be secured against unauthorized access. Any attempts to enable or disable FIPS mode from a virtual terminal session are denied.
To enable FIPS mode:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable FIPS mode from a console port.</td>
<td>fips mode enable</td>
<td>CONFIG</td>
</tr>
</tbody>
</table>

When the FIPS mode is enabled, the following actions are taken:

- If enabled, the SSH server will be disabled.
- All open SSH and Telnet sessions, as well as all SCP and FTP file transfers, will be closed.
- Any existing host keys (both RSA and RSA1) will be deleted from system memory and NVRAM storage.
- The FIPS mode is enabled.
  - If the SSH server was enabled when the `fips mode enable` command was entered, it will be re-enabled for version 2 only.
  - If the SSH server is re-enabled, a new RSA host key-pair will be generated automatically. This key-pair can also be created manually using the `crypto key generate` command.

**Note:** Under certain unusual circumstances, it is possible for the FIPS enable command to indicate a failure.

- This will occur if any of the self-tests fail when FIPS mode is enabled
- This will occur if there were existing SSH/Telnet sessions that could not be closed successfully in a reasonable amount of time. In general this can occur if a user at a remote host is in the process of establishing an SSH session to the local system, and has been prompted to accept a new host key or to enter a password, but is not responding to the request. Assuming this is a transient condition, attempting to enable FIPS mode again should be successful.

## Generating Host-Keys

When FIPS mode is enabled, or disabled, the system will delete the current public/private host-key pair, terminate any SSH sessions that are in progress (deleting all the per-session encryption key information), actually enable/test the FIPS mode, generate new host-keys, and re-enable the SSH server (assuming it was enabled before enabling FIPS). Refer to **SSH Server and SCP Commands** in the **Security** chapter of the *FTOS Command Line Interface Guide* for more information.
Monitoring FIPS Mode Status

The status of the current FIPS mode (Enabled/Disabled) can be viewed directly using either the `show fips status` command or the `show system` command as shown below.

FTOS#show fips status
FIPS Mode : Enabled

FTOS#show system
Stack MAC : 00:01:e8:8a:ff:0c
Reload Type : normal-reload [Next boot : normal-reload]
--  Unit 0 --
Unit Type : Management Unit
Status : online
Next Boot : online
Required Type : S4810 - 52-port GE/TE/FG (SE)
Current Type : S4810 - 52-port GE/TE/FG (SE)
Master priority : 0
Hardware Rev : 3.0
Num Ports : 64
Up Time : 7 hr, 3 min
FTOS Version : 4810-8-5-7-1061
Jumbo Capable : yes
POE Capable : no
FIPS Mode : enabled
Burned In MAC : 00:01:e8:8a:ff:0c
No Of MACs : 3
...

Disabling the FIPS Mode

Use the console port to disable FIPS mode.

To disable the FIPS mode:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>To disable FIPS mode from a console port.</td>
<td>no fips mode enable</td>
<td>CONFIG</td>
</tr>
</tbody>
</table>

The following Warning message displays:

WARNING: Disabling FIPS mode will close all SSH/Telnet connections, restart those servers, and destroy all configured host keys.

Proceed (y/n) ?

When the FIPS mode is disabled, the following changes occur:

- The SSH server is disabled.
- All open SSH and Telnet sessions, as well as all SCP and FTP file transfers, are closed.
- Any existing host keys (both RSA and RSA1) are deleted from system memory and NVRAM storage.
- The FIPS mode is disabled.
- The SSH server is re-enabled.
- The telnet server is re-enabled if it is present in the configuration
- New 1024-bit RSA and RSA1 host key-pairs are created.
FIP Snooping

FIP snooping is supported on the following platforms: Z S4810

This chapter describes the FIP snooping concepts and configuration procedures:

- Fibre Channel over Ethernet
- Ensuring Robustness in a Converged Ethernet Network
- FIP Snooping on Ethernet Bridges
- FIP Snooping in a Switch Stack
- Configuring FIP Snooping
- Displaying FIP Snooping Information
- FIP Snooping Configuration Example

Fibre Channel over Ethernet

Fibre Channel over Ethernet (FCoE) provides a converged Ethernet network that allows the combination of storage-area network (SAN) and LAN traffic on a Layer 2 link by encapsulating Fibre Channel data into Ethernet frames.

FCoE works with the Ethernet enhancements provided in data center bridging (DCB) to support lossless (no-drop) SAN and LAN traffic. In addition, DCB provides flexible bandwidth sharing for different traffic types, such as LAN and SAN, according to 802.1p priority classes of service. For more information, refer to the Data Center Bridging (DCB) chapter in the S4810 FTOS Configuration Guide.

Ensuring Robustness in a Converged Ethernet Network

Fibre Channel networks used for SAN traffic employ switches that operate as trusted devices. End devices log into the switch to which they are attached in order to communicate with other end devices attached to the Fibre Channel network. Because Fibre Channel links are point-to-point, a Fibre Channel switch controls all storage traffic that an end device sends and receives over the network. As a result, the switch can enforce zoning configurations, ensure that end devices use their assigned addresses, and secure the network from unauthorized access and denial-of-service attacks.
To ensure similar Fibre Channel robustness and security with FCoE in an Ethernet cloud network, the Fibre Channel over Ethernet initialization protocol (FIP) establishes virtual point-to-point links between FCoE end-devices (server ENodes and target storage devices) and FCoE forwarders (FCFs) over transit FCoE-enabled bridges.

Ethernet bridges commonly provide access control list (ACLs) that can emulate a point-to-point link by providing the traffic enforcement required to create a Fibre Channel-level of robustness. In addition, FIP serves as a Layer 2 protocol to:

- Operate between FCoE end-devices and FCFs over intermediate Ethernet bridges to prevent unauthorized access to the network and achieve the required security.
- Allow transit Ethernet bridges to efficiently monitor FIP frames passing between FCoE end-devices and an FCF, and use the FIP snooping data to dynamically configure ACLs on the bridge to only permit traffic authorized by the FCF.

FIP enables FCoE devices to discover one another, initialize and maintain virtual links over an Ethernet network, and access storage devices in a storage area network. FIP satisfies the Fibre Channel requirement for point-to-point connections by creating a unique virtual link for each connection between an FCoE end-device and an FCF via a transit switch.

FIP provides functionality for discovering and logging in to an FCF. After discovering and logging in, FIP allows FCoE traffic to be sent and received between FCoE end-devices (ENodes) and the FCF. FIP uses its own EtherType and frame format. The following figure shows the communication that occurs between an ENode server and an FCoE switch (FCF).

FIP performs the following functions:

- FIP virtual local area network (VLAN) discovery: FCoE devices (ENodes) discover the FCoE VLANs on which to transmit and receive FIP and FCoE traffic.
- FIP discovery: FCoE end-devices and FCFs are automatically discovered.
- Initialization: FCoE devices perform fabric login (FLOGI) and fabric discovery (FDISC) to create a virtual link with an FCoE switch.
- Maintenance: A valid virtual link between an FCoE device and an FCoE switch is maintained and the link termination logout (LOGO) functions properly.
FIP Snooping on Ethernet Bridges

In a converged Ethernet network, intermediate Ethernet bridges can snoop on FIP packets during the login process on an FCF. Then, using ACLs, a transit bridge can permit only authorized FCoE traffic to be transmitted between an FCoE end-device and an FCF. An Ethernet bridge that provides these functions is called a FIP snooping bridge (FSB).

On a FIP snooping bridge, ACLs are created dynamically as FIP login frames are processed. The ACLs are installed on switch ports configured for any of the following port modes:

- ENode mode for server-facing ports
- FCF mode for a trusted port directly connected to an FCF

You must enable FIP snooping on the switch, configure the FIP snooping parameters, and configure CAM allocation for FCoE (optional in FTOS version 9.1.(0.0)). When you enable FIP snooping, all ports on the switch by default become ENode ports.

Dynamic ACL generation on the switch operating as a FIP snooping bridge functions as follows:
- Port-based ACLs are applied on all three port modes: on ports directly connected to an FCF, server-facing ENode ports, and bridge-to-bridge links.
- Port-based ACLs take precedence over global ACLs.
- FCoE-generated ACLs take precedence over user-configured ACLs. A user-configured ACL entry cannot deny FCoE and FIP snooping frames.

Figure 18-2 shows a switch used as a FIP snooping bridge in a converged Ethernet network. The ToR switch operates as an FCF for FCoE traffic. Converged LAN and SAN traffic is transmitted between the ToR switch and an S4810 switch. The switch operates as a lossless FIP snooping bridge to transparently forward FCoE frames between the ENode servers and the FCF switch.

**Figure 18-2. FIP Snooping on an S4810 Switch**

![Diagram of FIP Snooping on an S4810 Switch](image)
The following sections describe how to configure the FIP snooping feature on a switch that functions as a FIP snooping bridge so that it can perform the following functions:

- Allocate CAM resources for FCoE (optional in FTOS version 9.1.(0.0)).
- Perform FIP snooping (allowing and parsing FIP frames) globally on all VLANs or on a per-VLAN basis.
- Set the FCoE MAC address prefix (FC-MAP) value used by an FCF to assign a MAC address to an FCoE end-device (server ENode or storage device) after a server successfully logs in. The FC-MAP value is used in the ACLs installed in bridge-to-bridge links on the switch.
- Set the FCF or Bridge-to-Bridge Port modes to provide additional port security on ports that are directly connected to an FCF and have links to other FIP snooping bridges.
- Check FIP snooping-enabled VLANs to ensure that they are operationally active.
- Process FIP VLAN discovery requests and responses, advertisements, solicitations, FLOGI/FDISC requests and responses, FLOGO requests and responses, keep-alive packets, and clear virtual-link messages.

**FIP Snooping in a Switch Stack**

FIP snooping supports switch stacking as follows:

- A switch stack configuration is synchronized with the standby stack unit.
- Dynamic population of the FCoE database (ENode, Session, and FCF tables) is synchronized with the standby stack unit. The FCoE database is maintained by snooping FIP keep-alive messages.
- In case of a failover, the new master switch starts the required timers for the FCoE database tables. Timers run only on the master stack unit.

**Configuring FIP Snooping**

The configuration of FIP snooping consists of the following tasks:

1. Enable the FIP snooping feature on a switch to maintain FIP snooping information on the switch.
2. Enable FIP snooping on all VLANs (globally) or individual VLANs on a FIP snooping bridge.
3. Configure the FC-Map value applied globally by the switch on all VLANs or an individual VLAN.
4. Configure FCoE-Trusted mode for a FIP snooping bridge-to-bridge link.
5. Configure FCF mode for a FIP snooping bridge-to-FCF link.

For a sample FIP snooping configuration, refer to Figure 18-11.
Enabling the FIP Snooping Feature

**Note:** FIP Snooping is disabled by default. To enable this feature, you must follow the Configuration Procedure.

As soon as you enable the FIP snooping feature on a switch-bridge, existing VLAN-specific and FIP snooping configurations are applied. The FCoE database is populated when the switch connects to a converged network adapter (CNA) or FCF port and compatible DCB configurations are synchronized. By default, all FCoE and FIP frames are dropped unless specifically permitted by existing FIP snooping-generated ACLs. You can reconfigure any of the FIP snooping settings.

If you disable FIP snooping, FIP and FCoE traffic are handled as normal Ethernet frames and no FIP snooping ACLs are generated. The VLAN-specific and FIP snooping configuration is disabled and stored until you re-enable FIP snooping and the configurations are re-applied.

Enabling FIP Snooping on VLANS

You can enable FIP snooping globally on a switch on all VLANs or on a specified VLAN. When you enable FIP snooping on VLANs:

- FIP frames are allowed to pass through the switch on the enabled VLANs and are processed to generate FIP snooping ACLs.
- FCoE traffic is allowed on VLANs only after a successful virtual-link initialization (fabric login FLOGI) between an ENode and an FCF. All other FCoE traffic is dropped.
- You must configure at least one interface for FCF (FIP snooping bridge-bridge) mode on a FIP snooping-enabled VLAN. You can configure multiple FCF trusted interfaces in a VLAN.
- A maximum of eight VLANS are supported for FIP snooping on the switch. When enabled globally, FIP snooping processes FIP packets in traffic only from the first eight incoming VLANs. When enabled on a per-VLAN basis, FIP snooping is supported on up to eight VLANs.

Configuring the FC-MAP Value

You can configure the FC-MAP value to be applied globally by the switch on all or individual FCoE VLANs to authorize FCoE traffic.

The configured FC-MAP value is used to check the FC-MAP value for the MAC address assigned to ENodes in incoming FCoE frames. If the FC-MAP value does not match, FCoE frames are dropped. A session between an ENode and an FCF is established by the switch-bridge only when the FC-MAP value on the FCF matches the FC-MAP value on the FIP snooping bridge.
Configuring a Port for a Bridge-to-Bridge Link

If a switch port is connected to another FIP snooping bridge, configure the FCoE-Trusted Port mode for bridge-bridge links. Initially, all FCoE traffic is blocked. Only FIP frames with the ALL_FCF_MAC and ALL_ENODE_MAC values in their headers are allowed to pass. After the switch learns the MAC address of a connected FCF, it allows FIP frames destined to or received from the FCF MAC address.

FCoE traffic is allowed on the port only after the switch learns the FC-MAP value associated with the specified FCF MAC address and verifies that it matches the configured FC-MAP value for the FCoE VLAN.

Configuring a Port for a Bridge-to-FCF Link

If a port is directly connected to an FCF, configure the port mode as FCF. Initially, all FCoE traffic is blocked; only FIP frames are allowed to pass.

FCoE traffic is allowed on the port only after a successful FLOGI request/response and confirmed use of the configured FC-MAP value for the VLAN.

Impact on Other Software Features

When you enable FIP snooping on a switch, other software features are impacted as follows:

- MAC address learning: MAC address learning is not performed on FIP and FCoE frames, which are denied by ACLs dynamically created by FIP snooping on server-facing ports in ENode mode.
- MTU auto-configuration: MTU size is set to mini-jumbo (2500 bytes) when a port is in Switchport mode, the FIP snooping feature is enabled on the switch, and FIP snooping is enabled on all or individual VLANs.
- Link aggregation group (LAG): FIP snooping is supported on port channels on ports on which PFC mode is on (PFC is operationally up).
- STP: If you enable an STP protocol (STP, RSTP, PVSTP, or MSTP) on the switch and ports enter a blocking state, when the state change occurs, the corresponding port-based ACLs are deleted. If a port is enabled for FIP snooping in ENode or FCF mode, the ENode/FCF MAC-based ACLs are deleted.

FIP Snooping Prerequisites

Before you configure FIP snooping on a switch, ensure that the following conditions are met:

- A FIP snooping bridge requires DCBX and PFC to be enabled on the switch for lossless Ethernet connections (refer to the Data Center Bridging (DCB) chapter in the S4810 FTOS Configuration Guide). Dell Force10 recommends that you also enable ETS; ETS is recommended but not required.
  If you enable DCBX and PFC mode is on (PFC is operationally up) in a port configuration, FIP snooping is operational on the port. If the PFC parameters in a DCBX exchange with a peer are not synchronized, FIP and FCoE frames are dropped on the port after you enable the FIP snooping feature.
• VLAN membership:
  • You must create the VLANs on the switch which handles FCoE traffic (interface vlan command).
  • You must configure each FIP snooping port to operate in Hybrid mode so that it accepts both tagged and untagged VLAN frames (portmode hybrid command).
  • You must configure tagged VLAN membership on each FIP snooping port that sends and receives FCoE traffic and has links with an FCF, ENode server, or another FIP snooping bridge (tagged port-type slot/port command).

The default VLAN membership of the port should continue to operate with untagged frames. FIP snooping is not supported on a port that is configured for non-default untagged VLAN membership.

**FIP Snooping Restrictions**

The following restrictions apply when you configure FIP snooping:

• The maximum number of FCoE VLANs supported on the switch is 8.
• The maximum number of FIP snooping sessions supported per ENode server is 16.
• The maximum number of FCFs supported per FIP snooping-enabled VLAN is 4.

**Configuration Procedure**

You can enable FIP snooping globally on all FCoE VLANs on a switch or on an individual FCoE VLAN. By default, FIP snooping is disabled.

To enable FIP snooping on the switch and configure FIP snooping parameters on ports, follow these steps:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure FCoE. The configuration files are stored in the flash memory in the CONFIG_TEMPLATE.</td>
<td>FCoE configuration: copy FCoE_DCB_Config running-config</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>2</td>
<td>Save the configuration on the switch.</td>
<td>write memory</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>3</td>
<td>Reload the switch to enable the configuration. After the switch is reloaded, DCB/DCBX will be enabled.</td>
<td>reload</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>4</td>
<td>Enable the FIP snooping feature on a switch.</td>
<td>feature fip-snooping</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>5</td>
<td>Enable FIP snooping on all VLANs or on a specified VLAN.</td>
<td>fip-snooping enable</td>
<td>CONFIGURATION Or VLAN INTERFACE</td>
</tr>
<tr>
<td>6</td>
<td>Configure the port for bridge-to-FCF links.</td>
<td>fip-snooping port-mode fcf</td>
<td>INTERFACE CONFIGURATION</td>
</tr>
</tbody>
</table>

**Note:** To disable the FIP snooping feature or FIP snooping on VLANs, use the no version of a command; for example, no feature fip-snooping or no fip-snooping enable.
Displaying FIP Snooping Information

Use the `show` commands in Table 18-1 to display information on FIP snooping.

**Table 18-1. Displaying FIP Snooping Information**

<table>
<thead>
<tr>
<th>Command</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>show fip-snooping sessions [interface vlan vlan-id]</code></td>
<td>Displays information on FIP-snooped sessions on all VLANs or a specified VLAN, including the ENode interface and MAC address, the FCF interface and MAC address, VLAN ID, FCoE MAC address and FCoE session ID number (FC-ID), worldwide node name (WWNN) and the worldwide port name (WWPN).</td>
</tr>
<tr>
<td><code>show fip-snooping config</code></td>
<td>Displays the FIP snooping status and configured FC-MAP values.</td>
</tr>
<tr>
<td><code>show fip-snooping enode [enode-mac-address]</code></td>
<td>Displays information on the ENodes in FIP-snooped sessions, including the ENode interface and MAC address, FCF MAC address, VLAN ID and FC-ID.</td>
</tr>
<tr>
<td><code>show fip-snooping fcf [fcf-mac-address]</code></td>
<td>Displays information on the FCFs in FIP-snooped sessions, including the FCF interface and MAC address, FCF interface, VLAN ID, FC-MAP value, FKA advertisement period, and number of ENodes connected.</td>
</tr>
<tr>
<td>`clear fip-snooping database interface vlan vlan-id [fcoe-mac-address</td>
<td>enode-mac-address</td>
</tr>
<tr>
<td>`show fip-snooping statistics [interface vlan vlan-id</td>
<td>interface port-type port/slot</td>
</tr>
<tr>
<td>`clear fip-snooping statistics [interface vlan vlan-id</td>
<td>interface port-type port/slot</td>
</tr>
<tr>
<td><code>show fip-snooping system</code></td>
<td>Display information on the status of FIP snooping on the switch (enabled or disabled), including the number of FCoE VLANs, FCFs, ENodes, and currently active sessions.</td>
</tr>
<tr>
<td><code>show fip-snooping vlan</code></td>
<td>Display information on the FCoE VLANs on which FIP snooping is enabled.</td>
</tr>
</tbody>
</table>

**Figure 18-3. show fip-snooping sessions Command Example**

```
FTOS#show fip-snooping sessions
Enode MAC  Enode Intf  FCF MAC  FCF Intf  VLAN
aa:bb:cc:00:00:00  Te 0/42  aa:bb:cd:00:00:00  Te 0/43  100
aa:bb:cc:00:00:00  Te 0/42  aa:bb:cd:00:00:00  Te 0/43  100
aa:bb:cc:00:00:00  Te 0/42  aa:bb:cd:00:00:00  Te 0/43  100
aa:bb:cc:00:00:00  Te 0/42  aa:bb:cd:00:00:00  Te 0/43  100
aa:bb:cc:00:00:00  Te 0/42  aa:bb:cd:00:00:00  Te 0/43  100
Fcoe MAC  FC-ID  Port WWPN  Port WWNN
0e:fc:00:10:00:01  01:00:00 31:00:0e:fc:00:00:00:00  21:00:0e:fc:00:00:00:00
0e:fc:00:10:00:02  01:00:00 41:00:0e:fc:00:00:00:00  21:00:0e:fc:00:00:00:00
0e:fc:00:10:00:03  01:00:00 41:00:0e:fc:00:00:00:00:01  21:00:0e:fc:00:00:00:00
0e:fc:00:10:00:04  01:00:00 41:00:0e:fc:00:00:00:00:02  21:00:0e:fc:00:00:00:00
0e:fc:00:10:00:05  01:00:00 41:00:0e:fc:00:00:00:00:03  21:00:0e:fc:00:00:00:00
```
Table 18-2. show fip-snooping sessions Command Description

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENode MAC</td>
<td>MAC address of the ENode.</td>
</tr>
<tr>
<td>ENode Interface</td>
<td>Slot/ port number of the interface connected to the ENode.</td>
</tr>
<tr>
<td>FCF MAC</td>
<td>MAC address of the FCF.</td>
</tr>
<tr>
<td>FCF Interface</td>
<td>Slot/ port number of the interface to which the FCF is connected.</td>
</tr>
<tr>
<td>VLAN</td>
<td>VLAN ID number used by the session.</td>
</tr>
<tr>
<td>FCoE MAC</td>
<td>MAC address of the FCoE session assigned by the FCF.</td>
</tr>
<tr>
<td>FC-ID</td>
<td>Fibre Channel ID assigned by the FCF.</td>
</tr>
<tr>
<td>Port WWPN</td>
<td>Worldwide port name of the CNA port.</td>
</tr>
<tr>
<td>Port WWNN</td>
<td>Worldwide node name of the CNA port.</td>
</tr>
</tbody>
</table>

Figure 18-4. show fip-snooping config Command Example

FTOS# show fip-snooping config
FIP Snooping Feature enabled Status: Enabled
FIP Snooping Global enabled Status: Enabled
Global FC-MAP Value: 0X0EFC00

FIP Snooping enabled VLANs
<table>
<thead>
<tr>
<th>VLAN</th>
<th>Enabled</th>
<th>FC-MAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>TRUE</td>
<td>0X0EFC00</td>
</tr>
</tbody>
</table>

Table 18-3. show fip-snooping enode Command Description

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENode MAC</td>
<td>MAC address of the ENode.</td>
</tr>
<tr>
<td>ENode Interface</td>
<td>Slot/ port number of the interface connected to the ENode.</td>
</tr>
<tr>
<td>FCF MAC</td>
<td>MAC address of the FCF.</td>
</tr>
<tr>
<td>VLAN</td>
<td>VLAN ID number used by the session.</td>
</tr>
<tr>
<td>FC-ID</td>
<td>Fibre Channel session ID assigned by the FCF.</td>
</tr>
</tbody>
</table>

Figure 18-5. show fip-snooping enode Command Example

FTOS# show fip-snooping enode
<table>
<thead>
<tr>
<th>ENode MAC</th>
<th>ENode Interface</th>
<th>FCF MAC</th>
<th>VLAN</th>
<th>FC-ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>d4:ae:52:1b:e3:cd</td>
<td>Te 0/11</td>
<td>54:7f:ee:37:34:40</td>
<td>100</td>
<td>62:00:11</td>
</tr>
</tbody>
</table>

Table 18-4. show fip-snooping fcf Command Description

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCF MAC</td>
<td>MAC address of the FCF.</td>
</tr>
<tr>
<td>FCF Interface</td>
<td>Slot/ port number of the interface to which the FCF is connected.</td>
</tr>
<tr>
<td>VLAN</td>
<td>VLAN ID number used by the session.</td>
</tr>
<tr>
<td>FC-MAP</td>
<td>Fibre Channel session ID assigned by the FCF.</td>
</tr>
<tr>
<td>FKA ADV_PERIOD</td>
<td></td>
</tr>
<tr>
<td>No. of Enodes</td>
<td></td>
</tr>
</tbody>
</table>

Figure 18-6. show fip-snooping fcf Command Example

FTOS# show fip-snooping fcf
<table>
<thead>
<tr>
<th>FCF MAC</th>
<th>FCF Interface</th>
<th>VLAN</th>
<th>FC-MAP</th>
<th>FKA ADV_PERIOD</th>
<th>No. of Enodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>54:7f:ee:37:34:40</td>
<td>Po 22</td>
<td>100</td>
<td>0e:fc:00</td>
<td>4000</td>
<td>2</td>
</tr>
<tr>
<td>Field</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCF MAC</td>
<td>MAC address of the FCF.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCF Interface</td>
<td>Slot/port number of the interface to which the FCF is connected.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VLAN</td>
<td>VLAN ID number used by the session.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FC-MAP</td>
<td>FC-Map value advertised by the FCF.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENode Interface</td>
<td>Slot/number of the interface connected to the ENode.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FKA_ADV_PERIOD</td>
<td>Period of time (in milliseconds) during which FIP keep-alive advertisements are transmitted.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No of ENodes</td>
<td>Number of ENodes connected to the FCF.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FC-ID</td>
<td>Fibre Channel session ID assigned by the FCF.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 18-7. show fip-snooping statistics (VLAN and port) Command Example

FTOS# show fip-snooping statistics interface vlan 100
Number of Vlan Requests :0
Number of Vlan Notifications :0
Number of Multicast Discovery Solicits :2
Number of Unicast Discovery Solicits :0
Number of FLOGI :2
Number of FDISC :16
Number of FLOGO :0
Number of Enode Keep Alive :9021
Number of VN Port Keep Alive :3349
Number of Multicast Discovery Advertisement :4437
Number of Unicast Discovery Advertisement :2
Number of FLOGI Accepts :2
Number of FLOGI Rejects :0
Number of FDISC Accepts :16
Number of FDISC Rejects :0
Number of FLOGO Accepts :0
Number of FLOGO Rejects :0
Number of CVL :0
Number of FCF Discovery Timeouts :0
Number of VN Port Session Timeouts :0
Number of Session failures due to Hardware Config :0
FTOS(conf)#

FTOS# show fip-snooping statistics int tengigabitethernet 0/11
Number of Vlan Requests :1
Number of Vlan Notifications :0
Number of Multicast Discovery Solicits :1
Number of Unicast Discovery Solicits :0
Number of FLOGI :1
Number of FDISC :16
Number of FLOGO :0
Number of Enode Keep Alive :4416
Number of VN Port Keep Alive :3136
Number of Multicast Discovery Advertisement :0
Number of Unicast Discovery Advertisement :0
Number of FLOGI Accepts :0
Number of FLOGI Rejects :0
Number of FDISC Accepts :0
Number of FDISC Rejects :0
Number of FLOGO Accepts :0
Number of FLOGO Rejects :0
Number of CVL :0
Number of FCF Discovery Timeouts :0
Number of VN Port Session Timeouts :0
Number of Session failures due to Hardware Config :0
Figure 18-8. show fip-snooping statistics (port channel) Command Example

FTOS# show fip-snooping statistics interface port-channel 22
Number of VLAN Requests : 0
Number of VLAN Notifications : 2
Number of Multicast Discovery Solicits : 0
Number of Unicast Discovery Solicits : 0
Number of FLOGI : 0
Number of FDISC : 0
Number of FLOGO : 0
Number of ENode Keep Alive : 0
Number of VN Port Keep Alive : 0
Number of Multicast Discovery Advertisement : 4451
Number of Unicast Discovery Advertisement : 2
Number of FLOGI Accepts : 2
Number of FLOGI Rejects : 0
Number of FDISC Accepts : 16
Number of FDISC Rejects : 0
Number of FLOGO Accepts : 0
Number of FLOGO Rejects : 0
Number of CVL : 0
Number of FCP Discovery Timeouts : 0
Number of VN Port Session Timeouts : 0
Number of Session failures due to Hardware Config : 0

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of VLAN Requests</td>
<td>Number of FIP snooped VLAN request frames received on the interface.</td>
</tr>
<tr>
<td>Number of VLAN Notifications</td>
<td>Number of FIP snooped VLAN notification frames received on the interface.</td>
</tr>
<tr>
<td>Number of Multicast Discovery Solicits</td>
<td>Number of FIP snooped multicast discovery solicit frames received on the interface.</td>
</tr>
<tr>
<td>Number of Unicast Discovery Solicits</td>
<td>Number of FIP snooped unicast discovery solicit frames received on the interface.</td>
</tr>
<tr>
<td>Number of FLOGI</td>
<td>Number of FIP snooped FLOGI request frames received on the interface.</td>
</tr>
<tr>
<td>Number of FDISC</td>
<td>Number of FIP snooped FDISC request frames received on the interface.</td>
</tr>
<tr>
<td>Number of FLOGO</td>
<td>Number of FIP snooped FLOGO frames received on the interface.</td>
</tr>
<tr>
<td>Number of ENode Keep Alives</td>
<td>Number of FIP snooped ENode keep-alive frames received on the interface.</td>
</tr>
<tr>
<td>Number of VN Port Keep Alives</td>
<td>Number of FIP snooped VN port keep-alive frames received on the interface.</td>
</tr>
<tr>
<td>Number of Multicast Discovery Advertisements</td>
<td>Number of FIP snooped multicast discovery advertisements received on the interface.</td>
</tr>
<tr>
<td>Number of Unicast Discovery Advertisements</td>
<td>Number of FIP snooped unicast discovery advertisements received on the interface.</td>
</tr>
<tr>
<td>Number of FLOGI Accepts</td>
<td>Number of FIP FLOGI accept frames received on the interface.</td>
</tr>
<tr>
<td>Number of FLOGI Rejects</td>
<td>Number of FIP FLOGI reject frames received on the interface.</td>
</tr>
<tr>
<td>Number of FDISC Accepts</td>
<td>Number of FIP FDISC accept frames received on the interface.</td>
</tr>
</tbody>
</table>
Table 18-5. show fip-snooping statistics Command Descriptions

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of FDISC Rejects</td>
<td>Number of FIP FDISC reject frames received on the interface.</td>
</tr>
<tr>
<td>Number of FLOGO Accepts</td>
<td>Number of FIP FLOGO accept frames received on the interface.</td>
</tr>
<tr>
<td>Number of FLOGO Rejects</td>
<td>Number of FIP FLOGO reject frames received on the interface.</td>
</tr>
<tr>
<td>Number of CVLs</td>
<td>Number of FIP clear virtual link frames received on the interface.</td>
</tr>
<tr>
<td>Number of FCF Discovery Timouts</td>
<td>Number of FCF discovery timeouts that occurred on the interface.</td>
</tr>
<tr>
<td>Number of VN Port Session Timouts</td>
<td>Number of VN port session timeouts that occurred on the interface.</td>
</tr>
<tr>
<td>Number of Session failures due to Hardware Config</td>
<td>Number of session failures due to hardware configuration that occurred on the interface.</td>
</tr>
</tbody>
</table>

Figure 18-9. show fip-snooping system Command Example

```
FTOS# show fip-snooping system
Global Mode : Enabled
FCOE VLAN List (Operational) : 1, 100
FCFs : 1
Enodes : 2
Sessions : 17
```

Figure 18-10. show fip-snooping vlan Command Example

```
FTOS# show fip-snooping vlan
* = Default VLAN

VLAN FC-MAP FCFs Enodes Sessions
----- ---- ---- ---- --------
    -    -    -   -       -
*1  -    -    -   -       -
100 0X0EFC00 1  2   17
```

FIP Snooping Configuration Example

Figure 18-11 shows a S4810 switch used as a FIP snooping bridge for FCoE traffic between an ENode (server blade) and an FCF (ToR switch). The ToR switch operates as an FCF and FCoE gateway.
In Figure 18-11, DCBX and PFC are enabled on the FIP snooping bridge and on the FCF ToR switch. On the FIP snooping bridge, DCBX is configured as follows:

- A server-facing port is configured for DCBX in an auto-downstream role.
- An FCF-facing port is configured for DCBX in an auto-upstream or configuration-source role.

The DCBX configuration on the FCF-facing port is detected by the server-facing port and the DCB PFC configuration on both ports is synchronized. For more information about how to configure DCBX and PFC on a port, refer to Data Center Bridging (DCB) in the S4810 FTOS Configuration Guide.

Figure 18-12 shows how to configure FIP snooping on FCoE VLAN 10, on an FCF-facing port (0/50), on an ENode server-facing port (0/1), and to configure the FIP snooping ports as tagged members of the FCoE VLAN enabled for FIP snooping.
Enable the FIP snooping feature on the switch (FIP snooping bridge):
FTOS(conf)# feature fip-snooping

Enable FIP snooping on FCoE VLAN 10:
FTOS(conf)# interface vlan 10
FTOS(conf-if-vl-10)# fip-snooping enable

Enable an FC-MAP value on VLAN 10:
FTOS(conf-if-vl-10)# fip-snooping fc-map 0x0EFC01
Note: Configuring an FC-MAP value is only required if you do not use the default FC-MAP value (0x0EFC00).

Configure the ENode server-facing port:
FTOS(conf)# interface tengigabitethernet 0/1
FTOS(conf-if-te-0/1)# portmode hybrid
FTOS(conf-if-te-0/1)# switchport
FTOS(conf-if-te-0/1)# protocol lldp
FTOS(conf-if-te-0/1-lldp)# dcbx port-role auto-downstream
Note: A port is enabled by default for bridge-ENode links.

Configure the FCF-facing port:
FTOS(conf)# interface tengigabitethernet 0/50
FTOS(conf-if-te-0/50)# portmode hybrid
FTOS(conf-if-te-0/50)# switchport
FTOS(conf-if-te-0/50)# fip-snooping port-mode fcf
FTOS(conf-if-te-0/50)# protocol lldp
FTOS(conf-if-te-0/50-lldp)# dcbx port-role auto-upstream

Configure FIP snooping ports as tagged members of FCoE VLAN:
FTOS(conf)# interface vlan 10
FTOS(conf-if-vl-10)# tagged tengigabitethernet 0/1
FTOS(conf-if-vl-10)# tagged tengigabitethernet 0/50
FTOS(conf-if-te-0/1)# no shut
FTOS(conf-if-te-0/50)# no shut
FTOS(conf-if-vl-10)# no shut

After FIP packets are exchanged between the ENode and the switch, a FIP snooping session is established. ACLS are dynamically generated for FIP snooping on the FIP snooping bridge/switch.
Force10 Resilient Ring Protocol (FRRP)

Force10 Resilient Ring Protocol (FRRP) is supported on platforms: E C S Z 54810

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. Refer to the illustration below for a simple example of this FRRP topology. Note that ring direction is determined by the Master node’s Primary and Secondary ports.
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 are allowed per system and multiple rings can be run on one system. However, it is not recommended on the S20/S50 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.

The S4810, S55, and S60 support up to 32 rings on a system (including stacked units).

**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 below, 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.

**Important FRRP Points**

FRRP provides a convergence time that can generally range between 150ms and 1500ms for Layer 2 networks. The master node originates a high-speed frame that circulates around the ring. This frame, appropriately, sets up or breaks down the ring.

- 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 per ring—all other nodes are Transit
- Each node has 2 member interfaces—Primary, Secondary
- No limit to the number of nodes on a ring
- Master node ring port states—blocking, pre-forwarding, forwarding, disabled
- Transit node ring port states—blocking, pre-forwarding, forwarding, disabled
- STP disabled on ring interfaces
- 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 19-1, "FRRP Components," in Force10 Resilient Ring Protocol (FRRP) lists some important FRRP concepts.

### Table 19-1. FRRP Components

<table>
<thead>
<tr>
<th>Concept</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ring ID</td>
<td>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 the illustration in Member VLAN Spanning Two Rings Connected by One Switch.</td>
</tr>
<tr>
<td>Control VLAN</td>
<td>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.</td>
</tr>
<tr>
<td>Member VLAN</td>
<td>Each ring maintains a list of member VLANs. Member VLANs must be consistent across the entire ring.</td>
</tr>
<tr>
<td>Port Role</td>
<td>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.</td>
</tr>
</tbody>
</table>
| 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. |
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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>protocol frrp</td>
<td>CONFIGURATION</td>
<td>Create the FRRP group with this Ring ID</td>
</tr>
<tr>
<td>ring-id</td>
<td></td>
<td>Ring ID: 1-255</td>
</tr>
</tbody>
</table>

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, refer to 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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interface vlan</td>
<td>CONFIGURATION</td>
<td>Create a VLAN with this ID number</td>
</tr>
<tr>
<td></td>
<td>vlan-id</td>
<td></td>
<td>VLAN ID: 1-4094</td>
</tr>
</tbody>
</table>
### 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, refer to [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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interface vlan vlan-id</td>
<td>CONFIGURATION</td>
<td>Create a VLAN with this ID number VLAN ID: 1-4094</td>
</tr>
</tbody>
</table>
| 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 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.
  - For a 40-Gigabit Ethernet interface, enter the keyword fortyGigE 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 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.
  - For a 40-Gigabit Ethernet interface, enter the keyword fortyGigE followed by the slot/port information.

Slot/Port: Slot and Port ID for the interface. VLAN ID: Identification number of the Control VLAN |
### Set FRRP Timers

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>timer &lt;hello-interval</td>
<td>CONFIG-FRRP</td>
<td>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.</td>
</tr>
</tbody>
</table>

### Clear FRRP counters

Use one of the following commands to clear the FRRP counters.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>clear frrp ring-id</td>
<td>EXEC PRIVELEGED</td>
<td>Clear the counters associated with this Ring ID Ring ID: 1-255</td>
</tr>
<tr>
<td>clear frrp</td>
<td>EXEC PRIVELEGED</td>
<td>Clear the counters associated with all FRRP groups</td>
</tr>
</tbody>
</table>

### Show FRRP configuration

Use the following command to view the configuration for the FRRP group.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show configuration</td>
<td>CONFIG-FRRP</td>
<td>Show the configuration for this FRRP group</td>
</tr>
</tbody>
</table>
Show FRRP information

Use one of the following commands show general FRRP information.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show frrp ring-id</td>
<td>EXEC or EXEC PRIVELEGED</td>
<td>Show the information for the identified FRRP group. Ring ID: 1-255</td>
</tr>
<tr>
<td>show frrp summary</td>
<td>EXEC or EXEC PRIVELEGED</td>
<td>Show the state of all FRRP groups. Ring ID: 1-255</td>
</tr>
</tbody>
</table>

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

The following illustration is an example of a basic FRRP topology. Below the illustration are the associated 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
GARP VLAN Registration Protocol (GVRP)

GARP VLAN Registration Protocol (GVRP) is supported on platforms: E C S Z 54810

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.
• On the E-Series, C-Series, and non-S60/S55/S4810 S-Series, Per-VLAN Spanning Tree (PVST+) or MSTP and GVRP cannot be enabled at the same time, as shown in the example below. If Spanning Tree and GVRP are both required, implement RSTP. The S60, S55, and S4810 systems do support enabling GVRP and MSTP at the same time.

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 the illustration below, that type 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.
Basic GVRP configuration is a 2-step process:

1. Enabling GVRP Globally.
2. Enabling GVRP on a Layer 2 Interface.

**Related Configuration Tasks**

- Configuring GVRP Registration
- Configuring a GARP Timer

**Enabling GVRP Globally**

Enable GVRP for the entire switch using the command `gvrp enable` in CONFIGURATION mode, as shown in the following example. Use the `show gvrp brief` command to inspect the global configuration.

```
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 the following example. Use `show config` from the INTERFACE mode to inspect the interface configuration, as shown in the following example, or use the `show gvrp interface` command in EXEC or EXEC Privilege mode.

```
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 below, 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.

```
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.

```
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**

```
Force10(conf)#garp timers join 300
% Error: Leave timer should be >= 3*Join timer.
```
High Availability

**High Availability** (HA) is supported on platforms: \[ \text{C} \text{ E} \text{ S} \text{ S610} \]

**Note:** High Availability is not supported on the S60 system.

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.

<table>
<thead>
<tr>
<th>Component</th>
<th>Boot Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-Series TeraScale RPM</td>
<td>2.4.2.1</td>
</tr>
<tr>
<td>E-Series TeraScale Line Card</td>
<td>2.3.2.1</td>
</tr>
<tr>
<td>E-Series ExaScale RPM</td>
<td>2.5.1.9</td>
</tr>
<tr>
<td>E-Series ExaScale Line Card</td>
<td>2.9.1.1</td>
</tr>
<tr>
<td>C-Series RPM</td>
<td>2.7.1.1</td>
</tr>
<tr>
<td>C-Series Line Card</td>
<td>2.6.0.2</td>
</tr>
<tr>
<td>S-Series RPM</td>
<td>2.8.2.0</td>
</tr>
<tr>
<td>S-Series Line Card</td>
<td>2.8.2.0</td>
</tr>
</tbody>
</table>

The features in this collection are:

- Component Redundancy
- Online Insertion and Removal
- Hitless Behavior
- Graceful Restart
- Software Resiliency
- Warm Upgrade
- Hot-lock Behavior
Component Redundancy

Dell Force10 systems eliminate 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
- Boot the chassis with dual RPMs
- Automatic and manual RPM failover
- Support for RPM redundancy by FTOS version
- RPM synchronization

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**

```plaintext
%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 21-1, "System Behavior with RPMs with Mismatched FTOS Versions," in High Availability. View the configuration loaded on each RPM using the command `show redundancy`, as shown in the example in Automatic and manual RPM failover.

<table>
<thead>
<tr>
<th>Mismatch Condition</th>
<th>Example</th>
<th>Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>different FTOS versions with only first two digits</td>
<td>Primary: 7.4.2.0</td>
<td>The link to the standby RPM is up, and FTOS block syncs only the</td>
</tr>
<tr>
<td>matching</td>
<td>Standby: 7.4.1.0</td>
<td>startup-config. The failover type is warm upgrade. FTOS displays</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Message 2.</td>
</tr>
<tr>
<td>different FTOS versions with first two digits not</td>
<td>Primary: 7.6.1.0</td>
<td>The link to the standby RPM is down, and the standby RPM is in a boot</td>
</tr>
<tr>
<td>matching</td>
<td>Standby: 7.5.1.0</td>
<td>loop. FTOS displays Message 3 and a boot fail prompt.</td>
</tr>
<tr>
<td>different FTOS versions with only first three digits</td>
<td>Primary: 7.4.2.0</td>
<td>The link to the peer RPM is up, and FTOS performs a complete block</td>
</tr>
<tr>
<td>matching</td>
<td>Standby: 7.4.2.1</td>
<td>sync. The failover type is hot failover. FTOS displays Message 2.</td>
</tr>
</tbody>
</table>

**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.

```
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
```

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_COMMDCOWN: 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 21-2, "Failover Behaviors," in High Availability describes the failover behavior for the possible failure scenarios.

**Table 21-2.  Failover Behaviors**

<table>
<thead>
<tr>
<th>Platform</th>
<th>Failover Trigger</th>
<th>Failover Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>C,E</td>
<td>CP task crash on the primary RPM</td>
<td>The standby RPM detects the IRC time out and initiates failover, and the failed RPM reboots itself after saving a CP application core dump.</td>
</tr>
<tr>
<td>C,E</td>
<td>CP IRC timeout for a non-task crash reason on the primary RPM</td>
<td>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.</td>
</tr>
<tr>
<td>E</td>
<td>RP task or kernel crash on the primary RPM</td>
<td>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.</td>
</tr>
</tbody>
</table>
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. Refer to:

- C-Series Debugging and Diagnostics, C-Series Debugging and Diagnostics
- E-Series TeraScale Debugging and Diagnostics, Inter-CPU timeouts

### Support for RPM redundancy by FTOS version

FTOS supports increasing levels of RPM redundancy (warm and hot) as described in Table 21-3, "Support for RPM Redundancy by FTOS Version," in High Availability.

### Table 21-3. Support for RPM Redundancy by FTOS Version

<table>
<thead>
<tr>
<th>Failover Type</th>
<th>Failover Behavior</th>
<th>Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warm Failover</td>
<td>The new primary RPM remains online, while the failed RPM, all line cards, and all SFMs reboot.</td>
<td>C E</td>
</tr>
<tr>
<td>Hot Failover</td>
<td>Only the failed RPM reboots.</td>
<td>C E</td>
</tr>
<tr>
<td></td>
<td>All line cards and SFMs remain online.</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td>All application tasks are spawned on the secondary RPM before failover.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The running configuration is synchronized at runtime so it does not need to be reapplied during failover.</td>
<td></td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Failover Type</th>
<th>Synchronized Data</th>
<th>Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warm Failover</td>
<td>some NVRAM information, startup-configuration, line card configurations, user-access configurations</td>
<td>ECES</td>
</tr>
<tr>
<td>Hot Failover</td>
<td>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</td>
<td>ECES</td>
</tr>
</tbody>
</table>

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 the example in the “Force an RPM failover” section.

```
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

```
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
- Linecard Online Insertion and Removal

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 the example below.
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.

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
```

**Linecard 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 the example below.

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 the example below.

```
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.

**Message 6** Line card Mismatch Error

```
%RPM0-P:CP %CHMGR-3-CARD_MISMATCH: Mismatch: line card 0 is type E48VB - type E48TB required
```

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 the example below.

```
%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 --

<table>
<thead>
<tr>
<th>Slot</th>
<th>Status</th>
<th>NxtBoot</th>
<th>ReqTyp</th>
<th>CurTyp</th>
<th>Version</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>online</td>
<td>online</td>
<td>E48VB</td>
<td>E48VB</td>
<td>7-5-1-71</td>
<td>48</td>
</tr>
</tbody>
</table>

[output omitted]

**Hitless Behavior**

Hitless Behavior is supported only on platforms: C E S4810

Hitless behavior is supported on S4810 with FTOS 8.3.12.0 and later or the E-Series ExaScale 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 and S4810: 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. Refer to Configure LACP as Hitless.
- Spanning Tree Protocol. Refer to Configuring Spanning Trees as Hitless.
- On the E-Series only, Bi-directional Forwarding Detection (line card ports). Refer to Bidirectional Forwarding Detection (BFD).
Graceful Restart

Graceful Restart is supported on platforms: E C S 54810

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.
- Open Shortest Path First.
- Protocol Independent Multicast—Sparse Mode.

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 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 identify 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 E-Series TeraScale Debugging and Diagnostics and E-Series ExaScale Debugging and Diagnostics chapters 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 E-Series TeraScale Debugging and Diagnostics, Respond to PCDFO events or E-Series ExaScale Debugging and Diagnostics.

*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 fail, 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, and 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 provide 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:

- C-Series Debugging and Diagnostics
- E-Series TeraScale Debugging and Diagnostics
- E-Series ExaScale Debugging and Diagnostics
- S-Series 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 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 Access Control Lists (ACLs).

- **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 Policy-based Routing.

## 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 21-4, "Control Plane and Data Plane Status during Warm Upgrade," in High Availability.

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, an 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 21-4, "Control Plane and Data Plane Status during Warm Upgrade," in High Availability show the warm upgrade and downtime impact, if any, which each step.

<table>
<thead>
<tr>
<th>Table 21-4. Control Plane and Data Plane Status during Warm Upgrade</th>
<th>Download 6.3.1.1 to RPMs</th>
<th>Reboot RPM1 to Upgrade</th>
<th>Initiate Warm Failover</th>
<th>Reboot RPM0 to Upgrade</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPM 0</td>
<td>7.6.1.0 Primary</td>
<td>7.6.1.0 Primary</td>
<td>7.6.1.0 Secondary</td>
<td>7.7.1.0 Secondary</td>
</tr>
<tr>
<td>RPM 1</td>
<td>7.6.1.0 Secondary</td>
<td>7.7.1.0 Secondary</td>
<td>7.7.1.0 Primary</td>
<td>7.7.1.0 Primary</td>
</tr>
<tr>
<td>Line Cards</td>
<td>7.6.1.0</td>
<td>7.6.1.0</td>
<td>7.7.1.0</td>
<td>7.7.1.0</td>
</tr>
<tr>
<td>Control Plane</td>
<td>Operational</td>
<td>Operational</td>
<td>Interruption</td>
<td>Operational</td>
</tr>
<tr>
<td>Forwarding State</td>
<td>Forwarding</td>
<td>Forwarding</td>
<td>Interruption</td>
<td>Forwarding</td>
</tr>
</tbody>
</table>
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 (shown below) to determine the version of your RPM. There is no hardware requirement for C-Series. In the example below, the relevant information appears in red.

```
FTOS#show rpm
-- RPM card 0 --
Status : active
Next Boot : online
Card Type : RPM - Route Processor Module (LC-EF3-RPM)
Hardware Rev : 2.21
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
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
```
2. The cache boot feature requires at least the boot code versions in Table 21-5, "Boot Code Requirements for Cache Boot," in High Availability. Use show rpm and show linecard commands to verify that you have the proper version.

Table 21-5. Boot Code Requirements for Cache Boot

<table>
<thead>
<tr>
<th>Component</th>
<th>Boot Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-Series TeraScale RPM</td>
<td>2.4.2.1</td>
</tr>
<tr>
<td>E-Series TeraScale Line Card</td>
<td>2.3.2.1</td>
</tr>
<tr>
<td>E-Series ExaScale RPM</td>
<td>2.5.0.3</td>
</tr>
<tr>
<td>E-Series ExaScale Line Card</td>
<td>2.9.0.5</td>
</tr>
<tr>
<td>C-Series RPM</td>
<td>2.7.1.1</td>
</tr>
<tr>
<td>C-Series Line Card</td>
<td>2.6.0.1</td>
</tr>
</tbody>
</table>

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 the example below. 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.

FTOS#upgrade system-image all A flash://FTOS-EF-7.8.1.0.bin

Current cache boot information in the system:

<table>
<thead>
<tr>
<th>Type</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>invalid</td>
<td>invalid</td>
</tr>
<tr>
<td>RP1</td>
<td>invalid[b][n]</td>
<td>invalid</td>
</tr>
<tr>
<td>RP2</td>
<td>invalid</td>
<td>invalid</td>
</tr>
<tr>
<td>linecard 0</td>
<td>invalid</td>
<td>invalid</td>
</tr>
<tr>
<td>linecard 1</td>
<td>is not present.</td>
<td></td>
</tr>
<tr>
<td>linecard 2</td>
<td>is not present.</td>
<td></td>
</tr>
<tr>
<td>linecard 3</td>
<td>is not present.</td>
<td></td>
</tr>
</tbody>
</table>
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 the example below.

FTOS#show boot system all

Current system image information in the system:
-----------------------------------------------
<table>
<thead>
<tr>
<th>Type</th>
<th>Boot Type</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>DOWNLOAD BOOT 4.7.5.427</td>
<td>invalid</td>
<td></td>
</tr>
<tr>
<td>RP1</td>
<td>DOWNLOAD BOOT 4.7.5.427</td>
<td>invalid</td>
<td></td>
</tr>
<tr>
<td>RP2</td>
<td>DOWNLOAD BOOT 4.7.5.427</td>
<td>invalid</td>
<td></td>
</tr>
<tr>
<td>linecard 0</td>
<td>DOWNLOAD BOOT 4.7.5.427</td>
<td>invalid</td>
<td></td>
</tr>
<tr>
<td>linecard 1 is not present.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>linecard 2 is not present.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>linecard 3 is not present.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>linecard 4</td>
<td>DOWNLOAD BOOT 4.7.5.427</td>
<td>6.5.1.8</td>
<td></td>
</tr>
<tr>
<td>linecard 5 is not present.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 the example below.

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
Process Restartability

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 attempting to restart the crashed process on primary RPM before executing the failover procedure as a last resort.

Currently, if a software exception occurs, FTOS executes a failover procedure. In a single-RPM system, the system generates a coredump and reboots; in a dual-RPM system, the system generates a coredump and fails over to the standby RPM.

With a system reload, the system must read and apply the entire startup-config file, which might take some time if the startup-config is large. Restarting a process saves time because only a portion of the configuration related to the crashed process is read and re-applied.

For a dual-RPMs system, restarting a process also precludes launching the failover process on the primary and standby RPMs. Recovery is attempted first locally on the primary RPM, which involves less CPU overhead, increasing the systems availability for other activities.

However, in both single and dual-RPM systems, even when Process Restart is configured, the coredump portion of failover is still executed.

The processes that can be restarted fall under three categories:

- **Interface-related processes**—TACACS+, RADIUS, CLI, and SSH, etc.
- **Protocol tasks**—OSPF, RIP, and ACL, etc. Process Restart is not currently available for protocol tasks; the failover procedure is executed immediately upon software exception.
- **Line card processes**—IPC, Event Log Agent, Line Card Manager, etc. Process Restart is not currently available for line card processes; the failover procedure is executed immediately upon software exception.
The restart time varies by process. In general, interface-related processes are hitless and can be restarted in seconds; if a restart is successful, traffic is not interrupted. Protocol tasks and line card processes are not hitless and take longer to restart. You can select which process may attempt to restart and the number of consecutive restart attempts before failover, but by default, every process fails over.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable Process Restartability for a process or task.</td>
<td><code>process restartable [process] [try number] [timestamp hours]</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Display the processes and tasks configured for restart.</td>
<td><code>show process restartable</code></td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

When a process restarts, FTOS displays **Message 9**.

**Message 9  System Message for Process Restarts**

```
```

Customers can specify the timestamp in hour(s) so that if the number of attempts to restart exceeds the max allowed within this timestamp, the restart mode will be changed into failover mode from this moment on. Meaning the next time the crashed process will NOT be restarted but failover to the standby RPM if it is on a dual RPM environment and rebooted if it is on a single RPM.
Internet Group Management Protocol (IGMP)

IGMP is supported on platforms: E C S Z S4810

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 S25/S50, 95 interfaces on the S4810, S55, and S60 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 the following illustration.

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 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). 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 the 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 are still sent to the all-systems address 224.0.0.1 as shown in the illustration below, but reports are sent to the all IGMP version 3-capable multicast routers address 244.0.0.22 as shown in the second illustration.
Joining and Filtering Groups and Sources

The following illustration 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. 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.
Membership Reports: Joining and Filtering

Leaving and Staying in Groups

The illustration below 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 the following illustration, 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.
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
- Selecting an IGMP Version
- Viewing IGMP Groups
- Adjusting Timers
- Configuring a Static IGMP Group
- Prevent a Host from Joining a Group
- Enabling IGMP Immediate-leave
- IGMP Snooping
- Fast Convergence after MSTP Topology Changes
- Designating a Multicast Router Interface
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` in the EXEC Privilege mode.

```
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 the following example.

```
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.

```
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 the example in Viewing IGMP Enabled Interfaces.

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 (refer to the illustration in IGMP version 2). 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.

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 the example in Viewing IGMP Groups.

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 the example in Selecting an IGMP Version.
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.

If IGMP snooping is enabled on a VLT unit, the IGMP snooping dynamically learned groups and multicast router ports are made to learn on the peer by explicitly tunneling the received IGMP control packets.

IGMP Snooping Implementation Information

- IGMP Snooping on FTOS uses IP multicast addresses not MAC addresses.
- 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 igmp` from CONFIGURATION mode, as shown in the following example. You can disable snooping on for a VLAN using the command `no ip igmp snooping` from INTERFACE VLAN mode.

There is no specific configuration needed for IGMP Snooping in conjunction with VLT.

```
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
- Disabling Multicast Flooding
- Specifying a Port as Connected to a Multicast Router
- Configuring the Switch as Querier
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 the example below.

```
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.
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: $E\ C\ S\ Z\ \{S4810\}$.

SONET interfaces are only supported on platform $E$.

Basic Interface Configuration:

- Interface Types
- View Basic Interface Information
- Enable a Physical Interface
- Physical Interfaces
- Management Interfaces
- VLAN Interfaces
- Loopback Interfaces
- Null Interfaces
- Port Channel Interfaces

Advanced Interface Configuration:

- Bulk Configuration
- Interface Range Macros
- Monitor and Maintain Interfaces
- Splitting QSFP ports to SFP+ ports
- Link Debounce Timer
- Link Dampening
- Link Bundle Monitoring
- Ethernet Pause Frames
- Configure MTU Size on an Interface
- Port-pipes
- Auto-Negotiation on Ethernet Interfaces
- View Advanced Interface Information
### Interface Types

<table>
<thead>
<tr>
<th>Interface Type</th>
<th>Modes Possible</th>
<th>Default Mode</th>
<th>Requires Creation</th>
<th>Default State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical</td>
<td>L2, L3</td>
<td>Unset</td>
<td>No</td>
<td>Shutdown (disabled)</td>
</tr>
<tr>
<td>Management</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No Shutdown (enabled)</td>
</tr>
<tr>
<td>Loopback</td>
<td>L3</td>
<td>L3</td>
<td>Yes</td>
<td>No Shutdown (enabled)</td>
</tr>
<tr>
<td>Null</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>Enabled</td>
</tr>
<tr>
<td>Port Channel</td>
<td>L2, L3</td>
<td>L3</td>
<td>Yes</td>
<td>Shutdown (disabled)</td>
</tr>
<tr>
<td>VLAN</td>
<td>L2, L3</td>
<td>L2</td>
<td>Yes (except default)</td>
<td>L2 - No Shutdown (enabled) L3 - Shutdown (disabled)</td>
</tr>
</tbody>
</table>

**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:** The CLI output may be incorrectly displayed as 0 (zero) for the Rx/Tx power values. Perform an SNMP query to obtain the correct power information.

### 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.

The following example displays the configuration and status information for one interface.

```plaintext
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
  0 runts, 0 giants, 0 throttles
  0 CRC, 0 overrun, 0 discarded

Output Statistics:
  3 packets, 192 bytes, 0 underruns
  3 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, 3 Broadcasts, 0 Unicasts
  0 Vlans, 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: 00:00:31

FTOS#

Use the `show ip interfaces brief` command in the EXEC Privilege mode to view which interfaces are enabled for Layer 3 data transmission. In the following example, 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.

FTOS# show ip interface brief

<table>
<thead>
<tr>
<th>Interface</th>
<th>IP-Address</th>
<th>OK? Method Status</th>
<th>Protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>GigabitEthernet 1/0</td>
<td>unassigned</td>
<td>NO Manual administratively down down</td>
<td></td>
</tr>
<tr>
<td>GigabitEthernet 1/1</td>
<td>unassigned</td>
<td>NO Manual administratively down down</td>
<td></td>
</tr>
<tr>
<td>GigabitEthernet 1/2</td>
<td>unassigned</td>
<td>YES Manual up</td>
<td>up</td>
</tr>
<tr>
<td>GigabitEthernet 1/3</td>
<td>unassigned</td>
<td>YES Manual up</td>
<td>up</td>
</tr>
<tr>
<td>GigabitEthernet 1/4</td>
<td>unassigned</td>
<td>YES Manual up</td>
<td>up</td>
</tr>
<tr>
<td>GigabitEthernet 1/5</td>
<td>10.10.10.1</td>
<td>YES Manual up</td>
<td>up</td>
</tr>
<tr>
<td>GigabitEthernet 1/6</td>
<td>unassigned</td>
<td>NO Manual administratively down down</td>
<td></td>
</tr>
<tr>
<td>GigabitEthernet 1/7</td>
<td>unassigned</td>
<td>NO Manual administratively down down</td>
<td></td>
</tr>
<tr>
<td>GigabitEthernet 1/8</td>
<td>unassigned</td>
<td>NO Manual administratively down down</td>
<td></td>
</tr>
</tbody>
</table>

Use the `show interfaces configured` command in the EXEC Privilege mode to view only configured interfaces. In the previous example, 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.

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 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.  
- For a 40-Gigabit Ethernet interface, enter the keyword fortyGigE 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 on each unit of the S4810. It provides dedicated management access to the system. The other S-Series (non-S4810) 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 and for more information on port channels, see Port Channel Interfaces.

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
- Configure Layer 2 (Data Link) Mode
- Management Interfaces
- Auto-Negotiation on Ethernet Interfaces
- Adjust the keepalive timer
- Clear interface counters

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 23-1. Interfaces Types

<table>
<thead>
<tr>
<th>Type of Interface</th>
<th>Possible Modes</th>
<th>Requires Creation</th>
<th>Default State</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/100/1000 Ethernet, Gigabit Ethernet, 10 Gigabit Ethernet</td>
<td>Layer 2, Layer 3</td>
<td>No</td>
<td>Shutdown (disabled)</td>
</tr>
<tr>
<td>SONET (PPP encapsulation)</td>
<td>Layer 3</td>
<td>No</td>
<td>Shutdown (disabled)</td>
</tr>
<tr>
<td>Management</td>
<td>n/a</td>
<td>No</td>
<td>Shutdown (disabled)</td>
</tr>
<tr>
<td>Loopback</td>
<td>Layer 3</td>
<td>Yes</td>
<td>No shutdown (enabled)</td>
</tr>
<tr>
<td>Null interface</td>
<td>n/a</td>
<td>No</td>
<td>Enabled</td>
</tr>
<tr>
<td>Port Channel</td>
<td>Layer 2, Layer 3</td>
<td>Yes</td>
<td>Shutdown (disabled)</td>
</tr>
<tr>
<td>VLAN</td>
<td>Layer 2, Layer 3</td>
<td>Yes, except for the default VLAN</td>
<td>No shutdown (active for Layer 2), Shutdown (disabled for Layer 3)</td>
</tr>
</tbody>
</table>

### 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 cannot configure switching or Layer 2 protocols such as spanning tree protocol on an interface unless the interface has been set to Layer 2 mode.

The following example displays the basic configuration found in a Layer 2 interface.

```plaintext
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>no shutdown</td>
<td>INTERFACE</td>
<td>Enable the interface.</td>
</tr>
<tr>
<td>switchport</td>
<td>INTERFACE</td>
<td>Place the interface in Layer 2 (switching) mode.</td>
</tr>
</tbody>
</table>

For information on enabling and configuring Spanning Tree Protocol, refer to Spanning Tree Protocol (STP). 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).

The example below shows how the `show config` command displays an 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. In the example below, 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.

```
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)#
```

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>no shutdown</code></td>
<td><code>INTERFACE</code></td>
<td>Enable the interface.</td>
</tr>
<tr>
<td><code>ip address ip-address mask [secondary]</code></td>
<td><code>INTERFACE</code></td>
<td>Configure a primary IP address and mask on the interface. The <code>ip-address</code> must be in dotted-decimal format (A.B.C.D) and the <code>mask</code> must be in slash format (/xx). Add the keyword <code>secondary</code> if the IP address is the interface’s backup IP address.</td>
</tr>
</tbody>
</table>
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 as shown in View Basic Interface Information.

To view IP information on an interface in Layer 3 mode, use the show ip interface command in the EXEC Privilege mode as shown in the example below.

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

The S4810 system supports the Management Ethernet interface as well as the standard S-Series interface on any port. Either method can be used to connect to the system.

Configure Management Interfaces on the E-Series, C-Series and the S4810

On the E-Series, C-Series, and S4810 the dedicated Management interface 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.

Note: On the S4810, a default IP address is assigned to the Management port. Use this IP address to set your laptop Ethernet port to the same network for test purposes.
To configure a Management interface, use the following command in the CONFIGURATION mode:

```
interface Managementethernet interface
```

Command Mode | Purpose
---|---
CONFIGURATION | Enter the slot and the port (0). ON the E-Series and C-Series, dual RPMs can be in use. Slot range:
\* C-Series, E-Series: 0-1
\* S4810: 0

To view the Primary RPM Management port, use the `show interface Managementethernet` command in the EXEC Privilege mode. If there are 2 RPMs, you cannot view information on that interface.

To configure IP addresses on a Management interface, use the following command in the MANAGEMENT INTERFACE mode:

```
ip address ip-address mask
```

Command Mode | Purpose
---|---
INTERFACE | Configure an IP address and mask on the interface.
\* `ip-address mask`: enter an address in dotted-decimal format (A.B.C.D), the mask must be in /prefix format (/x)

If there are 2 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. 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 recognize the virtual-IP address.

**Important Things to Remember — virtual-ip**

- `virtual-ip` is a CONFIGURATION mode command.
- When applied, the management port on the primary RPM assumes the virtual IP address. Executing `show interfaces` and `show ip interface brief` commands on the primary RPM management interface will display the virtual IP address and not the actual IP address assigned on that interface.
- 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 the following example, 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.

```
FTOS#show int gig 0/48
GigabitEthernet 0/48 is up, line protocol is up
Description: This is the Management 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

<table>
<thead>
<tr>
<th>Destination</th>
<th>Gateway</th>
<th>Dist/Metric</th>
<th>Last Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>*S 0.0.0.0/0</td>
<td>via 10.11.131.254, Gi 0/48</td>
<td>1/0</td>
<td>1d2h</td>
</tr>
<tr>
<td>C 10.11.130.0/23</td>
<td>Direct, Gi 0/48</td>
<td>0/0</td>
<td>1d2h</td>
</tr>
</tbody>
</table>

FTOS#
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 Layer 2 and Virtual LANs (VLAN)

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.

Note: Egress rate shaping and ingress rate policing cannot be simultaneously used on the same VLAN.

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip address ip-address mask [secondary]</td>
<td>INTERFACE</td>
<td>Configure an IP address and mask on the interface.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ip-address mask: enter an address in dotted-decimal format (A.B.C.D) and the mask must be in slash format (/24).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• secondary: the IP address is the interface’s backup IP address. You can configure up to eight secondary IP addresses.</td>
</tr>
</tbody>
</table>

The following example shows a sample configuration of a VLAN participating in an OSPF process.

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>interface loopback number</td>
<td>CONFIGURATION</td>
<td>Enter a number as the loopback interface.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range: 0 to 16383.</td>
</tr>
</tbody>
</table>

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.

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>interface null 0</td>
<td>CONFIGURATION</td>
<td>Enter the INTERFACE mode of the Null interface.</td>
</tr>
</tbody>
</table>

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 Link Aggregation Control Protocol (LACP).

<table>
<thead>
<tr>
<th>Platform</th>
<th>Port-channels</th>
<th>Members/Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-Series TeraScale</td>
<td>255</td>
<td>16</td>
</tr>
<tr>
<td>E-Series ExaScale</td>
<td>512</td>
<td>64</td>
</tr>
<tr>
<td>C-Series</td>
<td>128</td>
<td>8</td>
</tr>
<tr>
<td>S-Series: S25 and S50</td>
<td>52</td>
<td>8</td>
</tr>
<tr>
<td>S55, S60 and S4810</td>
<td>128</td>
<td>8</td>
</tr>
<tr>
<td>Z9000</td>
<td>128</td>
<td>8</td>
</tr>
</tbody>
</table>
**Note:** If you are using either 10G ports or 40G ports, the Z9000 supports 8 members per LAG

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 not 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 (mandatory)
- Reassign an interface to a new port channel (optional)
- Configure the minimum oper up links in a port channel (LAG) (optional)
- Add or remove a port channel from a VLAN (optional)
- Assign an IP address to a port channel (optional)
- Delete or disable a port channel (optional)
- Load balancing through port channels (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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interface port-channel id-number</td>
<td>CONFIGURATION</td>
<td>Create a port channel.</td>
</tr>
<tr>
<td>2</td>
<td>no shutdown</td>
<td>INTERFACE</td>
<td>Ensure that the port channel is active.</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>channel-member interface</td>
<td>INTERFACE PORT-CHANNEL</td>
<td>Add the interface to a port channel. The <code>interface</code> variable is the physical interface type and slot/port information.</td>
</tr>
<tr>
<td>2</td>
<td>show config</td>
<td>INTERFACE PORT-CHANNEL</td>
<td>Double check that the interface was added to the port channel.</td>
</tr>
</tbody>
</table>

To view the port channel’s status and channel members in a tabular format, use the `show interfaces port-channel brief` command in the EXEC Privilege mode, as shown in the following 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#
```

The example below 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.

```
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 the following example 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.

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)#
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>no channel-member interface</td>
<td>INTERFACE PORT-CHANNEL</td>
<td>Remove the interface from the first port channel.</td>
</tr>
<tr>
<td>2</td>
<td>interface port-channel id number</td>
<td>INTERFACE PORT-CHANNEL</td>
<td>Change to the second port channel INTERFACE mode.</td>
</tr>
<tr>
<td>3</td>
<td>channel-member interface</td>
<td>INTERFACE PORT-CHANNEL</td>
<td>Add the interface to the second port channel.</td>
</tr>
</tbody>
</table>

The following text displays an example of moving the GigabitEthernet 1/8 interface from port channel 4 to port channel 3.

```
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimum-links number</td>
<td>INTERFACE</td>
<td>Enter the number of links in a LAG that must be in “oper up” status.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default: 1</td>
</tr>
</tbody>
</table>

The following text displays an example of configuring five minimum “oper up” links in a port channel.

```
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>tagged port-channel id number</td>
<td>INTERFACE</td>
<td>Add the port channel to the VLAN as a tagged interface. An interface with tagging enabled can belong to multiple VLANs.</td>
</tr>
<tr>
<td></td>
<td>VLAN</td>
<td></td>
</tr>
<tr>
<td>untagged port-channel id number</td>
<td>INTERFACE</td>
<td>Add the port channel to the VLAN as an untagged interface. An interface without tagging enabled can belong to only one VLAN.</td>
</tr>
<tr>
<td></td>
<td>VLAN</td>
<td></td>
</tr>
</tbody>
</table>

To remove a port channel from a VLAN, use either of the following commands:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>no tagged port-channel id number</td>
<td>INTERFACE</td>
<td>Remove the port channel with tagging enabled from the VLAN.</td>
</tr>
<tr>
<td></td>
<td>VLAN</td>
<td></td>
</tr>
<tr>
<td>no untagged port-channel id number</td>
<td>INTERFACE</td>
<td>Remove the port channel without tagging enabled from the VLAN.</td>
</tr>
<tr>
<td></td>
<td>VLAN</td>
<td></td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip address ip-address mask [secondary]</td>
<td>INTERFACE</td>
<td>Configure an IP address and mask on the interface.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ip-address mask: enter an address in dotted-decimal format (A.B.C.D) and the mask must be in slash format (/24).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• secondary: the IP address is the interface’s backup IP address. You can configure up to eight secondary IP addresses.</td>
</tr>
</tbody>
</table>

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
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 23-3. Hash Methods as Applied to Port Channel Types

<table>
<thead>
<tr>
<th>Hash (Header Based)</th>
<th>Layer 2 Port Channel</th>
<th>Layer 3 Port Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-tuple</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>3-tuple</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Packet-based</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>MAC source address (SA) and destination address (DA)</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

On the E-Series, to change the 5-tuple default to 3-tuple, MAC, or packet-based, use the following command in CONFIGURATION mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| [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 23-1 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 23-4. 5-tuple and 3-tuple Keys

<table>
<thead>
<tr>
<th>Keys</th>
<th>5-tuple</th>
<th>3-tuple</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP source address (lower 32 bits)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>IP destination address (lower 32 bits)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Protocol type</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
The following example shows the configuration and show command for packet-based hashing on the E-Series.

```
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 23-1.

### I Pv 4, 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.

#### Figure 23-1. The load-balance Commands and Port Channel Types

<table>
<thead>
<tr>
<th>Configuration Commands</th>
<th>Switched IP Traffic</th>
<th>Routed IP Traffic (IPv4 only)</th>
<th>Switched Non-IP Traffic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default (IP 5-tuple)</td>
<td>IP 5-tuple (lower 32 bits)</td>
<td>IP 5-tuple</td>
<td>MAC-based</td>
</tr>
<tr>
<td>load-balance ip-selection 3-tuple</td>
<td>IP 3-tuple (lower 32 bits)</td>
<td>IP 3-tuple</td>
<td>MAC-based</td>
</tr>
<tr>
<td>load-balance ip-selection mac</td>
<td>MAC-based</td>
<td>IP 5-tuple</td>
<td>MAC-based</td>
</tr>
<tr>
<td>load-balance ip-selection 3-tuple</td>
<td>MAC-based</td>
<td>IP 3-tuple</td>
<td>MAC-based</td>
</tr>
<tr>
<td>load-balance ip-selection packet-based</td>
<td>MAC-based</td>
<td>Packet-based</td>
<td>MAC-based</td>
</tr>
</tbody>
</table>

Note: For IPV6, only the first 32 bits (LSB) of IP Source Address and IP Destination Address are used for hash generation.

Table 23-4. 5-tuple and 3-tuple Keys

<table>
<thead>
<tr>
<th>Keys</th>
<th>5-tuple</th>
<th>3-tuple</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCP/UDP source port</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TCP/UDP destination port</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[no] load-balance {ip-selection</td>
<td>CONFIGURATION</td>
<td>Replace the default IP 4-tuple method of balancing traffic over a port</td>
</tr>
<tr>
<td>dest-ip</td>
<td>source-ip}</td>
<td>ip-selection {dest-ip</td>
</tr>
<tr>
<td>mac [dest-mac</td>
<td>source-dest-mac</td>
<td>source-mac] {tcp-udp enable}</td>
</tr>
<tr>
<td>ing-port}</td>
<td>tcp-udp enable —Distribute traffic based on TCP/UDP source and destination ports.</td>
<td></td>
</tr>
<tr>
<td>ing-port}</td>
<td>ing-port —Distribute traffic based on the port ID of the IP source address.</td>
<td></td>
</tr>
</tbody>
</table>

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).

The following example shows a sample configuration for the `hash-algorithm` command.

```
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:

```
hash-algorithm ecmp {crc-upper} | {dest-ip} | {lsb}
```

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:

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

FTOS(config)# interface range gigabitethernet 5/1 - 23
FTOS(config-if-range-gi-5/1-23)# no shutdown
FTOS(config-if-range-gi-5/1-23)#

Create a multiple-range

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:

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:

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:

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)#

Add ranges

The example below shows how to use commas to add VLAN and port-channel interfaces to the range.
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:

```
FTOS(config)# define interface-range macro_name {vlan vlan_ID - vlan_ID} |
                   {gigabitethernet | tengigabitethernet | fortyGigE} slot/interface - interface} |
                   [ , {vlan vlan_ID - vlan_ID} |
                   {gigabitethernet | tengigabitethernet | fortyGigE} |
                   slot/interface - interface}]
```

**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
```

**Choose an Interface-range Macro**

To use an interface-range macro in the interface range command, enter this command:

```
interface range macro name
```

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>monitor interface interface</code></td>
<td>EXEC Privilege</td>
<td>View the interface’s statistics. Enter the type of interface and slot/port information:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10/100/1000 Ethernet interface, enter the keyword <code>GigabitEthernet</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a Gigabit Ethernet interface, enter the keyword <code>GigabitEthernet</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For the Management interface on the RPM, enter the keyword <code>ManagementEthernet</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a SONET interface, enter the keyword <code>sonet</code> followed by slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10 Gigabit Ethernet interface, enter the keyword <code>TenGigabitEthernet</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 40-Gigabit Ethernet interface, enter the keyword <code>fortyGigE</code> followed by the slot/port information.</td>
</tr>
</tbody>
</table>

The information displays in a continuous run, refreshing every 2 seconds by default as shown in the example below. 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

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

<table>
<thead>
<tr>
<th>Traffic statistics:</th>
<th>Current</th>
<th>Rate</th>
<th>Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input bytes:</td>
<td>0</td>
<td>0 Bps</td>
<td>0</td>
</tr>
<tr>
<td>Output bytes:</td>
<td>0</td>
<td>0 Bps</td>
<td>0</td>
</tr>
<tr>
<td>Input packets:</td>
<td>0</td>
<td>0 pps</td>
<td>0</td>
</tr>
<tr>
<td>Output packets:</td>
<td>0</td>
<td>0 pps</td>
<td>0</td>
</tr>
<tr>
<td>64B packets:</td>
<td>0</td>
<td>0 pps</td>
<td>0</td>
</tr>
<tr>
<td>Over 64B packets:</td>
<td>0</td>
<td>0 pps</td>
<td>0</td>
</tr>
<tr>
<td>Over 127B packets:</td>
<td>0</td>
<td>0 pps</td>
<td>0</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>tdr-cable-test gigabitethernet &lt;slot&gt;/&lt;port&gt;</code></td>
<td>EXEC Privilege</td>
<td>To test for cable faults on the GigabitEthernet cable.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Between two ports, the user must not start the test on both ends of the cable.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• The user must enable the interface before starting the test.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• The port should be enabled to run the test or the test prints an error message.</td>
</tr>
<tr>
<td>2</td>
<td><code>show tdr gigabitethernet &lt;slot&gt;/&lt;port&gt;</code></td>
<td>EXEC Privilege</td>
<td>Displays TDR test results.</td>
</tr>
</tbody>
</table>

### Splitting QSFP ports to SFP+ ports

Splitting QSFP ports to SFP+ ports is supported on platforms: [S4810](#) and [Z9000](#).

The S4810 and Z9000 platforms support splitting a single 40G QSFP port into four 10G SFP+ ports using one of the supported breakout cables (refer to the [Installation Guide](#) or the [Release Notes](#) for a list of supported cables).

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>stack-unit stack-unit port number portmode quad</code></td>
<td>CONFIGURATION</td>
<td>Split a single 40G port into 4-10G ports on the S4810 or Z9000.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>stack-unit</code>: Enter the stack member unit identifier of the stack member to reset.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range: 0 to 11</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>number</code>: Enter the port number of the 40G port to be split.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S4810 Range: 0 to 47 for 10G ports; 48, 52, 56 and 60 for 40G ports</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Z9000 Range: 0 to 31</td>
</tr>
</tbody>
</table>

### Important Points

- Splitting a 40G port into 4x10G port is supported only on a standalone unit.
- Split ports cannot be used as stack-link to stack an S4810 or a Z9000.
- Split ports cannot be a part of any stacked system.
- The unit number with the split ports must be the default (stack-unit 0)

This can be verified using the `show system brief` command. If the unit ID is different than 0, then it must be renumbered to 0 before ports are split by using the `stackunit id renumber 0` command in EXEC mode.

- The quad port must be in a default configuration before it can be split into 4x10G ports. The 40G port is lost in the config when the port is split, so be sure the port is also removed from other L2/L3 feature configurations.
- The system must be reloaded after issuing the CLI for the change to take effect.
## Link Debounce Timer

Link Debounce Timer is supported on platform

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

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 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                                         |

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.

Note: FTOS rounds the entered debounce time up to the nearest hundredth.
Note in the example in Assign a debounce time to an interface that the timer was set at 150 ms, but appears as 200 in the example in Show debounce times in an interface above.

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 the following example.

```
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 the following example.

```
FTOS# show interfaces dampening
InterfaceStateFlapsPenaltyHalf-LifeReuseSuppressMax-Sup
Gi 0/0Up005750250020
Gi 0/1Up21200205001500300
Gi 0/2Down4850306002000120
```

View a dampening summary for the entire system using the command `show interfaces dampening summary` from EXEC Privilege mode, as shown in the example below.

```
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 the following example.

```
FTOS# clear dampening interface Gi 0/1

FTOS# show interfaces dampening GigabitEthernet0/0
InterfaceStateFlapsPenaltyHalf-LifeReuseSuppressMax-Sup
Gi 0/0Up00205001500300
```

**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 12000 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 planning MTU sizes across a network.

Table 23-5 lists the range for each transmission media.

<table>
<thead>
<tr>
<th>Transmission Media</th>
<th>MTU Range (in bytes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethernet</td>
<td>594-12000 = link MTU</td>
</tr>
<tr>
<td></td>
<td>576-9234 = IP MTU</td>
</tr>
</tbody>
</table>

Link Bundle Monitoring

**Link Bundle Monitoring** is supported only on platform: S4810

Monitoring linked LAG bundles allows traffic distribution amounts in a link to be monitored for unfair distribution at any given time. A threshold of 60% is defined as an acceptable amount of traffic on a member link. Links are monitored in 15-second intervals for three consecutive instances. Any deviation within that time causes a Syslog to be sent and an alarm event to be generated. When the deviation clears, another Syslog is sent and a clear alarm event is generated.

The link bundle utilization is calculated as the total bandwidth of all links divided by the total bytes-per-second of all links. If monitoring is enabled, the utilization calculation is performed when the utilization of the link-bundle (not a link within a bundle) exceeds 60%.

Enable link bundle monitoring using the ecmp-group command.

View all LAG link bundles being monitored using the show running-config ecmp-group command.

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 (non-S4810) platforms, Ethernet Pause Frames TX should be enabled only after consulting with the Dell Force10 Technical Assistance Center.

**Note:** Changes in the flow-control values may not be reflected automatically in the `show interface` output. 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.

**Note:** The S4810 supports only the rx control option. The S4810 does not transmit pause frames.

**Note:** If rx flow control is disabled, Dell Force10 recommends rebooting the system.

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>flowcontrol rx [off</td>
<td>on] tx [off</td>
<td>on] [threshold {&lt;1-2047&gt; &lt;1-2013&gt; &lt;1-2013&gt;}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Defaults:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-Series: rx off tx off</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E-Series: rx on tx on</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S-Series: rx off tx off</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S4810: rx off</td>
</tr>
</tbody>
</table>

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 enough to include the Layer 2 header. For example, for VLAN packets, if the IP MTU is 1400, the Link MTU must be no less than 1422:

\[
1400\text{-byte IP MTU} + 22\text{-byte VLAN Tag} = 1422\text{-byte link MTU}
\]

The MTU range is 592-12000, with a default of 1500. On the E-Series, the user must enter the `ip mtu` command to manually configure the IP MTU to compensate for the Layer 2 header. The C-Series and S-Series automatically configure the IP MTU.

Table 23-6 lists the various Layer 2 overheads found in FTOS and the number of bytes.

**Table 23-6. Difference between Link MTU and IP MTU**

<table>
<thead>
<tr>
<th>Layer 2 Overhead</th>
<th>Difference between Link MTU and IP MTU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethernet (untagged)</td>
<td>18 bytes</td>
</tr>
<tr>
<td>VLAN Tag</td>
<td>22 bytes</td>
</tr>
<tr>
<td>Untagged Packet with VLAN-Stack Header</td>
<td>22 bytes</td>
</tr>
<tr>
<td>Tagged Packet with VLAN-Stack Header</td>
<td>26 bytes</td>
</tr>
</tbody>
</table>

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 23-7 presents these platform differences again.

**Table 23-7. Platform Differences Concerning Port-pipes**

<table>
<thead>
<tr>
<th>Chassis Type</th>
<th>Port-pipes / Slot</th>
<th>Channels / Port-pipe</th>
<th>Capacity of Each Channel (Gbps)</th>
<th>Raw Slot Capacity (Gbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1200/E1200i-AC/DC</td>
<td>2</td>
<td>9</td>
<td>3.125</td>
<td>56.25</td>
</tr>
<tr>
<td>E600/E600i</td>
<td>2</td>
<td>9</td>
<td>3.125</td>
<td>56.25</td>
</tr>
<tr>
<td>E300</td>
<td>1</td>
<td>8</td>
<td>3.125</td>
<td>25</td>
</tr>
</tbody>
</table>
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. In FTOS, the command speed 100 is an exact equivalent of speed auto 100 in IOS.

To discover whether the remote and local interface require manual speed synchronization, and to manually synchronize them if necessary, use the following command sequence as shown in the second example:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Determine the local interface status. Refer to the example below.</td>
<td>show interfaces [interface</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td></td>
<td></td>
<td>linecard slot-number] status</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Determine the remote interface status. Use the command on the remote system that is equivalent to the above command.</td>
<td>EXEC EXEC Privilege</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Access CONFIGURATION mode.</td>
<td>config</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>4</td>
<td>Access the port.</td>
<td>interface interface slot/port</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>5</td>
<td>Set the local port speed. speed {10</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>6</td>
<td>Optionally, set full- or half-duplex.</td>
<td>duplex {half</td>
<td>full}</td>
</tr>
<tr>
<td>7</td>
<td>Disable auto-negotiation on the port. If the speed was set to 1000, auto-negotiation does not need to be disabled.</td>
<td>no negotiation auto</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>8</td>
<td>Verify configuration changes.</td>
<td>show config</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>
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].

<table>
<thead>
<tr>
<th>Port</th>
<th>Description</th>
<th>Status</th>
<th>Speed</th>
<th>Duplex</th>
<th>Vlan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gi 0/0</td>
<td></td>
<td>Up</td>
<td>1000 Mbit</td>
<td>Auto</td>
<td>--</td>
</tr>
<tr>
<td>Gi 0/1</td>
<td></td>
<td>Down</td>
<td></td>
<td>Auto</td>
<td>1</td>
</tr>
<tr>
<td>Gi 0/2</td>
<td></td>
<td>Down</td>
<td>Auto</td>
<td>Auto</td>
<td>1</td>
</tr>
<tr>
<td>Gi 0/3</td>
<td></td>
<td>Down</td>
<td>Auto</td>
<td>Auto</td>
<td>--</td>
</tr>
<tr>
<td>Gi 0/4</td>
<td>Force10Port</td>
<td>Up</td>
<td>1000 Mbit</td>
<td>Auto</td>
<td>30-130</td>
</tr>
<tr>
<td>Gi 0/5</td>
<td></td>
<td>Down</td>
<td>Auto</td>
<td>Auto</td>
<td>--</td>
</tr>
<tr>
<td>Gi 0/6</td>
<td></td>
<td>Down</td>
<td>Auto</td>
<td>Auto</td>
<td>--</td>
</tr>
<tr>
<td>Gi 0/7</td>
<td></td>
<td>Up</td>
<td>1000 Mbit</td>
<td>Auto</td>
<td>1502,1504,1506-1508,1602</td>
</tr>
<tr>
<td>Gi 0/8</td>
<td></td>
<td>Down</td>
<td>Auto</td>
<td>Auto</td>
<td>--</td>
</tr>
<tr>
<td>Gi 0/9</td>
<td></td>
<td>Down</td>
<td>Auto</td>
<td>Auto</td>
<td>--</td>
</tr>
<tr>
<td>Gi 0/10</td>
<td></td>
<td>Down</td>
<td>Auto</td>
<td>Auto</td>
<td>--</td>
</tr>
<tr>
<td>Gi 0/11</td>
<td></td>
<td>Down</td>
<td>Auto</td>
<td>Auto</td>
<td>--</td>
</tr>
<tr>
<td>Gi 0/12</td>
<td></td>
<td>Down</td>
<td>Auto</td>
<td>Auto</td>
<td>--</td>
</tr>
</tbody>
</table>

In the example above, several ports display “Auto” in the Speed field, including port 0/1. In the following example, the speed of port 0/1 is set to 100Mb and then its auto-negotiation is disabled.

```
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.
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
show                    Show autoneg configuration information
FTOS(conf-if-autoneg)#mode ?
forced-master           Force port to master mode
forced-slave            Force port to slave mode
FTOS(conf-if-autoneg)#

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>keepalive [seconds]</td>
<td>INTERFACE</td>
<td>Change the default interval between keepalive messages.</td>
</tr>
</tbody>
</table>

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.

The following example lists the possible show commands that have the configured keyword available:

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 as shown in the example below displays the interface, whether the interface supports IEEE 802.1Q tagging or not, and the VLANs to which the interface belongs.

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:
  Vlan  2

Name: GigabitEthernet 13/3
  802.1QTagged: True
  Vlan membership:
  Vlan  2

--More--

**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.

The following example shows how to configure rate interval when changing the default value:

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

**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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>clear counters [interface]</td>
<td>EXEC Privilege</td>
<td>Clear the counters used in the show interface commands for all VRRP groups, VLANs, and physical interfaces or selected ones.</td>
</tr>
<tr>
<td>[vrrp [vrid]</td>
<td>learning-limit]</td>
<td>Without an interface specified, the command clears all interface counters.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(OPTIONAL) Enter the following interface keywords and slot/port or number information:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 1-Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a Loopback interface, enter the keyword loopback followed by a number from 0 to 16383.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a SONET interface, enter the keyword sonet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10-Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 40-Gigabit Ethernet interface, enter the keyword fortyGigE followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(OPTIONAL) Enter the keyword vrrp to clear statistics for all VRRP groups configured. Enter a number from 1 to 255 as the vrid.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(OPTIONAL) Enter the keyword learning-limit to clear unknown source address (SA) drop counters when MAC learning limit is configured on the interface.</td>
</tr>
</tbody>
</table>

When you enter this command, you must confirm that you want FTOS to clear the interface counters for that interface as shown in the following example.

```
FTOS#clear counters gi 0/0
Clear counters on GigabitEthernet 0/0 [confirm]
FTOS#
```
IPv4 Routing

IPv4 Routing is supported on platforms: E C S Z 54810

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.

- IP Addresses
- Directed Broadcast
- Resolution of Host Names
- ARP
- ICMP
- UDP Helper

Table 24-1 lists the defaults for the IP addressing features described in this chapter.

**Table 24-1. IP Defaults**

<table>
<thead>
<tr>
<th>IP Feature</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNS</td>
<td>Disabled</td>
</tr>
<tr>
<td>Directed Broadcast</td>
<td>Disabled</td>
</tr>
<tr>
<td>Proxy ARP</td>
<td>Enabled</td>
</tr>
<tr>
<td>ICMP Unreachable</td>
<td>Disabled</td>
</tr>
<tr>
<td>ICMP Redirect</td>
<td>Disabled</td>
</tr>
</tbody>
</table>

**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. Superenetting, 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,

0000101011010110010101110000011

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** (mandatory)
- **Configure static routes** (optional)
- **Configure static routes for the management interface** (optional)

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

**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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 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.  
* For a 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` followed by the slot/port information. |
| 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 in the INTERFACE mode as shown in the example below or `show ip interface` in the EXEC privilege mode as shown in the second example.

```
FTOS(conf-if)#show conf
!
interface GigabitEthernet 0/0
  ip address 10.11.1.1/24
  no shutdown
!
FTOS(conf-if)#
```

```
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.

```
FTOS#show ip route static

<table>
<thead>
<tr>
<th>Destination</th>
<th>Gateway</th>
<th>Dist/Metric</th>
<th>Last Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>S 2.1.2.0/24</td>
<td>Direct, Nu 0</td>
<td>0/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.0/24</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.2/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.3/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.4/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.5/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.6/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.7/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.8/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.9/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.10/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.11/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.12/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.13/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.14/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.15/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.16/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.17/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 11.1.1.0/24</td>
<td>Direct, Nu 0</td>
<td>0/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td></td>
<td>Direct, Lo 0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

---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:
To view the configured static routes for the management port, use the `show ip management-route` command in the EXEC privilege mode.

```
FTOS#show ip route static
```

<table>
<thead>
<tr>
<th>Destination</th>
<th>Gateway</th>
<th>Dist/Metric</th>
<th>Last Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>S 2.1.2.0/24</td>
<td>Direct, Nu 0</td>
<td>0/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.0/24</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.2/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.3/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.4/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.5/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.6/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.7/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.8/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.9/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.10/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.11/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.12/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.13/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.14/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.15/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.16/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 6.1.2.17/32</td>
<td>via 6.1.20.2, Te 5/0</td>
<td>1/0</td>
<td>00:02:30</td>
</tr>
<tr>
<td>S 11.1.1.0/24</td>
<td>Direct, Nu 0</td>
<td>0/0</td>
<td>00:02:30</td>
</tr>
</tbody>
</table>

```
```

### 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:

```
ip directed-broadcast
```

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip directed-broadcast</td>
<td>INTERFACE</td>
<td>Enable directed broadcast.</td>
</tr>
</tbody>
</table>

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 command.

- Enable dynamic resolution of host names
- Specify local system domain and a list of domains
- DNS with traceroute

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip domain-lookup</td>
<td>CONFIGURATION</td>
<td>Enable dynamic resolution of host names.</td>
</tr>
<tr>
<td>ip name-server ip-address</td>
<td>CONFIGURATION</td>
<td>Specify up to 6 name servers. The order you entered the</td>
</tr>
<tr>
<td>[ip-address2 ... ip-address6]</td>
<td></td>
<td>servers determines the order of their use.</td>
</tr>
</tbody>
</table>

To view current bindings, use the show hosts command.

```
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip domain-name name</td>
<td>CONFIGURATION</td>
<td>Enter up to 63 characters to configure one domain name for the E-Series.</td>
</tr>
</tbody>
</table>

To configure a list of domain names, use the following command in the CONFIGURATION mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip domain-list name</td>
<td>CONFIGURATION</td>
<td>Enter up to 63 characters to 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.</td>
</tr>
</tbody>
</table>

DNS with traceroute

To configure your switch to perform DNS with traceroute, follow the steps below in the CONFIGURATION mode.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip domain-lookup</td>
<td>CONFIGURATION</td>
<td>Enable dynamic resolution of host names.</td>
</tr>
<tr>
<td>ip name-server ip-address [ip-address2 ... ip-address6]</td>
<td>CONFIGURATION</td>
<td>Specify up to 6 name servers. The order you entered the servers determines the order of their use.</td>
</tr>
<tr>
<td>traceroute [host</td>
<td>ip-address ]</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
The following text is an example output of DNS using the **traceroute** command.

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
------------------------------------------------------------------------------------------
<table>
<thead>
<tr>
<th>TTL</th>
<th>Hostname</th>
<th>Probe1</th>
<th>Probe2</th>
<th>Probe3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.11.199.190</td>
<td>001.000 ms</td>
<td>001.000 ms</td>
<td>002.000 ms</td>
</tr>
<tr>
<td>2</td>
<td>gwegress-sjc-02.force10networks.com</td>
<td>005.000 ms</td>
<td>001.000 ms</td>
<td>001.000 ms</td>
</tr>
<tr>
<td>3</td>
<td>fw-sjc-01.force10networks.com</td>
<td>000.000 ms</td>
<td>000.000 ms</td>
<td>000.000 ms</td>
</tr>
<tr>
<td>4</td>
<td><a href="http://www.force10networks.com">www.force10networks.com</a></td>
<td>000.000 ms</td>
<td>000.000 ms</td>
<td>000.000 ms</td>
</tr>
</tbody>
</table>
FTOS#

**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 (optional)
- Enable Proxy ARP (optional)
- Clear ARP cache (optional)
- ARP Learning via Gratuitous ARP
- ARP Learning via ARP Request
- Configurable ARP Retries

For a complete listing of all ARP-related commands, refer to the *FTOS Command Line Reference*.  

IPv4 Routing | 523
Configure static ARP entries

ARP dynamically maps the MAC and IP addresses, and while most network hosts 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:

```
arp ip-address mac-address interface
```

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 in the EXEC privilege mode as shown below.

```
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:

```
ip 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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>clear arp-cache [interface</td>
<td>ip ip-address]</td>
<td>EXEC privilege</td>
</tr>
<tr>
<td></td>
<td>no-refresh]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 1-Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a SONET interface, enter the keyword sonet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10-Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a VLAN interface, enter the keyword vlan followed by a number between 1 and 4094.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 40-Gigabit Ethernet interface, enter the keyword fortyGigE followed by the slot/port information.</td>
</tr>
<tr>
<td>ip ip-address (OPTIONAL)</td>
<td></td>
<td>Enter the keyword ip followed by the IP address of the ARP entry you wish to clear.</td>
</tr>
<tr>
<td>no-refresh (OPTIONAL)</td>
<td></td>
<td>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.</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable ARP learning via gratuitous ARP.</td>
<td><code>arp learn-enable</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**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.

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.

Whether ARP Learning via Gratuitous ARP 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 default backoff interval remains at 20 seconds. On the S4810 platform, with FTOS version 8.3.8.0 and later, the time between ARP resend is configurable. This timer is an exponential backoff timer. Over the specified period, the time between ARP requests increases. This reduces the potential for the system to slow down while waiting for a multitude of ARP responses.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
</table>
| Set the number of ARP retries. | arp retries number  
Default: 5  
Range: 1-20 | CONFIGURATION |
| Set the an exponential timer for resending unresolved ARPs. | arp backoff-time  
Default: 30  
Range: 1 to 3600 | CONFIGURATION |
| Display all ARP entries learned via gratuitous ARP. | show arp retries | EXEC Privilege |

ICMP

For diagnostics, Internet Control Message Protocol (ICMP) provides 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
- Enable ICMP redirects

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip unreachable</td>
<td>INTERFACE</td>
<td>Set FTOS to create and send ICMP unreachable messages on the interface.</td>
</tr>
</tbody>
</table>

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 [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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip redirect</td>
<td>INTERFACE</td>
<td>Set FTOS to create and send ICMP redirect messages on the interface.</td>
</tr>
</tbody>
</table>

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.

### 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. Refer to Enabling UDP Helper.
2. Configure a broadcast address on interfaces that will receive UDP broadcast traffic. Refer to Configuring a Broadcast Address.

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 UDP helper using the command `ip udp-helper udp-ports`, as shown in the example below.

```
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 the following example.

```
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 the example below.

```plaintext
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 the following example.

```plaintext
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**
- **UDP Helper with Subnet Broadcast Addresses**
- **UDP Helper with Configured Broadcast Addresses**
- **UDP Helper with No Configured Broadcast Addresses**

**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 the illustration below:
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 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.

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 the following illustration, Packet 1 has the destination IP address 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 1.1.255. In this case, it is flooded on VLAN 101 in its original condition as the forwarding process is Layer 2.
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 the following illustration, 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.

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 the example below.

```
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 the following example.

```
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
```
IPv6 Routing

IPv6 Routing is supported on platforms E C S S4810.

Note: The IPv6 basic commands are supported on all platforms. However, not all features are supported on all platforms, nor for all releases. See Table 25-2 to determine the FTOS version supporting which features and platforms.

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
  - Extended Address Space
  - Stateless Autoconfiguration
  - IPv6 Headers
- Implementing IPv6 with FTOS
  - ICMPv6
  - Path MTU Discovery
  - IPv6 Neighbor Discovery
  - QoS for IPv6
  - IPv6 Multicast
  - SSH over an IPv6 Transport
- Configuration Task List for 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
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 IPv6 addresses by using either the MAC address or a private random number to build its unique IPv6 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 redirect 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 shown in Figure 25-1.

![IPv6 Header Fields Diagram](image)

**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 header 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 25-1). 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 25-1. Next Header field values

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Hop-by-Hop option header</td>
</tr>
<tr>
<td>4</td>
<td>IPv4</td>
</tr>
<tr>
<td>6</td>
<td>TCP</td>
</tr>
<tr>
<td>8</td>
<td>Exterior Gateway Protocol (EGP)</td>
</tr>
<tr>
<td>41</td>
<td>IPv6</td>
</tr>
<tr>
<td>43</td>
<td>Routing header</td>
</tr>
<tr>
<td>44</td>
<td>Fragmentation header</td>
</tr>
<tr>
<td>50</td>
<td>Encrypted Security</td>
</tr>
<tr>
<td>51</td>
<td>Authentication header</td>
</tr>
<tr>
<td>59</td>
<td>No Next Header</td>
</tr>
<tr>
<td>60</td>
<td>Destinations option header</td>
</tr>
</tbody>
</table>
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 IPv6 address for the packet originator.

Destination Address (128 bits)

The Destination Address field contains the intended recipient’s IPv6 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 by 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 25-1).

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 25-1).

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 25-1.

- **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::0000:0000:0000:1428:57ab 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::0000:0000:0000:1428:57ab 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 and/or trailing zeros in a group can also be omitted (as in ::1 for localhost, 1:: for network addresses and :: for unspecified addresses).

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 IPv6 addresses are manually assigned to a computer by an administrator. Dynamic IPv6 addresses are assigned either randomly or by a server using Dynamic Host Configuration Protocol (DHCP). Even though IPv6 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 IPv6 addresses. In this
In the case of a DHCP server being used, it is specifically configured to always assign the same IPv6 address to a particular computer and never to assign that IP address to another computer. This allows static IPv6 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 25-2 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 following symbols: C E S

<table>
<thead>
<tr>
<th>Feature and/or Functionality</th>
<th>FTOS Release Introduction</th>
<th>Documentation and Chapter Location</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E-Series TeraScale</td>
<td>E-Series ExaScale</td>
</tr>
<tr>
<td>Basic IPv6 Commands</td>
<td>7.4.1 8.2.1 7.8.1 7.8.1 8.3.10.0</td>
<td>IPv6 Basic Commands in the FTOS Command Line Interface Reference Guide</td>
</tr>
<tr>
<td>IPv6 Basic Addressing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPv6 address types: Unicast</td>
<td>7.4.1 8.2.1 7.8.1 7.8.1 8.3.10.0</td>
<td>Extended Address Space in this chapter</td>
</tr>
<tr>
<td>IPv6 neighbor discovery</td>
<td>7.4.1 8.2.1 7.8.1 7.8.1 8.3.10.0</td>
<td>IPv6 Neighbor Discovery in this chapter</td>
</tr>
<tr>
<td>IPv6 stateless autoconfiguration</td>
<td>7.4.1 8.2.1 7.8.1 7.8.1 8.3.10.0</td>
<td>Stateless Autoconfiguration in this chapter</td>
</tr>
<tr>
<td>IPv6 MTU path discovery</td>
<td>7.4.1 8.2.1 7.8.1 7.8.1 8.3.10.0</td>
<td>Path MTU Discovery in this chapter</td>
</tr>
<tr>
<td>IPv6 ICMPv6</td>
<td>7.4.1 8.2.1 7.8.1 7.8.1 8.3.10.0</td>
<td>ICMPv6 in this chapter</td>
</tr>
<tr>
<td>IPv6 ping</td>
<td>7.4.1 8.2.1 7.8.1 7.8.1 8.3.10.0</td>
<td>ICMPv6 in this chapter</td>
</tr>
<tr>
<td>IPv6 traceroute</td>
<td>7.4.1 8.2.1 7.8.1 7.8.1 8.3.10.0</td>
<td>ICMPv6 in this chapter</td>
</tr>
<tr>
<td>IPv6 Routing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Static routing</td>
<td>7.4.1 8.2.1 7.8.1 7.8.1 8.3.10.0</td>
<td>Assign a Static IPv6 Route in this chapter</td>
</tr>
</tbody>
</table>
### Table 25-2. FTOS and IPv6 Feature Support (continued)

<table>
<thead>
<tr>
<th>Feature</th>
<th>FTOS 7.4.1</th>
<th>FTOS 8.2.1</th>
<th>FTOS 7.8.1</th>
<th>FTOS 8.4.2</th>
<th>FTOS 8.3.10.0</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Route redistribution</td>
<td>7.4.1</td>
<td>8.2.1</td>
<td>7.8.1</td>
<td>8.4.2</td>
<td>8.3.10.0</td>
<td>OSPF, IS-IS, and IPv6 BGP chapters in the FTOS Command Line Reference Guide</td>
</tr>
<tr>
<td>Multiprotocol BGP extensions for IPv6</td>
<td>7.4.1</td>
<td>8.2.1</td>
<td>7.8.1</td>
<td>8.4.2</td>
<td>8.3.10.0</td>
<td>IPv6 BGP in the FTOS Command Line Reference Guide</td>
</tr>
<tr>
<td>IPv6 BGP MD5 Authentication</td>
<td>8.2.1.0</td>
<td>8.2.1.0</td>
<td>8.2.1.0</td>
<td>8.4.2</td>
<td>8.3.10.0</td>
<td>IPv6 BGP in the FTOS Command Line Reference Guide</td>
</tr>
<tr>
<td>IS-IS for IPv6</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>8.3.10.0</td>
<td>Chapter 27, “Intermediate System to Intermediate System,” on page 569 in the FTOS Configuration Guide</td>
</tr>
<tr>
<td>IS-IS for IPv6 support for redistribution</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>8.3.10.0</td>
<td>IPv6 IS-IS in the FTOS Command Line Reference Guide</td>
</tr>
<tr>
<td>ISIS for IPv6 support for distribute lists and administrative distance</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>8.3.10.0</td>
<td>IPv6 IS-IS in the FTOS Command Line Reference Guide</td>
</tr>
<tr>
<td>OSPF for IPv6 (OSPFv3)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>OSPFv3 in the FTOS Command Line Reference Guide</td>
</tr>
<tr>
<td>Equal Cost Multipath for IPv6</td>
<td>7.4.1</td>
<td>8.2.1</td>
<td>7.8.1</td>
<td>7.8.1</td>
<td>8.3.10.0</td>
<td></td>
</tr>
</tbody>
</table>

### IPv6 Services and Management

<table>
<thead>
<tr>
<th>Feature</th>
<th>FTOS 7.5.1</th>
<th>FTOS 8.2.1</th>
<th>FTOS 7.8.1</th>
<th>FTOS 7.8.1</th>
<th>FTOS 8.3.10.0</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telnet client over IPv6 (outbound Telnet)</td>
<td>7.5.1</td>
<td>8.2.1</td>
<td>7.8.1</td>
<td>7.8.1</td>
<td>8.3.10.0</td>
<td>Telnet with IPv6 in this chapter Control and Monitoring in the FTOS Command Line Reference Guide</td>
</tr>
<tr>
<td>Telnet server over IPv6 (inbound Telnet)</td>
<td>7.4.1</td>
<td>8.2.1</td>
<td>7.8.1</td>
<td>7.8.1</td>
<td>8.3.10.0</td>
<td>Telnet with IPv6 in this chapter Control and Monitoring in the FTOS Command Line Reference Guide</td>
</tr>
<tr>
<td>Secure Shell (SSH) client support over IPv6 (outbound SSH) Layer 3 only</td>
<td>7.5.1</td>
<td>8.2.1</td>
<td>7.8.1</td>
<td>7.8.1</td>
<td>8.3.10.0</td>
<td>SSH over an IPv6 Transport in this chapter</td>
</tr>
<tr>
<td>Secure Shell (SSH) server support over IPv6 (inbound SSH) Layer 3 only</td>
<td>7.4.1</td>
<td>8.2.1</td>
<td>7.8.1</td>
<td>7.8.1</td>
<td>8.3.10.0</td>
<td>SSH over an IPv6 Transport in this chapter</td>
</tr>
</tbody>
</table>
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 CE S Z 4810.

Path MTU (Maximum Transmission Unit), in accordance with RFC 1981, 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 25-2. MTU Discovery Path**

Source            | Router A          | Router B          | Destination
---                | ---               | ---               | ---
MTU 1600          | MTU 1400          | MTU 1200          |
Packet (MTU 1600)  | ICMPv6 (Type 2)   | Use MTU 1400      |
Packet (MTU 1400)  | ICMPv6 (Type 2)   | Use MTU 1200      |
Packet (MTU 1200)  | Packet Received   |
**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.

**Note:** If a neighboring node does not have an IPv6 address assigned, it must be manually pined to allow the IPv6 device to determine the relationship of the neighboring node.

**Note:** To avoid problems with network discovery, Dell Force10 recommends configuring the static route last or assigning an IPv6 address to the interface and assigning an address to the peer (the forwarding router’s address) less than 10 seconds apart.

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 25-3. NDP Router Redirect**

Send a Packet to Network 2001:db8::1428:57ab

Packet Destination (2001:db8::1428:57ab)

ICMPv6 Redirect (Data: Use Router C)

Packet Destination (Destination 2001:db8::1428:57ab)
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 `ipv6 nd mtu` command sets the value advertised to routers. It does not set the actual MTU rate. For example, if `ipv6 nd mtu` is set to 1280, the interface will still pass 1500-byte packets, if that is what is set with the `mtu` command.

QoS for IPv6

IPv6 QoS is supported on platform E

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.

IPv6 Multicast

IPv6 Multicast is supported on platforms E

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; 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 a 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>cam-profile ipv6-extacl</td>
<td>EXEC Privileged</td>
<td>Enable the CAM profile with IPv6 extended ACLs on the entire chassis or on a specific linecard</td>
</tr>
<tr>
<td>microcode ipv6-extacl</td>
<td></td>
<td>chassis changes the CAM profile for all linecards in the chassis</td>
</tr>
<tr>
<td>chassis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>linecard slot</td>
<td></td>
<td>linecard slot/port changes the CAM profile only for the specified slot</td>
</tr>
</tbody>
</table>

Figure 25-4 displays the IPv6 CAM profile summary for a chassis that already has IPv6 CAM profile configured. Figure 25-5 shows the full IPv6 CAM profiles. Refer to Chapter 11, Content Addressable Memory (CAM), on page 271 for complete information regarding CAM configuration.
Adjust your CAM-Profile on a C-Series or S-Series

The `cam-acl` command is supported on platforms C S.

Although this is not a mandatory step, 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>cam-acl { ipv6acl }</td>
<td>CONFIGURATION</td>
<td>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.</td>
</tr>
<tr>
<td>show cam-acl</td>
<td>EXEC</td>
<td>Show the current CAM settings.</td>
</tr>
<tr>
<td>show cam-acl-egress</td>
<td>EXEC Privilege</td>
<td>Provides information on FP groups allocated for the egress acl. You must allocate at least one group for L2ACL and IPv4 ACL. The total number of groups is 4.</td>
</tr>
</tbody>
</table>

Assign an IPv6 Address to an Interface

IPv6 Addresses are supported on platforms C E S 54810

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.

When you configure IPv6 addresses on multiple interfaces (ipv6 address command) and verify the configuration (show ipv6 interfaces command), the same link local (fe80) address is displayed for each IPv6 interface.

If the user attempts to configure more than two IPv6 addresses, the following error message appears:

Already configured maximum IPV6 addresses for this interface type
One of the existing IPv6 addresses must be deleted before a new IPv6 address can be configured.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipv6 address ipv6 address/mask</td>
<td>CONFIG-INTERFACE</td>
<td>Enter the IPv6 Address for the device. ipv6 address : x:x:x:x::x mask : prefix length 0 to 128</td>
</tr>
</tbody>
</table>

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.
**Note:** After you configure a static IPv6 route (`ipv6 route` command) and configure the forwarding router’s address (specified in the `ipv6 route` command) on a neighbor’s interface, the IPv6 neighbor is not displayed in the `show ipv6 route` command output.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ipv6 route prefix type {slot/port} forwarding router tag</code></td>
<td>CONFIGURATION</td>
<td>Set up IPv6 static routes</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>prefix</code>: IPv6 route prefix</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>type {slot/port}</code>: interface type and slot/port</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>forwarding router</code>: forwarding router’s address</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>tag</code>: route tag</td>
</tr>
</tbody>
</table>

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 40-Gigabit Ethernet interface, enter the keyword `fortyGigE` 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>telnet ipv6 address</td>
<td>EXEC or EXEC Privileged</td>
<td>Enter the IPv6 Address for the device. ipv6 address : x:x:x:x:x mask : prefix length 0-128</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show ipv6 ?</td>
<td>EXEC or</td>
<td>List the IPv6 show options</td>
</tr>
<tr>
<td></td>
<td>EXEC Privileged</td>
<td></td>
</tr>
</tbody>
</table>

FTOS#show ipv6 ?
accounting IPv6 accounting information
cam IPv6 CAM Entries
fib IPv6 FIB Entries
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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show ipv6 interface</td>
<td>EXEC</td>
<td>Show the currently running configuration for the specified interface</td>
</tr>
<tr>
<td>type {slot/port}</td>
<td></td>
<td>Enter the keyword <code>interface</code> followed by the type of interface and slot/port information:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For all brief summary of IPv6 status and configuration, enter the keyword <code>brief</code>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For all IPv6 configured interfaces, enter the keyword <code>configured</code>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10/100/1000 Ethernet interface, enter the keyword <code>GigabitEthernet</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a Gigabit Ethernet interface, enter the keyword <code>GigabitEthernet</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10 Gigabit Ethernet interface, enter the keyword <code>TenGigabitEthernet</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 40-Gigabit Ethernet interface, enter the keyword <code>fortyGigE</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a loopback interface, enter the keyword <code>loopback</code> followed by the loopback number</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a linecard interface, enter the keyword <code>linecard</code> followed by the slot number</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a port-channel interface, enter the keyword <code>port-channel</code> followed by the port-channel number</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a VLAN interface, enter the keyword <code>vlan</code> followed by the VLAN ID</td>
</tr>
</tbody>
</table>

Figure 25-6 illustrates the `show ipv6 interface` command output.

**Figure 25-6. 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show ipv6 route</td>
<td>EXEC</td>
<td>Show IPv6 routing information for the specified route type. Enter the keyword:</td>
</tr>
<tr>
<td>type</td>
<td></td>
<td>• To display information about a network, enter the ipv6 address (X:X:X:X::X).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To display information about a host, enter the hostname.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To display information about all IPv6 routes (including non-active routes), enter all.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To display information about all connected IPv6 routes, enter connected.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To display information about brief summary of all IPv6 routes, enter summary.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To display information about Border Gateway Protocol (BGP) routes, enter bgp.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To display information about ISO IS-IS routes, enter isis.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To display information about Open Shortest Path First (OSPF) routes, enter ospf.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To display information about Routing Information Protocol (RIP), enter rip.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To display information about static IPv6 routes, enter static.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To display information about an IPv6 Prefix lists, enter list and the prefix-list name.</td>
</tr>
</tbody>
</table>

Figure 25-7 illustrates the show ipv6 route command output.

```
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
C    2002::/120 [0/0]    Direct, Gi 1/1, 00:28:49
C    2003::/120 [0/0]    Direct, Gi 1/1, 00:28:49
```

Figure 25-8 illustrates the show ipv6 route summary command output.
Figure 25-8. Command Example: show ipv6 route summary

FTOS#show ipv6 route summary

Route Source       Active Routes Non-active Routes
connected          5               0
static             0               0
Total              5               0

Figure 25-9 illustrates the show ipv6 route static command output.

Figure 25-9. Command Example: show ipv6 route static

FTOS#show ipv6 route static

Destination Dist/Metric, Gateway, Last Change
-----------------------------------------------------
S 8888:9999:5555:6666:1111::/96 [1/0]  via 2222:2222:3333:3333::1, Gi 9/1, 00:03:16

Show the Running-Configuration for an Interface

View the configuration for any interface with the following command.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show running-config</td>
<td>EXEC</td>
<td>Show the currently running configuration for the specified interface. Enter the keyword <code>interface</code> followed by the type of interface and slot/port information:</td>
</tr>
<tr>
<td>interface type {slot/port}</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10/100/1000 Ethernet interface, enter the keyword <code>GigabitEthernet</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a Gigabit Ethernet interface, enter the keyword <code>GigabitEthernet</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For the Management interface on the RPM, enter the keyword <code>ManagementEthernet</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10 Gigabit Ethernet interface, enter the keyword <code>TenGigabitEthernet</code> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 40-Gigabit Ethernet interface, enter the keyword <code>fortyGigE</code> followed by the slot/port information.</td>
</tr>
</tbody>
</table>

Figure 25-10 illustrates the show running-config command output. Note the IPv6 address listed.
Clear IPv6 Routes

Use the clear IPv6 route command to clear routes from the IPv6 routing table.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>clear ipv6 route {*</td>
<td>ipv6 address prefix-length}</td>
<td>EXEC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>* : all routes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ipv6 address : x:x:x:x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>mask : prefix length 0-128</td>
</tr>
</tbody>
</table>

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.
iSCSI Optimization

iSCSI Optimization is supported on platform [S4810].

This chapter describes how to configure internet small computer system interface (iSCSI) optimization, which enables quality-of-service (QoS) treatment for iSCSI traffic. The topics covered in this chapter include:

- iSCSI Optimization Overview
- Default iSCSI Optimization Values
- iSCSI Optimization Prerequisites
- Configuring iSCSI Optimization
- Displaying iSCSI Optimization Information

iSCSI Optimization Overview

iSCSI is a TCP/IP-based protocol for establishing and managing connections between IP-based storage devices and initiators in a storage area network (SAN). iSCSI optimization enables the networking switch to auto-detect Dell’s iSCSI storage arrays and triggers a self-configuration of several key network configurations that enables the network to be optimized for better storage traffic throughput.

iSCSI optimization also provides a means of monitoring iSCSI sessions and applying QoS policies on iSCSI traffic. When enabled, iSCSI optimization allows a switch to monitor (snoop) the establishment and termination of iSCSI connections. The switch uses the snooped information to detect iSCSI sessions and connections established through the switch.

iSCSI session monitoring over VLT will synchronize the ISCSI session information between the VLT peers, thereby allowing session information to be available in both the VLT peers.

iSCSI optimization functions as follows:

- Auto-detection of EqualLogic storage arrays — The switch detects any active EqualLogic array directly attached to its ports.
- Manual configuration to detect Compellent storage arrays where auto-detection is not supported.
- Automatic configuration of switch ports after detection of storage arrays.
- iSCSI monitoring sessions — The switch monitors and tracks active iSCSI sessions in connections on the switch, including port information and iSCSI session information.
• iSCSI QoS—A user-configured iSCSI class of service (CoS) profile is applied to all iSCSI traffic. Classifier rules are used to direct the iSCSI data traffic to queues that can be given preferential QoS treatment over other data passing through the switch. Preferential treatment helps to avoid session interruptions during times of congestion that would otherwise cause iSCSI packets to be dropped.
• iSCSI DCBX TLVs are supported.

Figure 26-1 shows iSCSI optimization between servers and a storage array in which a stack of three switches connect installed servers (iSCSI initiators) to a storage array (iSCSI targets) in a SAN network. iSCSI optimization running on the master switch is configured to use dot1p priority-queue assignments to ensure that iSCSI traffic in these sessions receives priority treatment when forwarded on stacked switch hardware.

Figure 26-1. iSCSI Optimization Example
Monitoring iSCSI Traffic Flows

The switch snoops iSCSI session-establishment and termination packets by installing classifier rules that trap iSCSI protocol packets to the CPU for examination. Devices that initiate iSCSI sessions usually use well-known TCP ports 3260 or 860 to contact targets. When you enable iSCSI optimization, by default the switch identifies IP packets to or from these ports as iSCSI traffic.

You can configure the switch to monitor traffic for additional port numbers or a combination of port number and target IP address, and you can remove the well-known port numbers from monitoring.

Application of Quality of Service to iSCSI Traffic Flows

The iSCSI CoS mode is user-configurable and controls whether CoS (dot1p priority) queue assignment and/or packet marking is performed on iSCSI traffic. When you enable iSCSI CoS mode, the CoS policy is applied to iSCSI traffic. When you disable iSCSI CoS mode, iSCSI sessions and connections are still detected and displayed in the status tables, but no CoS policy is applied to iSCSI traffic.

You can configure whether the iSCSI optimization feature uses the VLAN priority or IP DSCP mapping to determine the traffic class queue. By default, iSCSI flows are assigned to dot1p priority 4. Use the CoS dot1p-priority command to map incoming iSCSI traffic on an interface to a dot1p priority-queue other than 4 (refer to QoS dot1p Traffic Classification and Queue Assignment). Dell Force10 recommends setting the CoS dot1p priority-queue to 0 (zero).

You can configure whether iSCSI frames are re-marked to contain the configured VLAN priority tag or IP DSCP when forwarded through the switch.

**Note:** On a switch in which a large proportion of traffic is iSCSI, CoS queue assignments may interfere with other network control-plane traffic, such as ARP or LACP. Preferential treatment of iSCSI traffic needs to be balanced against the needs of other critical data in the network.

Information Monitored in iSCSI Traffic Flows

iSCSI optimization examines the following data in packets and uses the data to track the session and create the classifier entries that enable QoS treatment:

- Initiator’s IP Address
- Target’s IP Address
- ISID (Initiator defined session identifier)
- Initiator’s IQN (iSCSI qualified name)
- Target’s IQN
- Initiator’s TCP Port
- Target’s TCP Port
- Connection ID
- Aging
- Up Time
If no iSCSI traffic is detected for a session during a user-configurable aging period, the session data is cleared.

If more than 256 simultaneous sessions are logged continuously, the following message displays indicating the queue rate limit has been reached.


Only sessions observed by the switch will be learnt; sessions flowing through an adjacent switch will not be learnt. Session monitoring learns sessions that actually flow through the switch, it does not learn all sessions in the entire topology.

After a switch is reloaded, any information exchanged during the initial handshake is not available. If the switch picks up the communication after reloading, it would detect a session was in progress but could not obtain complete information for it. Any incomplete information of this type would not be available in the “show” commands.

Detection and Auto-configuration for Dell EqualLogic Arrays

The iSCSI optimization feature includes auto-provisioning support with the ability to detect directly connected Dell EqualLogic storage arrays and automatically reconfigure the switch to enhance storage traffic flows.

The switch uses the link layer discovery protocol (LLDP) to discover Dell EqualLogic devices on the network. LLDP is enabled by default. For more information about LLDP, refer to Chapter 30, Link Layer Discovery Protocol (LLDP).

The following message is displayed the first time a Dell EqualLogic array is detected and describes the configuration changes that are automatically performed:

%STKUNIT0-M:CP %IFMGR-5-IFM_ISCSI_AUTO_CONFIG: This switch is being configured for optimal conditions to support iSCSI traffic which will cause some automatic configuration to occur including jumbo frames and flow-control on all ports; no storm control and spanning-tree port fast to be enabled on the port of detection.

The following syslog message is generated the first time an EqualLogic array is detected:

%STKUNIT0-M:CP %LLDP-5-LLDP_EQL_DETECTED: EqualLogic Storage Array detected on interface Te 1/43

• At the first detection of an EqualLogic array, an MTU of 12000 is enabled on all ports and port-channels (if it is has not already been enabled).
• Spanning-tree portfast is enabled on the interface identified by LLDP.
• Unicast storm control is disabled on the interface identified by LLDP.
Detection and Port Configuration for Dell Compellent Arrays

Switches support the `iscsi profile-compellent` command to configure a port connected to a Dell Compellent storage array. The command configures a port for the best iSCSI traffic conditions and must be entered in INTERFACE Configuration mode.

The following message is displayed the first time you use the `iscsi profile-compellent` command to configure a port connected to a Dell Compellent storage array and describes the configuration changes that are automatically performed:

```
%STKUNIT0-M:CP %IFMGR-5-IFM_ISCSI_AUTO_CONFIG: This switch is being configured for optimal conditions to support iSCSI traffic which will cause some automatic configuration to occur including jumbo frames and flow-control on all ports; no storm control and spanning-tree portfast to be enabled on the port of detection.
```

After you execute the `iscsi profile-compellent` command, the following actions occur:

- Jumbo frame size is set to 12000 for all interfaces on all ports and port-channels, if it is not already enabled.
- Spanning-tree portfast is enabled on the interface.
- Unicast storm control is disabled on the interface.

You must enter the `iscsi profile-compellent` command in INTERFACE configuration mode; for example:

```
FTOS(conf-if-te-0/50# iscsi profile-compellent
```

Synchronizing iSCSI Sessions Learned on VLT-Lags with VLT-Peer

The following behavior occurs during synchronization of iSCSI sessions:

- If the iSCSI login request packet is received on a port belonging to a VLT lag, the information is synced to the VLT peer and the connection is associated with this interface.
- Additional updates to connections (including aging updates) that are learnt on VLT lag members are synced to the peer.
- When receiving an iSCSI login request on a non-VLT interface followed by a response from a VLT interface, the session is not synced since it is initially learnt on a non-VLT interface through the request packet.
- A new connection log is generated by the peer that sees the login response packet. If the login response packet uses the ICL path, it will be seen by both the peers, which in turn generate logs for this connection.
Enabling and Disabling iSCSI Optimization

**Note:** iSCSI monitoring is disabled by default. iSCSI auto-configuration and auto-detection is enabled by default.

If iSCSI is enabled, flow control will be automatically enabled on all interfaces. To disable the flow control on all interfaces, enter the command “no flow control rx on tx off” and save the configuration. To disable iSCSI optimization, which can turn on flow control again on reboot, enter the command “no iscsi enable” and save the configuration.

When you enable iSCSI on the switch, the following actions occur:

- Link-level flow control is globally enabled, if it is not already enabled, and PFC is disabled.
- iSCSI session snooping is enabled.
- iSCSI LLDP monitoring starts to automatically detect EqualLogic arrays.

The following message is displayed when you enable iSCSI on a switch and describes the configuration changes that are automatically performed:

```
%STKUNIT0-M:CP %IFMGR-5-IFM_ISCSI_ENABLE: iSCSI has been enabled causing flow control to be enabled on all interfaces. EQL detection and enabling iscsi profile-compellent on an interface may cause some automatic configurations to occur like jumbo frames on all ports and no storm control and spanning tree port-fast on the port of detection.
```

You can reconfigure any of the auto-provisioned configuration settings that result when you enable iSCSI on a switch.

When you disable the iSCSI feature, iSCSI resources are released and the detection of EqualLogic arrays using LLDP is disabled. Disabling iSCSI does not remove the MTU, flow control, portfast, or storm control configuration applied as a result of enabling iSCSI.

**Note:** By default, CAM allocation for iSCSI is set to 0. This disables session monitoring.

### Default iSCSI Optimization Values

**Table 26-1** shows the default values for the iSCSI optimization feature.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Default Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>iSCSI Optimization global setting</td>
<td>Enabled</td>
</tr>
<tr>
<td>iSCSI CoS mode (802.1p priority queue mapping)</td>
<td>Enabled: dot1p priority 4 without <strong>remark</strong> setting</td>
</tr>
<tr>
<td>iSCSI CoS Packet classification</td>
<td>iSCSI packets are classified by VLAN instead of by DSCP values.</td>
</tr>
</tbody>
</table>
iSCSI Optimization Prerequisites

- iSCSI optimization requires LLDP on the switch. LLDP is enabled by default (refer to Chapter 30, Link Layer Discovery Protocol (LLDP)).
- iSCSI optimization requires two ingress ACL groups to be configured. The ACL groups are allocated after iSCSI Optimization is configured. (refer to When to Use CAM Profiling).

Configuring iSCSI Optimization

To configure iSCSI optimization on a switch, follow these steps:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>For a non-DCB environment: Enable session monitoring. Configuring fcoeacl and iscioptacl are optional in FTOS version 9.1.(0.0). For a DCB environment: Configure iSCSI Optimization. The configuration files are stored in the flash memory in the CONFIG_TEMPLATE. Note: DCB/DCBX is enabled when the iSCSI configuration is applied. For OpenFlow: Allocate CAM space.</td>
<td>cam acl 12acl 4 ipv4acl 4 ipv6acl 0 ipv4qos 2 12 qos 1 12pt 0 ipmacacl 0 vman-qos 0 ecfmacl 0 fcoeacl 2 iscioptacl 2</td>
<td>CONFIGURATION EXEC Privilege</td>
</tr>
<tr>
<td>2</td>
<td>Save the configuration on the switch.</td>
<td>write memory</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>3</td>
<td>Reload the switch. After the switch is reloaded, DCB/DCBX and iSCSI monitoring will be enabled.</td>
<td>reload</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Step</td>
<td>Task</td>
<td>Command</td>
<td>Command Mode</td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>--------------</td>
</tr>
<tr>
<td>4</td>
<td>(Optional) Configure the iSCSI target ports and optionally the IP addresses on which iSCSI communication will be monitored, where:</td>
<td>[no] iscsi target port tcp-port-1 tcp-port-2...tcp-port-16 [address ip-address]</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td></td>
<td>• tcp-port-n is the TCP port number or a list of TCP port numbers on which the iSCSI target listens to requests. You can configure up to 16 target TCP ports on the switch in one command or multiple commands. Default: 860, 3260. Separate port numbers with a comma.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• ip-address specifies the IP address of the iSCSI target. When you enter the no form of the command, and the TCP port to be deleted is one bound to a specific IP address, the IP address value must be included in the command.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>(Optional) Set the QoS policy that will be applied to iSCSI flows, where:</td>
<td>[no] iscsi cos {enable</td>
<td>disable</td>
</tr>
<tr>
<td></td>
<td>• enable: enables the application of preferential QoS treatment to iSCSI traffic so that iSCSI packets are scheduled in the switch with a dot1p priority 4 regardless of the VLAN priority tag in the packet. Default: iSCSI packets are handled with dot1p priority 4 without remark.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• disable: disables the application of preferential QoS treatment to iSCSI frames.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• dot1p vlan-priority-value: specifies the VLAN priority tag assigned to incoming packets in an iSCSI session. Range: 0 to 7. Default: The dot1p value in ingress iSCSI frames is not changed and is used in iSCSI TLV advertisements if the iscsi priority-bits command is not entered (Step 5).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• dscp dscp-value: specifies the DSCP value assigned to incoming packets in an iSCSI session. Range: 0 to 63. Default: The DSCP value in ingress packets is not changed.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• remark: marks incoming iSCSI packets with the configured dot1p or DSCP value when they egress the switch. Default: The dot1 and DSCP values in egress packets are not changed.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>(Optional) Set the aging time for iSCSI session monitoring. Range: 5 to 43,200 minutes. Default: 10 minutes.</td>
<td>[no] iscsi aging time time</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>7</td>
<td>(Optional) Configures DCBX to send iSCSI TLV advertisements. You can configure iSCSI TLVs to be sent either globally or on a specified interface. The interface configuration takes priority over global configuration. Default: Enabled.</td>
<td>[no] advertise dcbx-app-tlv iscsi</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td></td>
<td>or INTERFACE</td>
<td>INTERFACE</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>8</td>
<td>(Optional) Configures the priority bitmap to be advertised in iSCSI application TLVs. Default: 4 (0x10 in the bitmap).</td>
<td>[no] iscsi priority-bits</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
Displaying iSCSI Optimization Information

Use the `show` commands in Table 26-2 to display information on iSCSI optimization.

Table 26-2. Displaying iSCSI Optimization Information

<table>
<thead>
<tr>
<th>Command</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>show iscsi</code> (Figure 26-2)</td>
<td>Displays the currently configured iSCSI settings.</td>
</tr>
<tr>
<td><code>show iscsi sessions</code> (Figure 26-3)</td>
<td>Displays information on active iSCSI sessions on the switch.</td>
</tr>
<tr>
<td><code>show iscsi sessions detailed [session isid]</code> (Figure 26-4)</td>
<td>Displays detailed information on active iSCSI sessions on the switch. To display detailed information on specified iSCSI session, enter the session’s iSCSI ID.</td>
</tr>
<tr>
<td><code>show run iscsi</code></td>
<td>Displays all globally-configured non-default iSCSI settings in the current FTOS session.</td>
</tr>
</tbody>
</table>

Figure 26-2. show iscsi Command Example

```
FTOS#show iscsi
iSCSI is enabled
iSCSI session monitoring is disabled
iSCSI COS : dot1p is 4 no-remark
Session aging time: 10
Maximum number of connections is 256
-----------------------------------------------
iSCSI Targets and TCP Ports:
-----------------------------------------------
TCP Port  Target IP Address
3260
860
```
**Figure 26-3. show iscsi session Command Example**

VLT PEER1

FTOS# show iscsi session
Session 0:
-----------------------------------------------------------------------------------------
Target: iqn.2001-05.com.equallogic:0-8a0906-0e70c2002-10a0018426a48c94-iom010
ISID: 400001370000

VLT PEER2
Session 0:
-----------------------------------------------------------------------------------------
Target: iqn.2001-05.com.equallogic:0-8a0906-0f60c2002-0360018428d48c94-iom011
ISID: 400001370000

**Figure 26-4. show iscsi session detailed Command Example**

VLT PEER1

FTOS# show iscsi session detailed
Session 0:
-----------------------------------------------------------------------------------------
Up Time: 00:00:01:28 (DD:HH:MM:SS)
Time for aging out: 00:00:09:34 (DD:HH:MM:SS)
ISID: 806978696102
Initiator  Initiator  Target  Target  Connection
  IP Address   TCP Port   IP Address   TCPPort   ID
  10.10.0.44   33345     10.10.0.101  3260      0

VLT PEER2
Session 0:
-----------------------------------------------------------------------------------------
Up Time: 00:00:01:28 (DD:HH:MM:SS)
Time for aging out: 00:00:09:34 (DD:HH:MM:SS)
ISID: 806978696102
Initiator  Initiator  Target  Target  Connection
  IP Address   TCP Port   IP Address   TCPPort   ID
  10.10.0.53   33432     10.10.0.101  3260      0
Intermediate System to Intermediate System

Intermediate System to Intermediate System is supported on the following platforms: E Z S4810 platforms.

IS-IS is supported on the E-Series ExaScale platform with FTOS 8.1.1.0 and later. It is supported on the S4810 with FTOS 8.3.10.0 and on the Z platform with FTOS 9.0.0.0

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 it is detailed in this chapter.

- Protocol Overview on page 569
- IS-IS Addressing on page 570
- Multi-Topology IS-IS on page 571
- Graceful Restart on page 572
- Implementation Information on page 573
- Configuration Information on page 574
- IS-IS Metric Styles on page 593
- Sample Configuration on page 597

IS-IS protocol standards are listed in the Chapter 50, 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 hierarchically 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.

**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 27-1 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 27-1. ISO Address Format**
Multi-Topology IS-IS

FTOS 7.8.1.0 and later support Multi-Topology Routing IS-IS.

E-Series ExaScale platform supports Multi-Topology IS-IS with FTOS 8.2.1.0 and later.

S-Series platform supports Multi-Topology IS-IS with FTOS 8.3.10.0 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 the E and S4810 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 functionally 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 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 27-1 displays the default values for IS-IS.

### Table 27-1. IS-IS Default Values

<table>
<thead>
<tr>
<th>IS-IS Parameter</th>
<th>Default Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete Sequence Number PDU (CSNP) interval</td>
<td>10 seconds</td>
</tr>
<tr>
<td>IS-to-IS hello PDU interval</td>
<td>10 seconds</td>
</tr>
<tr>
<td>IS-IS interface metric</td>
<td>10</td>
</tr>
<tr>
<td>Metric style</td>
<td>Narrow</td>
</tr>
<tr>
<td>Designated Router priority</td>
<td>64</td>
</tr>
<tr>
<td>Circuit Type</td>
<td>Level 1 and Level 2</td>
</tr>
<tr>
<td>IS Type</td>
<td>Level 1 and Level 2</td>
</tr>
<tr>
<td>Equal Cost Multi Paths</td>
<td>16</td>
</tr>
</tbody>
</table>

**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 on page 575
- Configure Multi-Topology IS-IS (MT IS-IS) on page 578
- Configure IS-IS Graceful Restart on page 579
- Change LSP attributes on page 582
- Configure IS-IS metric style and cost on page 583
- Change the IS-type on page 585
- Control routing updates on page 586
- Configure authentication passwords on page 590
• Set the overload bit on page 591
• Debug IS-IS on page 592

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
</table>
| 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 |
<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
</table>
| 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: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 ipv6-address mask` | INTERFACE |
| 6    | Enable IS-IS on the IPv4 interface. If you configure a `tag` variable, it must be the same as the `tag` variable assigned in step 1. | `ip router isis [tag]` | ROUTER ISIS |
| 7    | Enable IS-IS on the IPv6 interface. If you configure a `tag` variable, it must be the same as the `tag` variable assigned in step 1. | `ipv6 router isis [tag]` | ROUTER ISIS |

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.
Use the `show isis traffic` command in EXEC Privilege mode to view IS-IS protocol statistics.

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)

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enable Multi-Topology IS-IS for IPv6.</td>
<td>multi-topology [transition]</td>
<td>ROUTER ISIS AF IPV6</td>
</tr>
<tr>
<td></td>
<td>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.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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.                         | isis ipv6 metric metric-value [level-1 | level-2 | level-1-2] | ROUTER ISIS AF IPV6 |
|      | To configure wide or wide transition metric style, the cost can be a between 0 and 16,777,215. |                                                   |                        |
Configure Multi-Topology IS-IS (MT IS-IS)

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enable Multi-Topology IS-IS for IPv6.</td>
<td><code>multi-topology [transition]</code></td>
<td>ROUTER ISIS AF IPV6</td>
</tr>
<tr>
<td></td>
<td>Enter the <code>transition</code> 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.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Excluded this router from other router’s SPF calculations.</td>
<td><code>set-overload-bit</code></td>
<td>ROUTER ISIS AF IPV6</td>
</tr>
<tr>
<td>3</td>
<td>Set the minimum interval between SPF calculations.</td>
<td>`spf-interval [level-1</td>
<td>level-2</td>
</tr>
<tr>
<td></td>
<td>This command is used for IPv6 route computation only when multi-topology is enabled. If using single-topology mode, use the <code>spf-interval</code> command in CONFIG ROUTER ISIS mode to apply to both IPv4 and IPv6 route computations.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Implement a wide metric-style globally.</td>
<td>`isis ipv6 metric metric-value [level-1</td>
<td>level-2</td>
</tr>
<tr>
<td></td>
<td>To configure wide or wide transition metric style, the cost can be a between 0 and 16,777,215.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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.

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>graceful-restart ietf</code></td>
<td>ROUTER-ISIS</td>
<td>Enable Graceful Restart on ISIS processes</td>
</tr>
<tr>
<td><code>graceful-restart interval minutes</code></td>
<td>ROUTER-ISIS</td>
<td>Configure the period of time during which the Graceful Restart attempt will be prevented. Range: 1-120 minutes Default: 5 minutes</td>
</tr>
<tr>
<td>Command Syntax</td>
<td>Command Mode</td>
<td>Purpose</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>-----------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><code>graceful-restart restart-wait seconds</code></td>
<td>ROUTER-ISIS</td>
<td>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</td>
</tr>
</tbody>
</table>
| `graceful-restart t1 {interval seconds | 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) |
| `| retry-times value}` |                 |                                                                                                                                      |
| `graceful-restart t2 {level-1 | 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 |
| `seconds` |                 |                                                                                                                                      |
| `graceful-restart t3 {adjacency | 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 this option.  
  - **manual**: allows you to specify a fixed value that the restarting router should use.  
   Range: 50-120 seconds Default: 30 seconds | manual seconds} |                 |                                                                                                                                      |
| `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 27-4. 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.
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>`lsp-gen-interval [level-1</td>
<td>level-2] seconds`</td>
<td>ROUTER ISIS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>seconds</code> range: 0 to 120</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default is 5 seconds.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default level is Level 1.</td>
</tr>
<tr>
<td><code>lsp-mtu size</code></td>
<td>ROUTER ISIS</td>
<td>Set the LSP size.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>size</code> range: 128 to 9195.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default is 1497.</td>
</tr>
<tr>
<td><code>lsp-refresh-interval seconds</code></td>
<td>ROUTER ISIS</td>
<td>Set the LSP refresh interval.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>seconds</code> range: 1 to 65535.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default is 900 seconds.</td>
</tr>
<tr>
<td><code>max-lsp-lifetime seconds</code></td>
<td>ROUTER ISIS</td>
<td>Set the maximum time LSPs lifetime.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>seconds</code> range: 1 to 65535.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default is 1200 seconds.</td>
</tr>
</tbody>
</table>

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).
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>
<thead>
<tr>
<th>Metric Style</th>
<th>Characteristics</th>
<th>Cost Range Supported on IS-IS Interfaces</th>
</tr>
</thead>
<tbody>
<tr>
<td>narrow</td>
<td>Sends and accepts narrow or old TLVs (Type Length Value).</td>
<td>0 to 63</td>
</tr>
<tr>
<td>wide</td>
<td>Sends and accepts wide or new TLVs</td>
<td>0 to 16777215</td>
</tr>
<tr>
<td>transition</td>
<td>Sends both wide (new) and narrow (old) TLVs.</td>
<td>0 to 63</td>
</tr>
<tr>
<td>narrow transition</td>
<td>Sends narrow (old) TLVs and accepts both narrow (old) and wide (new) TLVs</td>
<td>0 to 63</td>
</tr>
<tr>
<td>wide transition</td>
<td>Sends wide (new) TLVs and accepts both narrow (old) and wide (new) TLVs</td>
<td>0 to 16777215</td>
</tr>
</tbody>
</table>

Use the following command in ROUTER ISIS mode to change the IS-IS metric style of the IS-IS process.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>metric-style {narrow [transition]</td>
<td>ROUTER ISIS</td>
<td>Set the metric style for the IS-IS process.</td>
</tr>
<tr>
<td>transition</td>
<td>wide [transition]} [level-1</td>
<td>set the metric style for the IS-IS process.</td>
</tr>
<tr>
<td></td>
<td>level-2]</td>
<td>Default: narrow</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default: Level 1 and Level 2 (level-1-2)</td>
</tr>
</tbody>
</table>

Use the `show isis protocol` command (Figure 476) in EXEC Privilege mode to view which metric types are generated and received.
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.

```
FTOS(config)#isis metric default-metric [level-1 | level-2]
```

**Table 27-3. Correct Value Range for the isis metric command**

<table>
<thead>
<tr>
<th>Metric Style</th>
<th>Correct Value Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>wide</td>
<td>0 to 16777215</td>
</tr>
<tr>
<td>narrow</td>
<td>0 to 63</td>
</tr>
<tr>
<td>wide transition</td>
<td>0 to 16777215</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>`is-type {level-1</td>
<td>level-1-2</td>
<td>level-2-only}`</td>
</tr>
</tbody>
</table>

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).
Control routing updates

Use the following commands in ROUTER ISIS mode to control the source of IS-IS route information.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 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 6, Access Control Lists (ACLs).

**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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>distribute-list prefix-list-name in [interface]</td>
<td>ROUTER ISIS</td>
<td>Apply a configured prefix list to all incoming IPv4 IS-IS routes. Enter the type of interface and slot/port information:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 1-Gigabit Ethernet interface, enter the keyword <strong>GigabitEthernet</strong> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For the Loopback interface on the RPM, enter the keyword <strong>loopback</strong> followed by a number from 0 to 16383.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a port channel, enter the keyword <strong>port-channel</strong> followed by a number from 1 to 255 for TeraScale and ExaScale, 1 to 32 for EtherScale.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a SONET interface, enter the keyword <strong>sonet</strong> followed by slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10-Gigabit Ethernet interface, enter the keyword <strong>TenGigabitEthernet</strong> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a VLAN, enter the keyword <strong>vlan</strong> 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.</td>
</tr>
<tr>
<td>distribute-list prefix-list-name out [bgp as-number</td>
<td>ROUTER ISIS</td>
<td>Apply a configured prefix list to all outgoing IPv4 IS-IS routes. You can configure one of the optional parameters:</td>
</tr>
<tr>
<td>connected</td>
<td>ospf process-id</td>
<td>rip</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>ospf process-id</strong>: for OSPF routes only.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>rip</strong>: for RIP routes only.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>static</strong>: for user-configured routes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>bgp</strong>: for BGP routes only</td>
</tr>
<tr>
<td>distribute-list redistributed-override in</td>
<td>ROUTER ISIS</td>
<td>Deny RTM download for pre-existing redistributed IPv4 routes</td>
</tr>
</tbody>
</table>
**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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>distribute-list prefix-list-name <strong>in</strong> [interface]</td>
<td>ROUTER ISIS-AF</td>
<td>Apply a configured prefix list to all incoming IPv6 IS-IS routes. Enter the type of interface and slot/port information:</td>
</tr>
<tr>
<td></td>
<td>IPV6</td>
<td>• For a 1-Gigabit Ethernet interface, enter the keyword <strong>GigabitEthernet</strong> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For the Loopback interface on the RPM, enter the keyword <strong>loopback</strong> followed by a number from 0 to 16383.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a port channel, enter the keyword <strong>port-channel</strong> followed by a number from 1 to 255 for TeraScale, 1 to 32 for EtherScale.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a SONET interface, enter the keyword <strong>sonet</strong> followed by slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10-Gigabit Ethernet interface, enter the keyword <strong>TenGigabitEthernet</strong> followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a VLAN, enter the keyword <strong>vlan</strong> followed by a number from 1 to 4094.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANs.</td>
</tr>
<tr>
<td>distribute-list prefix-list-name <strong>out</strong> [bgp as-number</td>
<td>ROUTER ISIS-AF</td>
<td>Apply a configured prefix list to all outgoing IPv6 IS-IS routes. You can configure one of the optional parameters:</td>
</tr>
<tr>
<td></td>
<td>connected</td>
<td>IPV6</td>
</tr>
<tr>
<td></td>
<td>ospf process-id</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rip</td>
<td></td>
</tr>
<tr>
<td></td>
<td>static</td>
<td></td>
</tr>
<tr>
<td></td>
<td>bgp</td>
<td></td>
</tr>
<tr>
<td>distribute-list redistributed-override <strong>in</strong></td>
<td>ROUTER ISIS-AF</td>
<td>Deny RTM download for pre-existing redistributed IPv6 routes</td>
</tr>
<tr>
<td>IPV6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 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, level-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, level-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 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 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, level-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, level-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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>area-password [hmac-md5]</code></td>
<td>ROUTER ISIS</td>
<td>Configure authentication password for an area. FTOS supports HMAC-MD5 authentication. This password is inserted in Level 1 LSPs, Complete SNPs, and Partial SNPs.</td>
</tr>
<tr>
<td><code>password</code></td>
<td></td>
<td></td>
</tr>
<tr>
<td>`domain-password [encryption-type</td>
<td>ROUTER ISIS</td>
<td>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.</td>
</tr>
<tr>
<td>hmac-md5] password`</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>set-overload-bit</code></td>
<td>ROUTER ISIS</td>
<td>Set the overload bit in LSPs. This prevents other routers from using it as an intermediate hop in their shortest path first (SPF) calculations.</td>
</tr>
</tbody>
</table>

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 27-9, 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 27-9. 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 0x61A7 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>debug isis</td>
<td>EXEC Privilege</td>
<td>View all IS-IS information.</td>
</tr>
<tr>
<td>debug isis adj-packets</td>
<td>EXEC Privilege</td>
<td>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:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• interface: Enter the type of interface and slot/port information to view IS-IS information on that interface only.</td>
</tr>
<tr>
<td>debug isis local-updates</td>
<td>EXEC Privilege</td>
<td>View information about IS-IS local update packets. To view specific information, enter one of the following optional parameters:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• interface: Enter the type of interface and slot/port information to view IS-IS information on that interface only.</td>
</tr>
<tr>
<td>debug isis snp-packets</td>
<td>EXEC Privilege</td>
<td>View IS-IS SNP packets, include CSNPs and PSNPs. To view specific information, enter one of the following optional parameters:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• interface: Enter the type of interface and slot/port information to view IS-IS information on that interface only.</td>
</tr>
<tr>
<td>debug isis spf-triggers</td>
<td>EXEC Privilege</td>
<td>View the events that triggered IS-IS shortest path first (SPF) events for debugging purposes.</td>
</tr>
</tbody>
</table>
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.

### IS-IS Metric Styles

The following sections provide additional information on IS-IS Metric Styles.

- IS-IS Metric Styles on page 593
- Configure Metric Values on page 593

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:

- Maximum Values in the Routing Table on page 594
- Changing the IS-IS Metric Style in One Level Only on page 594
- Leaking from One Level to Another on page 596
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 27-4. Correct Value Range for the `isis metric` Command

<table>
<thead>
<tr>
<th>Metric Style</th>
<th>Correct Value Range for the <code>isis metric</code> Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>wide</td>
<td>0 to 16777215</td>
</tr>
<tr>
<td>narrow</td>
<td>0 to 63</td>
</tr>
<tr>
<td>wide transition</td>
<td>0 to 16777215</td>
</tr>
<tr>
<td>narrow transition</td>
<td>0 to 63</td>
</tr>
<tr>
<td>transition</td>
<td>0 to 63</td>
</tr>
</tbody>
</table>

### 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 27-5. Metric Value when Metric Style Changes

<table>
<thead>
<tr>
<th>Beginning metric style</th>
<th>Final metric style</th>
<th>Resulting IS-IS metric value</th>
</tr>
</thead>
<tbody>
<tr>
<td>wide</td>
<td>narrow</td>
<td>default value (10) if the original value is greater than 63. A message is sent to the console.</td>
</tr>
<tr>
<td>wide</td>
<td>transition</td>
<td>truncated value(^1) (the truncated value appears in the LSP only.) The original <code>isis metric</code> value is displayed in the <code>show config</code> and <code>show running-config</code> commands and is used if you change back to transition metric style.</td>
</tr>
<tr>
<td>wide</td>
<td>narrow transition</td>
<td>default value (10) if the original value is greater than 63. A message is sent to the console.</td>
</tr>
<tr>
<td>wide</td>
<td>wide transition</td>
<td>original value</td>
</tr>
<tr>
<td>narrow</td>
<td>wide</td>
<td>original value</td>
</tr>
<tr>
<td>narrow</td>
<td>transition</td>
<td>original value</td>
</tr>
<tr>
<td>narrow</td>
<td>narrow transition</td>
<td>original value</td>
</tr>
<tr>
<td>narrow</td>
<td>wide transition</td>
<td>original value</td>
</tr>
</tbody>
</table>
Table 27-6. Metric Value when Metric Style Changes (continued)

<table>
<thead>
<tr>
<th>Beginning metric style</th>
<th>Final metric style</th>
<th>Resulting IS-IS metric value</th>
</tr>
</thead>
<tbody>
<tr>
<td>transition</td>
<td>wide</td>
<td>original value</td>
</tr>
<tr>
<td>transition</td>
<td>narrow</td>
<td>original value</td>
</tr>
<tr>
<td>transition</td>
<td>narrow transition</td>
<td>original value</td>
</tr>
<tr>
<td>transition</td>
<td>wide transition</td>
<td>original value</td>
</tr>
<tr>
<td>narrow transition</td>
<td>wide</td>
<td>original value</td>
</tr>
<tr>
<td>narrow transition</td>
<td>narrow</td>
<td>original value</td>
</tr>
<tr>
<td>narrow transition</td>
<td>wide transition</td>
<td>original value</td>
</tr>
<tr>
<td>narrow transition</td>
<td>transition</td>
<td>original value</td>
</tr>
<tr>
<td>wide transition</td>
<td>wide</td>
<td>original value</td>
</tr>
<tr>
<td>wide transition</td>
<td>narrow</td>
<td>default value (10) if the original value is greater than 63. A message is sent to the console.</td>
</tr>
<tr>
<td>wide transition</td>
<td>narrow transition</td>
<td>default value (10) if the original value is greater than 63. A message is sent to the console.</td>
</tr>
<tr>
<td>wide transition</td>
<td>transition</td>
<td>truncated value (the truncated value appears in the LSP only.) The original <em>isis metric</em> value is displayed in the <em>show config</em> and <em>show running-config</em> commands and is used if you change back to transition metric style.</td>
</tr>
</tbody>
</table>

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 27-6).

Table 27-6. Metric Value when Metric Style Changes Multiple Times

<table>
<thead>
<tr>
<th>Beginning metric style</th>
<th>next isis metric style</th>
<th>resulting isis metric value</th>
<th>Next metric style</th>
<th>final isis metric value</th>
</tr>
</thead>
<tbody>
<tr>
<td>wide</td>
<td>transition</td>
<td>truncated value</td>
<td>wide</td>
<td>original value is recovered</td>
</tr>
<tr>
<td>wide transition</td>
<td>transition</td>
<td>truncated value</td>
<td>wide transition</td>
<td>original value is recovered</td>
</tr>
<tr>
<td>wide</td>
<td>transition</td>
<td>truncated value</td>
<td>narrow</td>
<td>default value (10) A message is sent to the logging buffer</td>
</tr>
<tr>
<td>wide transition</td>
<td>transition</td>
<td>truncated value</td>
<td>narrow transition</td>
<td>default value (10) A message is sent to the logging buffer</td>
</tr>
</tbody>
</table>
Leaking from One Level to Another

In the following scenarios, each IS-IS level is configured with a different metric style.

Table 27-7. Metric Value with Different Levels Configured with Different Metric Styles

<table>
<thead>
<tr>
<th>Level-1 metric style</th>
<th>Level-2 metric style</th>
<th>Resulting isis metric value</th>
</tr>
</thead>
<tbody>
<tr>
<td>narrow</td>
<td>wide</td>
<td>original value</td>
</tr>
<tr>
<td>narrow</td>
<td>wide transition</td>
<td>original value</td>
</tr>
<tr>
<td>narrow</td>
<td>narrow transition</td>
<td>original value</td>
</tr>
<tr>
<td>narrow</td>
<td>transition</td>
<td>original value</td>
</tr>
<tr>
<td>wide</td>
<td>narrow</td>
<td>truncated value</td>
</tr>
<tr>
<td>wide</td>
<td>narrow transition</td>
<td>truncated value</td>
</tr>
<tr>
<td>wide</td>
<td>wide transition</td>
<td>original value</td>
</tr>
<tr>
<td>wide</td>
<td>transition</td>
<td>truncated value</td>
</tr>
<tr>
<td>narrow transition</td>
<td>wide</td>
<td>original value</td>
</tr>
<tr>
<td>narrow transition</td>
<td>narrow</td>
<td>original value</td>
</tr>
<tr>
<td>narrow transition</td>
<td>wide transition</td>
<td>original value</td>
</tr>
<tr>
<td>narrow transition</td>
<td>transition</td>
<td>original value</td>
</tr>
<tr>
<td>transition</td>
<td>wide</td>
<td>original value</td>
</tr>
<tr>
<td>transition</td>
<td>narrow</td>
<td>original value</td>
</tr>
<tr>
<td>transition</td>
<td>wide transition</td>
<td>original value</td>
</tr>
<tr>
<td>transition</td>
<td>narrow transition</td>
<td>original value</td>
</tr>
<tr>
<td>wide transition</td>
<td>wide</td>
<td>original value</td>
</tr>
<tr>
<td>wide transition</td>
<td>narrow</td>
<td>truncated value</td>
</tr>
<tr>
<td>wide transition</td>
<td>narrow transition</td>
<td>truncated value</td>
</tr>
<tr>
<td>wide transition</td>
<td>transition</td>
<td>truncated value</td>
</tr>
</tbody>
</table>
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 the clear isis command. The clear isis command 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 27-10 is a sample configuration for enabling IPv6 IS-IS. Figure 27-13 illustrates the topology created with that CLI configuration.

IPv6 IS-IS routes can be configured in one of the following three different methods:

- Congruent Topology: Both IPv4 and IPv6 addresses must be configured on the interface. The commands `ip router isis` and `ipv6 router isis` must be enabled on the interface. You must enable the `wide-metrics` parameter in the router isis configuration mode.
- Multi-topology: The IPv6 address must be configured. Configuring the IPv4 address is optional. The command `ipv6 router isis` must be enabled on the interface. If IPv4 is configured, the command `ip router isis` must also be enabled. In the router isis configuration mode, enable multi-topology under address-family ipv6 unicast.
- Multi-topology Transition: The IPv6 address must be configured. Configuring the IPv4 address is optional. The command `ipv6 router isis` must be enabled on the interface. If IPv4 is configured, the command `ip router isis` must also be enabled. In the router isis configuration mode, enable multi-topology transition under address-family ipv6 unicast.
Figure 27-10.  IS-IS Sample Configuration - Congruent Topology

FTOS(conf-if-te-3/17)#show config
!
interface TenGigabitEthernet 3/17
ip address 24.3.1.1/24
ipv6 address 24:3::1/76
ip router isis
ipv6 router isis
no shutdown
FTOS (conf-if-te-3/17)#

FTOS (conf-router_isis)#show config
!
router isis
metric-style wide level-1
metric-style wide level-2
net 34.0000.0000.AAAA.00
FTOS (conf-router_isis)#

Figure 27-11.  IS-IS Sample Configuration - Multi-topology

FTOS (conf-if-te-3/17)#show config
!
interface TenGigabitEthernet 3/17
ipv6 address 24:3::1/76
ipv6 router isis
no shutdown
FTOS (conf-if-te-3/17)#

FTOS (conf-router_isis)#show config
!
router isis
net 34.0000.0000.AAAA.00
!
address-family ipv6 unicast
multi-topology
exit-address-family
FTOS (conf-router_isis)#

Figure 27-12.  IS-IS Sample Configuration - Multi-topology Transition

FTOS (conf-if-te-3/17)#show config
!
interface TenGigabitEthernet 3/17
ipv6 address 24:3::1/76
ipv6 router isis
no shutdown
FTOS (conf-if-te-3/17)#

FTOS (conf-router_isis)#show config
!
router isis
net 34.0000.0000.AAAA.00
!
address-family ipv6 unicast
multi-topology transition
exit-address-family
FTOS (conf-router_isis)#
Figure 27-13. IPv6 IS-IS Sample Topography
Link Aggregation Control Protocol (LACP)

Link Aggregation Control Protocol (LACP) is supported on platforms: \[\text{E} \quad \text{C} \quad \text{S} \quad \text{S4810}\]

The major sections in the chapter are:

- Introduction to Dynamic LAGs and LACP
- LACP Configuration Tasks
- Shared LAG State Tracking
- Configure LACP as Hitless
- LACP Basic Configuration Example

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 in the Interfaces chapter.

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

- 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 for more information.

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[no] lACP system-priority priority-value</td>
<td>CONFIGURATION</td>
<td>Configure the system priority. Range: 1–65535 (the higher the number, the lower the priority) Default: 32768</td>
</tr>
<tr>
<td>[no] port-channel-protocol lACP</td>
<td>INTERFACE</td>
<td>Enable or disable LACP on any LAN port:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Default is “LACP disabled”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• This command creates a new context.</td>
</tr>
<tr>
<td>[no] port-channel number mode [active</td>
<td>passive</td>
<td>off]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Default is “LACP active”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• number cannot statically contain any links</td>
</tr>
<tr>
<td>[no] lACP port-priority priority-value</td>
<td>LACP</td>
<td>Configure port priority.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Ranges: 1 – 65535 (the higher the number, the lower the priority)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Default: 32768</td>
</tr>
</tbody>
</table>

LACP Configuration Tasks

The tasks covered in this section are:

- Create a LAG
- Configure the LAG interfaces as dynamic
- Set the LACP long timeout
- Monitor and Debugging LACP
- Configure Shared LAG State Tracking

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 the following example, which uses the example of LAG 32:

```
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 as shown in the example below:

```
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. The following example 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`.

```
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 as shown in the example below:

1. Set the LACP timeout value to 30 seconds.

   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.

Monitor and Debugging LACP

The system log (syslog) records faulty LACP actions.

To debug LACP, use the following command:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[no] debug lacp [config</td>
<td>events</td>
<td>pdu in</td>
</tr>
</tbody>
</table>
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 the following illustration, 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.

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:

Note: If a LAG interface is part of a redundant pair, it cannot be used as a member of a failover group created for Shared LAG State Tracking.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter port-channel failover group mode.</td>
<td>port-channel failover-group</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Create a failover group and specify the two port-channels that will be members of the group.</td>
<td>group number port-channel number port-channel number</td>
<td>CONFIG-PO-FAILOVER-GRP</td>
</tr>
</tbody>
</table>

In the following example, LAGs 1 and 2 have been placed into to the same failover group.
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 the example below.

R2#show running-config po-failover-group
! port-channel failover-group
    group 1 port-channel 1 port-channel 2

In the following illustration, 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 1, in which a console message declares both LAGs down at the same time.

![Diagram of network topology](fnC0049mp)

**Message 1**  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 the following example.

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
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. Dynamic LAG interfaces in a redundant pair can be in a hitless LACP configuration.

Configure LACP to be hitless using the command `redundancy protocol lacp` from CONFIGURATION mode, as shown in the following example.

```
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 the following illustration. Two routers are named ALPHA and BRAVO, and their hostname prompts reflect those names.
The sections are:

- Configuring a LAG on ALPHA
- Summary of the configuration on ALPHA
- Summary of the configuration on BRAVO

**Configuring a LAG on ALPHA**

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)#

Inspecting a LAG port configuration on ALPHA.

Alpha#sh int gig 2/31
GigabitEthernet 2/31 is up, line protocol is up
Port is part of Port-channel 10
Hardware is Force10Eth, address is 00:01:e8:06:95:c0
   Current address is 00:01:e8:06:95:c0
Interface Index is 109101113
Port will not be disabled on partial SFM failure
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 1000 Mbit, Mode full duplex, Slave
Flowcontrol rx on tx on
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:02:11
Queueing strategy: fifo
Input statistics:
    132 packets, 163668 bytes
    0 Vlans
    0 64-byte pkts, 12 over 64-byte pkts, 120 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    132 Multicasts, 0 Broadcasts
    0 runts, 0 giants, 0 throttles
    0 CRC, 0 overrun, 0 discarded
Output Statistics
    136 packets, 16718 bytes, 0 underruns
    0 64-byte pkts, 15 over 64-byte pkts, 121 over 127-byte pkts
    0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
    136 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: 00:02:14
GigabitEthernet 2/31 is up, line protocol is up
Port is part of Port-channel 10
Hardware is Force10Eth, address is 00:01:e8:06:95:c0
Current address is 00:01:e8:06:95:c0
Interface index is 10910113
Port will not be disabled on partial SFM failure
Internet address is not set
MTU 1554 bytes, IP MTU 1500 bytes
LineSpeed 1000 Mbit, Mode full duplex, Slave
Flowcontrol rx on tx on
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of "show interface" counters 00:02:11
Queueing strategy: fifo
Input Statistics:
  132 packets, 16368 bytes
  0 Vlans
  0 64-byte pkts, 12 over 64-byte pkts, 120 over 127-byte pkts
  0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
  132 Multicasts, 0 Broadcasts
  0 runts, 0 giants, 0 throttles
  0 CRC, 0 overrun, 0 discarded
Output Statistics:
  136 packets, 16718 bytes, 0 underruns
  0 64-byte pkts, 15 over 64-byte pkts, 121 over 127-byte pkts
  0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
  136 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, 0.000% of line-rate
  Output 00.00 Mbits/sec, 0.000% of line-rate
Time since last interface status change: 00:02:14

Shows the status of this physical interface, and shows it is part of port channel 10.
Shows the speed of this physical interface. Also shows it is the slave of the GigE link.
Inspecting configuration of LAG 10 on ALPHA.

Alpha#show int port-channel 10
Port-channel 10 is up, line protocol is up
Created by LACP protocol
Hardware address is 00:01:e8:06:96:63, Current address is 00:01:e8:06:96:63
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 2/31(U) Gi 2/32(U) Gi 2/33(U)
ARP type: ARPA, ARP Timeout 04:00:00
Last clearing of “show interface” counters 00:04:09
Queueing strategy: fifo
Input Statistics:
  621 packets, 78732 bytes
  0 Vlans
  0 64-byte pkts, 18 over 64-byte pkts, 603 over 127-byte pkts
  0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
  621 Multicasts, 0 Broadcasts
  0 runts, 0 giants, 0 throttles
  0 CRC, 0 overrun, 0 discarded
Output Statistics:
  630 packets, 79284 bytes, 0 underruns
  0 64-byte pkts, 30 over 64-byte pkts, 600 over 127-byte pkts
  0 over 255-byte pkts, 0 over 511-byte pkts, 0 over 1023-byte pkts
  630 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: 00:03:38

Indicates the MAC address assigned to the LAG. This does NOT match any of the physical interface MAC addresses.

Confirms the number of links to bring up the LAG and that this is a switch port instead of a router port.

Confirms the total bandwidth for this LAG and which interfaces are active.
Using the **show lacp** command to verify LAG 10 status on ALPHA.

```
Alpha#show lacp 10
Port-channel 10 admin up, oper up, mode lacp
  Actor System ID: Priority 32768, Address 0001.e806.953e
  Partner System ID: Priority 32768, Address 0001.e809.c24a
  Actor Admin Key 10, Oper Key 10, Partner Oper Key 10
LACP LAG 10 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 2/31 is enabled, LACP is enabled and mode is lacp
  Actor Admin: State ACEHJLMP Key 10 Priority 32768
  Oper: State ACEGIKNP Key 10 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
  Oper: State ACEGIKNP Key 10 Priority 32768

Port Gi 2/32 is enabled, LACP is enabled and mode is lacp
  Actor Admin: State ACEHJLMP Key 10 Priority 32768
  Oper: State ACEGIKNP Key 10 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
  Oper: State ACEGIKNP Key 10 Priority 32768

Port Gi 2/33 is enabled, LACP is enabled and mode is lacp
  Actor Admin: State ACEHJLMP Key 10 Priority 32768
  Oper: State ACEGIKNP Key 10 Priority 32768
  Partner Admin: State BDFHJLMP Key 0 Priority 0
  Oper: State ACEGIKNP Key 10 Priority 32768
```

Interfaces participating in the LAG are included here.
Summary of the configuration on ALPHA

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

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
```
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
```

---

Also shows it is part of LAG 10.

Shows that this is a Layer 2 port.

Shows the speed of this physical interface.

Also shows it is the Master of the GigE link.
Using the `show interfaces port-channel` command to inspect LAG 10.

```
FTOS#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
```

Indicates the MAC address assigned to the LAG. This does NOT match any of the physical interface MAC addresses.

Confirms the number of links to bring up the LAG and that this is a switch port instead of a router port.

Confirms the total bandwidth for this LAG and which interfaces are active.
Using the `show lacp` command to inspect LAG status.

```
FTOS#show lacp 10
Port-channel 10 admin up, oper up, mode lacp
Actor   System ID: Priority 32768, Address 0001.e809.c24a
Partner System ID: Priority 32768, Address 0001.e806.953e
Actor Admin Key 10, Oper Key 10, Partner Oper Key 10
LACP LAG 10 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 3/21 is enabled, LACP is enabled and mode is lacp
Actor   Admin: State ACEHJLMP Key 10 Priority 32768
         Oper: State ACEGIKNP Key 10 Priority 32768
Partner Admin: State BDFHJLMP Key 0 Priority 0
              Oper: State ACEGIKNP Key 10 Priority 32768

Port Gi 3/22 is enabled, LACP is enabled and mode is lacp
Actor   Admin: State ACEHJLMP Key 10 Priority 32768
         Oper: State ACEGIKNP Key 10 Priority 32768
Partner Admin: State BDFHJLMP Key 0 Priority 0
              Oper: State ACEGIKNP Key 10 Priority 32768

Port Gi 3/23 is enabled, LACP is enabled and mode is lacp
Actor   Admin: State ACEHJLMP Key 10 Priority 32768
         Oper: State ACEGIKNP Key 10 Priority 32768
Partner Admin: State BDFHJLMP Key 0 Priority 0
              Oper: State ACEGIKNP Key 10 Priority 32768
FTOS#
```

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.
Layer 2

Layer 2 features are supported on platforms: E C S Z S4810

This chapter describes the following Layer 2 features:

• Managing the MAC Address Table
• MAC Learning Limit
• NIC Teaming
• Microsoft Clustering
• Configuring Redundant Pairs
• Restricting Layer 2 Flooding
• Far-end Failure Detection

Managing the MAC Address Table

FTOS provides the following management activities for the MAC address table:

• Clear the MAC Address Table
• Set the Aging Time for Dynamic Entries
• Configure a Static MAC Address
• Display the MAC Address Table

Clear the MAC Address Table

You may clear the MAC address table of dynamic entries:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear a MAC address table of dynamic entries.</td>
<td>\texttt{clear mac-address-table {dynamic</td>
<td>sticky} {address</td>
</tr>
</tbody>
</table>

• 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
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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disable MAC address aging for all dynamic entries.</td>
<td><code>mac-address-table aging-time 0</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Specify an aging time.</td>
<td><code>mac-address-table aging-time</code>&lt;br&gt;Range: 10-1000000</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify an aging time.</td>
<td><code>mac-address-table aging-time</code>&lt;br&gt;Range: 1-1000000</td>
<td>INTERFACE VLAN</td>
</tr>
</tbody>
</table>

**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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create a static MAC address entry in the MAC address table.</td>
<td><code>mac-address-table static</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
Display the MAC Address Table

To display the contents of the MAC address table:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display the contents of the MAC address table.</td>
<td>show mac-address-table [address</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>• address</td>
<td>aging-time [vlan vlan-id]] count</td>
<td></td>
</tr>
<tr>
<td>• aging-time</td>
<td>dynamic</td>
<td>interface</td>
</tr>
<tr>
<td>• count</td>
<td>displays the specified entry.</td>
<td></td>
</tr>
<tr>
<td>• dynamic</td>
<td>displays the configured aging-time.</td>
<td></td>
</tr>
<tr>
<td>• interface</td>
<td>displays only dynamic entries</td>
<td></td>
</tr>
<tr>
<td>• static</td>
<td>displays only static entries.</td>
<td></td>
</tr>
<tr>
<td>• vlan</td>
<td>displays only entries for the specified VLAN.</td>
<td></td>
</tr>
</tbody>
</table>

MAC Learning Limit

This section has the following sub-sections:

• mac learning-limit dynamic
• mac learning-limit mac-address-sticky
• mac learning-limit station-move
• Learning Limit Violation Actions
• Station Move Violation Actions
• Recovering from Learning Limit and Station Move Violations
• Per-VLAN MAC Learning Limit
MAC Address Learning Limit is a method of port security on Layer 2 port-channel and physical interfaces, and VLANs. It enables you to set an upper limit on the number of MAC addresses that 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, 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify the number of MAC addresses that the system can learn off a Layer 2 interface.</td>
<td><code>mac learning-limit address_limit</code></td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

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`

The MAC address table is stored on the Layer 2 FIB region of the CAM (and the Layer 2 ACL region on the E-Series). On the C-Series and S-Series the Layer 2 FIB region allocates space for static MAC address entries and dynamic MAC address entries (all MAC address entries on the E-Series are dynamic). When MAC Learning Limit is enabled, entries created on this port are static by default. When you configure the **dynamic** option, learned MAC addresses are stored in the dynamic region and are subject to aging. Entries created before this option is set are not affected.

**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 mac-address-sticky

Using sticky MAC addresses allows you to associate a specific port with MAC addresses from trusted devices. If sticky MAC is enabled, the specified port will retain any dynamically-learned addresses and prevent them from being transferred or learned on other ports.

If `mac-learning-limit` is configured and sticky MAC is enabled, all dynamically-learned addresses are converted to sticky MAC addresses for the selected port. Any new MAC addresses learned on this port will be converted to sticky MAC addresses.

To save all sticky MAC addresses into a configuration file that can be used as a startup configuration file, use the `write config` command. If the number of existing MAC addresses is fewer than the configured mac learn limit, any additional MAC addresses will be converted to sticky MACs on that interface. To remove all sticky MAC addresses from the running config file, disable sticky MAC and use the `write config` command.

When sticky mac is enabled on an interface, dynamically-learned MAC addresses will not age, even if `mac-learning-limit dynamic` is enabled. If `mac-learning-limit` and `mac-learning-limit dynamic` are configured and sticky MAC is disabled, any dynamically-learned MAC addresses will age.

mac learning-limit station-move

mac learning-limit station-move is available only on platforms: C S Z

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.

Learning Limit Violation Actions

Learning Limit Violation Actions are supported only on platforms: E S60 S4810.

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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generate a system log message when the MAC learning limit is exceeded.</td>
<td><code>learn-limit-violation log</code></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>Shut down the interface and generate a system log message when the MAC learning limit is exceeded.</td>
<td><code>learn-limit-violation shutdown</code></td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>
Station Move Violation Actions

Station Move Violation Actions are supported only on platforms: S-Series (S25/S50)

no-station-move is the default behavior. 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generate a system log message indicating a station move.</td>
<td>station-move-violation log</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>Shut down the first port to learn the MAC address.</td>
<td>station-move-violation shutdown-original</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>Shut down the second port to learn the MAC address.</td>
<td>station-move-violation shutdown-offending</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>Shut down both the first and second port to learn the MAC address.</td>
<td>station-move-violation shutdown-both</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

To display a list of interfaces configured with MAC learning limit or station move violation actions:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display a list of all of the interfaces configured with MAC learning limit or station move violation.</td>
<td>show mac learning-limit violate-action</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reset interfaces in ERR_Disabled state caused by a learning limit violation or station move violation.</td>
<td>mac learning-limit reset</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Reset interfaces in ERR_Disabled state caused by a learning limit violation.</td>
<td>mac learning-limit reset learn-limit-violation [interface</td>
<td>all]</td>
</tr>
<tr>
<td>Reset interfaces in ERR_Disabled state caused by a station move violation.</td>
<td>mac learning-limit reset station-move-violation [interface</td>
<td>all]</td>
</tr>
</tbody>
</table>

Note: Alternatively, you can reset the interface by shutting it down using the shutdown command and then reenabling it using the command no shutdown.
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 the following illustration, an Internet Exchange Point (IXP) connects multiple Internet Service Provider (ISP). An IXP can provide several types of services to its customers including public and private peering. Public peering means that all customers are connected to one VLAN. 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 the following illustration, 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.

ISP A, B, and C are all public peers through VLAN 10. In addition, ISP A and C are private peers on a separate VLAN, VLAN 20. Since the access ports for ISP A and C are members of multiple VLANs, Per-VLAN MAC Learning Limit can be applied to those ports.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure a MAC learning limit on a VLAN.</td>
<td><strong>mac learning-limit</strong> <em>limit</em> <strong>vlan</strong> <em>vlan-id</em></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>Display the MAC learning limit counters for a VLAN.</td>
<td><strong>show mac learning-limit</strong> [interface <em>slot/port</em> [vlan <em>vlan-id</em>]]</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>
NIC Teaming

NIC Teaming is available on the following platforms: C E S

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.

The following illustration 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 29-1. 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 29-2) and Port 0/5 is the failover port. When the NIC fails, the system automatically sends an ARP request for the gateway or host NIC to resolve the ARP and refresh the egress interface. When the ARP is resolved, the same MAC address is learned on the same port where the ARP is resolved.
(in the above example, this is Port 0/5 of the switch). To ensure the MAC address is disassociated with one port and re-associated with another port in the ARP table, 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 29-2. 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 29-3); 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 29-4).

Figure 29-3. Server Clustering: Multiple ARP Replies

![Figure 29-3. Server Clustering: Multiple ARP Replies](image)

Figure 29-4. Server Clustering: Failover and Balancing Not Preserved

![Figure 29-4. Server Clustering: Failover and Balancing Not Preserved](image)

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 `vlan-flooding` on the Dell Force10 switch at the time that the Microsoft cluster is configured (Figure 29-5).
As shown in Figure 29-5, 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 `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 29-5.  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 platforms: \[ E \ C \ S \ Z \]

Networks that employ switches that do not support Spanning Tree (STP) — for example, networks with Digital Subscriber Line Access Multiplexers (DSLAM) — cannot have redundant links between switches because they create switching loops (Figure 29-6). The Redundant Pairs feature enables you to create redundant links in networks that do not use STP by configuring backup interfaces for the interfaces on either side of the primary link.

Note: For details on STP, see Chapter 50, “Spanning Tree Protocol (STP),” on page 1003.

Assign a backup interface to an interface using the command `switchport backup`. The backup interface remains in down state until the primary fails, at which point it transitions to up state. If the primary interface fails, and later comes up, it becomes the backup interface for the redundant pair. FTOS supports Gigabit, 10-Gigabit, and 40-Gigabit interfaces as backup interfaces.

You must apply all other configurations to each interface in the redundant pair such that their configurations are identical, so that transition to the backup interface in the event of a failure is transparent to rest of the network.

Figure 29-6. Configuring Redundant Layer 2 Pairs without Spanning Tree

Redundant links create a switching loop. Without STP broadcast storms occurs.

Use backup interfaces to create redundant links in networks without STP.
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, 10-Gigabit, and 40-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 Chapter 23, Interfaces, Port Channel Interfaces on page 480 for more information.) If the interface is a member link of a LAG, the following primary/backup interfaces are also supported:

- primary interface is a physical interface, the backup interface can be a physical interface
- primary interface is a physical interface, the backup interface can be a static or dynamic LAG
- primary interface is a static or dynamic LAG, the backup interface can be a physical interface
- primary interface is a static or dynamic LAG, the backup interface can be a static or dynamic LAG

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 physical 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**

- You may not configure any interface to be a backup for more than one interface, no interface can have more than one backup, and a backup interface may not have a backup interface.
- Neither the active nor the backup interface may be a member of a LAG.
- The active and standby do not have to be of the same type (1G, 10G, etc).
- You may not enable any Layer 2 protocol on any interface of a redundant pair or to ports connected to them.

In Figure 29-7, interface 3/41 is a backup interface for 3/42, and 3/42 is in the down state, 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 29-7. CLI for Configuring Redundant Layer 2 Pairs without Spanning Tree**

```plaintext
FTOS#show interfaces port-channel brief
Codes: L - LACP Port-channel
LAG  Mode  Status       Uptime      Ports
1    L2    up           00:08:33    Te 0/0     (Up)
2    L2    up           00:00:02    Te 0/1     (Up)

FTOS#configure
FTOS(conf)#interface port-channel 1
FTOS(conf-if-po-1)#switchport backup interface GigabitEthernet 3/42
Fig 29-8: CLI for Redundant Pair in Port-channel on S4810

FTOS(conf-if-range-gi-3/41-42)#do show ip int brief | find 3/41
GigabitEthernet 3/41 unassigned YES Manual up
GigabitEthernet 3/42 unassigned NO Manual up
down

FTOS#configure
FTOS(conf)#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
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#configure
FTOS(conf-if-gi-3/41)#shutdown

00:24:55: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: 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 YES Manual up
down
GigabitEthernet 3/42 unassigned NO Manual administratively down
down

FTOS(conf-if-range-gi-3/41-42)#do show ip int brief | find 3/41
GigabitEthernet 3/41 unassigned YES Manual up
down
GigabitEthernet 3/42 unassigned NO Manual administratively down
down

```

**Figure 29-8. CLI for Redundant Pair in Port-channel on S4810**

```plaintext
FTOS#show interfaces switchport backup
Interface Status Paired Interface     Status
Port-channel 1 Active Port-channel 2   Active
Port-channel 2 Active Port-channel 1   Active
```
Restricting Layer 2 Flooding

Restricting Layer 2 Flooding is supported only on platform: E

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 the following example, 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 29-9. 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 (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. FEFD can be enabled globally or locally on an interface basis. Disabling the global FEFD configuration does not disable the interface configuration.

**Figure 29-10. Configuring Far-end Failure Detection**

```
FTOS(conf-if-gi-4/0)#show config
!
interface GigabitEthernet 4/0
  no ip address
  switchport
  fefd
  no shutdown

FTOS(conf-if-gi-1/0)#show config
!
interface GigabitEthernet 1/0
  no ip address
  switchport
  fefd
  no shutdown
```

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 has two operational modes, Normal and Aggressive. When Normal mode is enabled on an interface and a far-end failure is detected, no intervention is required to reset the interface to bring it back to an FEFD operational state. When Aggressive mode is enabled on an interface in the same state, manual intervention is required to reset the interface.

FEFD enabled systems (comprised of one or more interfaces) will automatically switch between four different states: Idle, Unknown, Bi-directional, and Err-disabled.

1. An interface on which FEFD is not configured is in Normal mode by default.
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 29-1. State Changes When Configuring FEFD

<table>
<thead>
<tr>
<th>Local Event</th>
<th>Mode</th>
<th>Local State</th>
<th>Remote State</th>
<th>Local Admin Status</th>
<th>Local Protocol Status</th>
<th>Remote Admin Status</th>
<th>Remote Protocol Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shutdown</td>
<td>Normal</td>
<td>Admin Shutdown</td>
<td>Unknown</td>
<td>Down</td>
<td>Down</td>
<td>Up</td>
<td>Down</td>
</tr>
<tr>
<td>Shutdown</td>
<td>Aggressive</td>
<td>Admin Shutdown</td>
<td>Err-disabled</td>
<td>Up</td>
<td>Down</td>
<td>Up</td>
<td>Down</td>
</tr>
<tr>
<td>FEFD enable</td>
<td>Normal</td>
<td>Bi-directional</td>
<td>Bi-directional</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
</tr>
<tr>
<td>FEFD enable</td>
<td>Aggressive</td>
<td>Bi-directional</td>
<td>Bi-directional</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
</tr>
<tr>
<td>FEFD + FEFD disable</td>
<td>Normal</td>
<td>Locally disabled</td>
<td>Unknown</td>
<td>Up</td>
<td>Down</td>
<td>Up</td>
<td>Down</td>
</tr>
<tr>
<td>FEFD + FEFD disable</td>
<td>Aggressive</td>
<td>Locally disabled</td>
<td>Err-disabled</td>
<td>Up</td>
<td>Down</td>
<td>Up</td>
<td>Down</td>
</tr>
<tr>
<td>Link Failure</td>
<td>Normal</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Up</td>
<td>Down</td>
<td>Up</td>
<td>Down</td>
</tr>
<tr>
<td>Link Failure</td>
<td>Aggressive</td>
<td>Err-disabled</td>
<td>Err-disabled</td>
<td>Up</td>
<td>Down</td>
<td>Up</td>
<td>Down</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Setup two or more connected interfaces for Layer 2 or Layer 3 use</td>
<td><code>ip address ip address, switchport</code></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>2</td>
<td>Activate the necessary ports administratively</td>
<td><code>no shutdown</code></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>3</td>
<td>Enable fefd globally</td>
<td>`fefd {interval</td>
<td>mode}`</td>
</tr>
</tbody>
</table>

Entering the `show fefd` command in EXEC privilege mode displays information about the state of each interface.

Figure 29-11. Show FEFD global outputs

```
FTOS#show fefd
FEFD is globally 'ON', interval is 3 seconds, mode is 'Normal'.

INTERFACE     MODE           INTERVAL (second)   STATE
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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Setup two or more connected interfaces for Layer 2 or Layer 3 use</td>
<td><code>ip address ip address, switchport</code></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>2</td>
<td>Activate the necessary ports administratively</td>
<td><code>no shutdown</code></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>3</td>
<td>Enable FEFD on each interface</td>
<td>`fefd {disable</td>
<td>interval</td>
</tr>
</tbody>
</table>

**Figure 29-12. 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 29-13. Debug FEFD events display**

FTOS#debug fefd events
FTOS#config
FTOS(config)#int gi 1/0
FTOS(config-if-gi-1/0)#shutdown
2w1d22h: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Gi 1/0
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 29-14. 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 29-15. FEFD state change during an RPM failover**

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Event Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>02-05-2009</td>
<td>12:40:38</td>
<td>Local7.Debug 10.16.151.12 Feb 5 07:06:19: %RPM1-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Gi 0/45</td>
</tr>
<tr>
<td>02-05-2009</td>
<td>12:40:38</td>
<td>Local7.Debug 10.16.151.12 Feb 5 07:06:19: %RPM1-P:CP %FEFD-5-FEFD-BIDIRECTION-LINK-DETECTED: Interface Gi 0/45 has bidirectional link with its peer</td>
</tr>
</tbody>
</table>
Link Layer Discovery Protocol (LLDP)

Link Layer Discovery Protocol (LLDP) is supported only on platforms:

This chapter contains the following sections:

- 802.1AB (LLDP) Overview
- TIA-1057 (LLDP-MED) Overview
- Configuring LLDP

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 30-1 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
TLVs are encapsulated in a frame called an LLDP Data Unit (LLDPDU) (Figure 30-2), 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>
<thead>
<tr>
<th>Type</th>
<th>TLV</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>End of LLDPDU</td>
<td>Marks the end of an LLDPDU</td>
</tr>
<tr>
<td>1</td>
<td>Chassis ID</td>
<td>An administratively assigned name that identifies the LLDP agent</td>
</tr>
<tr>
<td>2</td>
<td>Port ID</td>
<td>An administratively assigned name that identifies a port through which TLVs are sent and received</td>
</tr>
<tr>
<td>3</td>
<td>Time to Live</td>
<td>A value that tells the receiving agent how long the information contained in the TLV Value field is valid</td>
</tr>
<tr>
<td>—</td>
<td>Optional</td>
<td>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.</td>
</tr>
</tbody>
</table>
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 30-2.

Organizationally Specific TLVs

Organizationally specific TLVs can be defined by a professional organization or a vendor. They have two mandatory fields (Figure 30-3) in addition to the basic TLV fields (Figure 30-1):

- 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 30-3. Organizationally Specific TLV
IEEE Organizationally Specific TLVs

Eight TLV types have been defined by the IEEE 802.1 and 802.3 working groups (Table 30-2) 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 30-2. Optional TLV Types

<table>
<thead>
<tr>
<th>Type</th>
<th>TLV</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Optional TLVs</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Port description</td>
<td>A user-defined alphanumeric string that describes the port. FTOS does not currently support this TLV.</td>
</tr>
<tr>
<td>5</td>
<td>System name</td>
<td>A user-defined alphanumeric string that identifies the system.</td>
</tr>
<tr>
<td>6</td>
<td>System description</td>
<td>A user-defined alphanumeric string that describes the system</td>
</tr>
<tr>
<td>7</td>
<td>System capabilities</td>
<td>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</td>
</tr>
<tr>
<td>8</td>
<td>Management address</td>
<td>Indicates the network address of the management interface. FTOS does not currently support this TLV.</td>
</tr>
</tbody>
</table>

IEEE 802.1 Organizationally Specific TLVs

<table>
<thead>
<tr>
<th>Type</th>
<th>TLV</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>127</td>
<td>Port-VLAN ID</td>
<td>On Dell Force10 systems, indicates the untagged VLAN to which a port belongs</td>
</tr>
<tr>
<td>127</td>
<td>Port and Protocol VLAN ID</td>
<td>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)</td>
</tr>
<tr>
<td>127</td>
<td>VLAN Name</td>
<td>Indicates the user-defined alphanumeric string that identifies the VLAN. This TLV is supported on C-Series only.</td>
</tr>
<tr>
<td>127</td>
<td>Protocol Identity</td>
<td>Indicates the protocols that the port can process. FTOS does not currently support this TLV.</td>
</tr>
</tbody>
</table>

IEEE 802.3 Organizationally Specific TLVs

<table>
<thead>
<tr>
<th>Type</th>
<th>TLV</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>127</td>
<td>MAC/PHY Configuration/Status</td>
<td>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.</td>
</tr>
<tr>
<td>127</td>
<td>Power via MDI</td>
<td>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.</td>
</tr>
<tr>
<td>127</td>
<td>Link Aggregation</td>
<td>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.</td>
</tr>
<tr>
<td>127</td>
<td>Maximum Frame Size</td>
<td>Indicates the maximum frame size capability of the MAC and PHY.</td>
</tr>
</tbody>
</table>
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 30-3* describes the five types of TIA-1057 Organizationally Specific TLVs.
Table 30-3. TIA-1057 (LLDP-MED) Organizationally Specific TLVs

<table>
<thead>
<tr>
<th>Type</th>
<th>Sub-type</th>
<th>TLV</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>127</td>
<td>1</td>
<td>LLDP-MED Capabilities</td>
<td>Indicates:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• whether the transmitting device supports LLDP-MED</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• what LLDP-MED TLVs it supports</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• LLDP device class</td>
</tr>
<tr>
<td>127</td>
<td>2</td>
<td>Network Policy</td>
<td>Indicates the application type, VLAN ID, Layer 2 Priority, and DSCP value</td>
</tr>
<tr>
<td>127</td>
<td>3</td>
<td>Location Identification</td>
<td>Indicates the physical location of the device expressed in one of three possible formats:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Coordinate Based LCI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Civic Address LCI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Emergency Call Services ELIN</td>
</tr>
<tr>
<td>127</td>
<td>4</td>
<td>Extended Power via MDI</td>
<td>Indicates power requirements, priority, and power status</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inventory Management TLVs</td>
<td>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.</td>
</tr>
<tr>
<td>127</td>
<td>5</td>
<td>Inventory - Hardware Revision</td>
<td>Indicates the hardware revision of the LLDP-MED device</td>
</tr>
<tr>
<td>127</td>
<td>6</td>
<td>Inventory - Firmware Revision</td>
<td>Indicates the firmware revision of the LLDP-MED device</td>
</tr>
<tr>
<td>127</td>
<td>7</td>
<td>Inventory - Software Revision</td>
<td>Indicates the software revision of the LLDP-MED device</td>
</tr>
<tr>
<td>127</td>
<td>8</td>
<td>Inventory - Serial Number</td>
<td>Indicates the device serial number of the LLDP-MED device</td>
</tr>
<tr>
<td>127</td>
<td>9</td>
<td>Inventory - Manufacturer Name</td>
<td>Indicates the manufacturer of the LLDP-MED device</td>
</tr>
<tr>
<td>127</td>
<td>10</td>
<td>Inventory - Model Name</td>
<td>Indicates the model of the LLDP-MED device</td>
</tr>
<tr>
<td>127</td>
<td>11</td>
<td>Inventory - Asset ID</td>
<td>Indicates a user specified device number to manage inventory</td>
</tr>
<tr>
<td>127</td>
<td>12-255</td>
<td>Reserved</td>
<td>—</td>
</tr>
</tbody>
</table>

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 30-4), each bit represents an LLDP-MED capability (Table 30-4).
- The possible values of the LLDP-MED Device Type is listed in Table 30-5. 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.
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

Table 30-4. FTOS LLDP-MED Capabilities

<table>
<thead>
<tr>
<th>Bit Position</th>
<th>TLV</th>
<th>FTOS Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>LLDP-MED Capabilities</td>
<td>Yes</td>
</tr>
<tr>
<td>1</td>
<td>Network Policy</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>Location Identification</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>Extended Power via MDI-PSE</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>Extended Power via MDI-PD</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>Inventory</td>
<td>No</td>
</tr>
<tr>
<td>6-15</td>
<td>reserved</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 30-5. LLDP-MED Device Types

<table>
<thead>
<tr>
<th>Value</th>
<th>Device Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Type Not Defined</td>
</tr>
<tr>
<td>1</td>
<td>Endpoint Class 1</td>
</tr>
<tr>
<td>2</td>
<td>Endpoint Class 2</td>
</tr>
<tr>
<td>3</td>
<td>Endpoint Class 3</td>
</tr>
<tr>
<td>4</td>
<td>Network Connectivity</td>
</tr>
<tr>
<td>5-255</td>
<td>Reserved</td>
</tr>
</tbody>
</table>
The application type is represented by an integer (the Type integer in Table 30-6), 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 650).

**Note:** With regard to Table 30-6, 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>
<thead>
<tr>
<th>Type</th>
<th>Application</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Reserved</td>
<td>—</td>
</tr>
<tr>
<td>1</td>
<td>Voice</td>
<td>Specify this application type for dedicated IP telephony handsets and other appliances supporting interactive voice services.</td>
</tr>
<tr>
<td>2</td>
<td>Voice Signaling</td>
<td>Specify this application type only if voice control packets use a separate network policy than voice data.</td>
</tr>
<tr>
<td>3</td>
<td>Guest Voice</td>
<td>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.</td>
</tr>
<tr>
<td>4</td>
<td>Guest Voice Signaling</td>
<td>Specify this application type only if guest voice control packets use a separate network policy than voice data.</td>
</tr>
<tr>
<td>5</td>
<td>Softphone Voice</td>
<td>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.</td>
</tr>
<tr>
<td>6</td>
<td>Video Conferencing</td>
<td>Specify this application type for dedicated video conferencing and other similar appliances supporting real-time interactive video.</td>
</tr>
<tr>
<td>7</td>
<td>Streaming Video</td>
<td>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.</td>
</tr>
<tr>
<td>8</td>
<td>Video Signaling</td>
<td>Specify this application type only if video control packets use a separate network policy than video data.</td>
</tr>
<tr>
<td>9-255</td>
<td>Reserved</td>
<td>—</td>
</tr>
</tbody>
</table>

**Figure 30-5. 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 30-6. Extended Power via MDI TLV**

<table>
<thead>
<tr>
<th>TLV Type (137)</th>
<th>TLV Length (27)</th>
<th>Organizationaly Unique ID (00-12-86)</th>
<th>Organizationaly Defined Sub-type (4)</th>
<th>Power Type (0)</th>
<th>Power Source (1)</th>
<th>Power Priority (2)</th>
<th>Power Value (130)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 bits</td>
<td>9 bits</td>
<td>3 octets</td>
<td>1 octet</td>
<td>2 bits</td>
<td>2 bits</td>
<td>4 bits</td>
<td>2 octets</td>
</tr>
</tbody>
</table>

**Configuring LLDP**

Configuring LLDP is a two-step process:

1. Enable LLDP globally.
2. Advertise TLVs out of an interface.

**Related Configuration Tasks**

- Viewing the LLDP Configuration
- Viewing Information Advertised by Adjacent LLDP Agents
- Configuring LLDPDU Intervals
- Configuring Transmit and Receive Mode
- Configuring a Time to Live
- Debugging LLDP
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 the CONFIGURATION mode and INTERFACE mode.

- Configurations made at the CONFIGURATION level are global, that is, they affect all interfaces on the system.
- Configurations made at the INTERFACE level affect only the specific interface, and they override CONFIGURATION level configurations.
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter Protocol LLDP mode.</td>
<td>protocol lldp</td>
<td>CONFIGURATION or INTERFACE</td>
</tr>
<tr>
<td>2</td>
<td>Enable LLDP.</td>
<td>no disable</td>
<td>PROTOCOL LLDP</td>
</tr>
</tbody>
</table>

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`. 

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:

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter LLDP mode.</td>
<td>protocol lldp</td>
<td>CONFIGURATION or INTERFACE</td>
</tr>
<tr>
<td>2</td>
<td>Advertise one or more TLVs. Include the keyword for each TLV you want to advertise.</td>
<td>advertise {management-tlv</td>
<td>dot1-tlv</td>
</tr>
</tbody>
</table>

Note: The vlan-name command is supported on C-Series and S-Series only.

In Figure 30-8, LLDP is enabled globally. R1 and R2 are transmitting periodic LLDPDUs that contain management, 802.1, and 802.3 TLVs.
Viewing the LLDP Configuration

Display the LLDP configuration using the command `show config` in either the CONFIGURATION or INTERFACE mode, as shown in Figure 30-9 and Figure 30-10, respectively.

**Figure 30-9. 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)#
```
Viewing Information Advertised by Adjacent LLDP Agents

Display brief information about adjacent devices using the command `show lldp neighbors`, as shown in Figure 30-11. Display all of the information that neighbors are advertising using the command `show lldp neighbors detail`, as shown in Figure 30-12.

```
R1(conf-if-gi-1/31-lldp)#show config
!
protocol lldp
R1(conf-if-gi-1/31-lldp)#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
```
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 `hello` command (Figure 30-13).
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 30-14).
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 multiplier command. Return to the default multiplier value using the no multiplier command (Figure 30-15).
Figure 30-15. 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 30-16. debug lldp detail—LLDPDU Packet Dissection
FTOS supports all IEEE 802.1AB MIB objects.

- **Table 30-7** lists the objects associated with received and transmitted TLVs.
- **Table 30-8** lists the objects associated with the LLDP configuration on the local agent.
- **Table 30-9** lists the objects associated with IEEE 802.1AB Organizationally Specific TLVs.
- **Table 30-10** lists the objects associated with received and transmitted LLDP-MED TLVs.
## Table 30-7. LLDP Configuration MIB Objects

<table>
<thead>
<tr>
<th>MIB Object Category</th>
<th>LLDP Variable</th>
<th>LLDP MIB Object</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LLDP Configuration</td>
<td>adminStatus</td>
<td>lldpPortConfigAdminStatus</td>
<td>Whether the local LLDP agent is enabled for transmit, receive, or both</td>
</tr>
<tr>
<td></td>
<td>msgTxHold</td>
<td>lldpMessageTxHoldMultiplier</td>
<td>Multiplier value</td>
</tr>
<tr>
<td></td>
<td>msgTxInterval</td>
<td>lldpMessageTxInterval</td>
<td>Transmit Interval value</td>
</tr>
<tr>
<td></td>
<td>rxInfoTTL</td>
<td>lldpRxInfoTTL</td>
<td>Time to Live for received TLVs</td>
</tr>
<tr>
<td></td>
<td>txInfoTTL</td>
<td>lldpTxInfoTTL</td>
<td>Time to Live for transmitted TLVs</td>
</tr>
<tr>
<td>Basic TLV Selection</td>
<td>mIBBasicTLVsTxEnable</td>
<td>lldpPortConfigTLVsTxEnable</td>
<td>Indicates which management TLVs are enabled for system ports</td>
</tr>
<tr>
<td></td>
<td>mIBMgmtAddrInstanceTxEnable</td>
<td>lldpManAddrPortsTxEnable</td>
<td>The management addresses defined for the system and the ports through which they are enabled for transmission</td>
</tr>
<tr>
<td>LLDP Statistics</td>
<td>statsAgeoutsTotal</td>
<td>lldpStatsRxPortAgeoutsTotal</td>
<td>Total number of times that a neighbors information is deleted on the local system due to an rxInfoTTL timer expiration</td>
</tr>
<tr>
<td></td>
<td>statsFramesDiscardedTotal</td>
<td>lldpStatsRxPortFramesDiscardedTotal</td>
<td>Total number of LLDP frames received then discarded</td>
</tr>
<tr>
<td></td>
<td>statsFramesInErrorsTotal</td>
<td>lldpStatsRxPortFramesErrors</td>
<td>Total number of LLDP frames received on a port with errors</td>
</tr>
<tr>
<td></td>
<td>statsFramesInTotal</td>
<td>lldpStatsRxPortFramesTotal</td>
<td>Total number of LLDP frames received through the port</td>
</tr>
<tr>
<td></td>
<td>statsFramesOutTotal</td>
<td>lldpStatsTxCPortFramesTotal</td>
<td>Total number of LLDP frames transmitted through the port</td>
</tr>
<tr>
<td></td>
<td>statsTLVsDiscardedTotal</td>
<td>lldpStatsRxPortTLVsDiscardedTotal</td>
<td>Total number of TLVs received then discarded</td>
</tr>
<tr>
<td></td>
<td>statsTLVsUnrecognizedTotal</td>
<td>lldpStatsRxPortTLVsUnrecognizedTotal</td>
<td>Total number of all TLVs the local agent does not recognize</td>
</tr>
</tbody>
</table>

## Table 30-8. LLDP System MIB Objects

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Name</th>
<th>TLV Variable</th>
<th>System</th>
<th>LLDP MIB Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Chassis ID</td>
<td>chassis ID subtype</td>
<td>Local</td>
<td>lldpLocChassisIdSubtype</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemChassisIdSubtype</td>
</tr>
<tr>
<td></td>
<td></td>
<td>chassid ID</td>
<td>Local</td>
<td>lldpLocChassisId</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemChassisId</td>
</tr>
</tbody>
</table>
## Table 30-8. LLDP System MIB Objects

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Name</th>
<th>TLV Variable</th>
<th>System</th>
<th>LLDP MIB Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Port ID</td>
<td>port subtype</td>
<td>Local</td>
<td>lldpLocPortIdSubtype</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemPortIdSubtype</td>
</tr>
<tr>
<td></td>
<td></td>
<td>port ID</td>
<td>Local</td>
<td>lldpLocPortId</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemPortId</td>
</tr>
<tr>
<td>4</td>
<td>Port Description</td>
<td>port description</td>
<td>Local</td>
<td>lldpLocPortDesc</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemPortDesc</td>
</tr>
<tr>
<td>5</td>
<td>System Name</td>
<td>system name</td>
<td>Local</td>
<td>lldpLocSysName</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemSysName</td>
</tr>
<tr>
<td>6</td>
<td>System Description</td>
<td>system description</td>
<td>Local</td>
<td>lldpLocSysDesc</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemSysDesc</td>
</tr>
<tr>
<td>7</td>
<td>System Capabilities</td>
<td>system capabilities</td>
<td>Local</td>
<td>lldpLocSysCapSupported</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemSysCapSupported</td>
</tr>
<tr>
<td>8</td>
<td>Management Address</td>
<td>enabled capabilities</td>
<td>Local</td>
<td>lldpLocSysCapEnabled</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemSysCapEnabled</td>
</tr>
<tr>
<td></td>
<td></td>
<td>management address length</td>
<td>Local</td>
<td>lldpLocManAddrLen</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemManAddrLen</td>
</tr>
<tr>
<td></td>
<td></td>
<td>management address subtype</td>
<td>Local</td>
<td>lldpLocManAddrSubtype</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemManAddrSubtype</td>
</tr>
<tr>
<td></td>
<td></td>
<td>management address</td>
<td>Local</td>
<td>lldpLocManAddr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemManAddr</td>
</tr>
<tr>
<td></td>
<td></td>
<td>interface numbering subtype</td>
<td>Local</td>
<td>lldpLocManAddrIfSubtype</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemManAddrIfSubtype</td>
</tr>
<tr>
<td></td>
<td></td>
<td>interface number</td>
<td>Local</td>
<td>lldpLocManAddrIfId</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemManAddrIfId</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OID</td>
<td>Local</td>
<td>lldpLocManAddrOID</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpRemManAddrOID</td>
</tr>
</tbody>
</table>

## Table 30-9. LLDP 802.1 Organizationally Specific TLV MIB Objects

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Name</th>
<th>TLV Variable</th>
<th>System</th>
<th>LLDP MIB Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>127</td>
<td>Port-VLAN ID</td>
<td>PVID</td>
<td>Local</td>
<td>lldpXdot1LocPortVlanId</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXdot1RemPortVlanId</td>
</tr>
</tbody>
</table>
### Table 30-9. LLDP 802.1 Organizationally Specific TLV MIB Objects

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>TLV Name</th>
<th>TLV Variable</th>
<th>System</th>
<th>LLDP MIB Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>127</td>
<td>Port and Protocol VLAN ID</td>
<td>port and protocol VLAN supported</td>
<td>Local</td>
<td>lldpXdot1LocProtoVlanSupported</td>
</tr>
<tr>
<td></td>
<td></td>
<td>port and protocol VLAN enabled</td>
<td>Local</td>
<td>lldpXdot1LocProtoVlanEnabled</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPVID</td>
<td>Local</td>
<td>lldpXdot1LocProtoVlanId</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXdot1RemProtoVlanSupported</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXdot1RemProtoVlanEnabled</td>
</tr>
<tr>
<td></td>
<td>VLAN Name</td>
<td>VID</td>
<td>Local</td>
<td>lldpXdot1LocVlanId</td>
</tr>
<tr>
<td>127</td>
<td>VLAN name length</td>
<td></td>
<td>Local</td>
<td>lldpXdot1LocVlanName</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXdot1RemVlanName</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Local</td>
<td>lldpXdot1LocVlanName</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXdot1RemVlanName</td>
</tr>
</tbody>
</table>

### Table 30-10. LLDP-MED System MIB Objects

<table>
<thead>
<tr>
<th>TLV Sub-Type</th>
<th>TLV Name</th>
<th>TLV Variable</th>
<th>System</th>
<th>LLDP-MED MIB Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LLDP-MED</td>
<td>LLDP-MED Capabilities</td>
<td>Local</td>
<td>lldpXMedPortCapSupported</td>
</tr>
<tr>
<td></td>
<td>Capabilities</td>
<td></td>
<td></td>
<td>lldpXMedPortConfigTLVsTxEnable</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMedRemCapSupported,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>lldpXMedRemConfigTLVs.TxEnable</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Local</td>
<td>lldpXMedLocDeviceClass</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMedRemDeviceClass</td>
</tr>
<tr>
<td>TLV Sub-Type</td>
<td>TLV Name</td>
<td>TLV Variable</td>
<td>System</td>
<td>LLDP-MED MIB Object</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>-------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>2</td>
<td>Network Policy</td>
<td>Application Type</td>
<td>Local</td>
<td>lldpXMEdLocMediaPolicyAppType</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemMediaPolicyAppType</td>
</tr>
<tr>
<td></td>
<td>Unknown Policy Flag</td>
<td>Local</td>
<td>lldpXMEdLocMediaPolicyUnknown</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemMediaPolicyUnknown</td>
</tr>
<tr>
<td></td>
<td>Tagged Flag</td>
<td></td>
<td>Local</td>
<td>lldpXMEdLocMediaPolicyTagged</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemMediaPolicyTagged</td>
</tr>
<tr>
<td></td>
<td>VLAN ID</td>
<td></td>
<td>Local</td>
<td>lldpXMEdLocMediaPolicyVlanID</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemMediaPolicyVlanID</td>
</tr>
<tr>
<td></td>
<td>L2 Priority</td>
<td></td>
<td>Local</td>
<td>lldpXMEdLocMediaPolicyPriority</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemMediaPolicyPriority</td>
</tr>
<tr>
<td></td>
<td>DSCP Value</td>
<td></td>
<td>Local</td>
<td>lldpXMEdLocMediaPolicyDscp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemMediaPolicyDscp</td>
</tr>
<tr>
<td>3</td>
<td>Location Identifier</td>
<td>Location Data Format</td>
<td>Local</td>
<td>lldpXMEdLocLocationSubtype</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemLocationSubtype</td>
</tr>
<tr>
<td></td>
<td>Location ID Data</td>
<td></td>
<td>Local</td>
<td>lldpXMEdLocLocationInfo</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemLocationInfo</td>
</tr>
<tr>
<td>TLV Sub-Type</td>
<td>TLV Name</td>
<td>TLV Variable</td>
<td>System</td>
<td>LLDP-MED MIB Object</td>
</tr>
<tr>
<td>--------------</td>
<td>----------</td>
<td>--------------</td>
<td>--------</td>
<td>--------------------</td>
</tr>
<tr>
<td>4</td>
<td>Extended Power via MDI</td>
<td>Power Device Type</td>
<td>Local</td>
<td>lldpXMEdLocXPoEDeviceType</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemXPoEDeviceType</td>
</tr>
<tr>
<td></td>
<td>Power Source</td>
<td></td>
<td>Local</td>
<td>lldpXMEdLocXPoEPSEPSource, lldpXMEdLocXPoEPDPowerSource</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemXPoEPSEPSource, lldpXMEdRemXPoEPDPowerSource</td>
</tr>
<tr>
<td></td>
<td>Power Priority</td>
<td></td>
<td>Local</td>
<td>lldpXMEdLocXPoEPDPowerPriority, lldpXMEdLocXPoEPSEPortPDPriority</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemXPoEPSEPowerPriority, lldpXMEdRemXPoEPDPowerPriority</td>
</tr>
<tr>
<td></td>
<td>Power Value</td>
<td></td>
<td>Local</td>
<td>lldpXMEdLocXPoEPSEPortPowerAv, lldpXMEdLocXPoEPDPowerReq</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Remote</td>
<td>lldpXMEdRemXPoEPSEPowerAv, lldpXMEdRemXPoEPDPowerReq</td>
</tr>
</tbody>
</table>
Multicast Source Discovery Protocol (MSDP)

Multicast Source Discovery Protocol (MSDP) is supported on platform E and S4810.

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 31-1) 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 31-1. 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 31-2. MSDP SA Message Format**

<table>
<thead>
<tr>
<th>Source Port</th>
<th>Dest. Port</th>
<th>Seq. Number</th>
<th>Ack. Number</th>
<th>Offset</th>
<th>Reserved</th>
<th>Flags</th>
<th>Window</th>
<th>Checksum</th>
<th>Urgent</th>
<th>Options</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Anycast RP**

Using Multicast Source Discovery Protocol (MSDP), Anycast RP provides load sharing and redundancy in Protocol Independent Multicast sparse mode (PIM-SM) networks. Anycast RP allows two or more rendezvous points (RPs) to share the load for source registration and the ability to act as hot backup routers for each other.

Anycast RP allows two or more RPs to be configured with the same IP address on loopback interfaces. The Anycast RP loopback address are configured with a 32-bit mask, making it a host address. All downstream routers are configured to know that the Anycast RP loopback address is the IP address of their local RP. IP routing automatically selects the closest RP for each source and receiver. Assuming that the sources are evenly spaced around the network, an equal number of sources register with each RP. Consequently, all the RPs in the network share the process of registering the sources equally. Since a source may register with one RP and receivers may join to a different RP, a method is needed for the RPs to exchange information about active sources. This information exchange is done with MSDP.

With Anycast RP, all the RPs are configured to be MSDP peers of each other. When a source registers with one RP, an SA message is sent to the other RPs informing them that there is an active source for a particular multicast group. The result is that each RP is aware of the active sources in the area of the other RPs. If any of the RPs fail, IP routing converges and one of the RPs becomes the active RP in more than one area. New sources register with the backup RP. Receivers join toward the new RP and connectivity is maintained.

**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 31-5 and MSDP Sample Configurations on page 686 show the OSPF-BGP configuration used in this chapter for MSDP. Otherwise, see Chapter 34, Open Shortest Path First (OSPFv2) and Chapter 9, Border Gateway Protocol.

2. Configure PIM-SM within each EGP routing domain.
   
   Figure 31-5 and MSDP Sample Configurations show the PIM-SM configuration in this chapter for MSDP. Otherwise, see Chapter 35, “PIM Sparse-Mode (PIM-SM),” on page 775.

3. Enable MSDP. See page 670.

4. Peer the RPs in each routing domain with each other. See Enable MSDP.

Related Configuration Tasks

- Enable MSDP
- Manage the Source-active Cache
- Accept Source-active Messages that fail the RFP Check
- Limit the Source-active Messages from a Peer
- Prevent MSDP from Caching a Local Source
- Prevent MSDP from Caching a Remote Source
- Prevent MSDP from Advertising a Local Source
- Terminate a Peership
- Clear Peer Statistics
- Clear Peer Statistics
- Debug MSDP
- MSDP with Anycast RP
- MSDP Sample Configurations
Multicast Source Discovery Protocol (MSDP)

Figure 31.3 Configuring Interfaces for MSDP

```
interface GigabitEthernet 1/1
  ip pim sparse-mode
  ip address 10.11.3.1/24
  no shutdown

interface GigabitEthernet 1/2
  ip pim sparse-mode
  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 address 192.168.0.1/32
  no shutdown

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

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 address 192.168.0.3/32
  no shutdown

interface GigabitEthernet 4/1
  ip pim sparse-mode
  ip address 10.11.5.1/24
  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
```

PC 1: 10.11.3.2/24
PC 2: 10.11.4.2/24
PC 3: 10.11.5.2/24
R1
R2
R3
R4

Figure 31-4. Configuring OSPF and BGP for MSDP
Figure 31-5. Configuring PIM in Multiple Routing Domains

```
R1
  +----+----+
  | 4/1 | 4/2 |
  +----+----+
  | 1/1  | 1/2  |
  +----+----+
R2
  +----+----+
  | 3/1  | 3/2  |
  +----+----+
R3
  +----+----+
  | 2/1  | 2/2  |
  +----+----+
R4
  +----+----+
  | 1/1  | 1/2  |
  +----+----+
  +----+----+
  | 2/3  | 2/4  |
  +----+----+

PC 1
  +----+----+
  | 1/1  | 1/2  |
  +----+----+

PC 2
  +----+----+
  | 2/1  | 2/2  |
  +----+----+

Source: 239.0.0.1
Receiver: 239.0.0.1

AS 100
  +----+----+
  | 1/1  | 1/2  |
  +----+----+

AS 200
  +----+----+
  | 1/1  | 1/2  |
  +----+----+

ip multicast-routing
ip pim rp-address 192.168.0.1 group-address 224.0.0.0/4
PIM + IGMP
ip pim rp-address 192.168.0.3 group-address 224.0.0.0/4
```
Multicast Source Discovery Protocol (MSDP)
Enable MSDP

Enable MSDP by peering RPs in different administrative domains.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enable MSDP.</td>
<td>ip multicast-msdp</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Peer PIM systems in different administrative domains.</td>
<td>ip msdp peer connect-source</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**Figure 31-7. Configuring an MSDP Peer**

```bash
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
```

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>View details about a peer.</td>
<td>show ip msdp peer</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

**Figure 31-8. Displaying Details about a Peer**

```bash
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

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>View the SA cache.</td>
<td>show ip msdp sa-cache</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

Figure 31-9. Displaying the MSDP Source-active Cache

<table>
<thead>
<tr>
<th>GroupAddr</th>
<th>SourceAddr</th>
<th>RPAddr</th>
<th>LearnedFrom</th>
<th>Expire UpTime</th>
</tr>
</thead>
<tbody>
<tr>
<td>239.0.0.1</td>
<td>10.11.4.2</td>
<td>192.168.0.1</td>
<td>192.168.0.1</td>
<td>76 00:10:44</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limit the number of sources that can be stored in the SA cache.</td>
<td>show ip msdp sa-limit</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear the SA cache of all, local, or rejected entries, or entries for a specific group.</td>
<td>clear ip msdp sa-cache [group-address local rejected-sa]</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cache rejected sources.</td>
<td>ip msdp cache-rejected-sa</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**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 31-10, 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 31-10. MSDP Default Peer

Scenario 1

<table>
<thead>
<tr>
<th>Group</th>
<th>Source</th>
<th>RP</th>
<th>Peer</th>
</tr>
</thead>
<tbody>
<tr>
<td>G2</td>
<td>S2</td>
<td>RP2</td>
<td>R2</td>
</tr>
<tr>
<td>G3</td>
<td>S3</td>
<td>RP3</td>
<td>R3</td>
</tr>
<tr>
<td>G4</td>
<td>S4</td>
<td>RP4</td>
<td>R4</td>
</tr>
<tr>
<td>G5</td>
<td>S5</td>
<td>RP5</td>
<td>R5</td>
</tr>
</tbody>
</table>

Scenario 2

<table>
<thead>
<tr>
<th>Group</th>
<th>Source</th>
<th>RP</th>
<th>Peer</th>
</tr>
</thead>
<tbody>
<tr>
<td>G2</td>
<td>S2</td>
<td>RP2</td>
<td>R2</td>
</tr>
<tr>
<td>G3</td>
<td>S3</td>
<td>RP3</td>
<td>R3</td>
</tr>
<tr>
<td>G4</td>
<td>S4</td>
<td>RP4</td>
<td>R4</td>
</tr>
<tr>
<td>G5</td>
<td>S5</td>
<td>RP5</td>
<td>R5</td>
</tr>
</tbody>
</table>

Scenario 3

Group Source RP Peer
G2 S2 RP2 R2
G3 S3 RP3 R3
G4 S4 RP4 R4
G5 S5 RP5 R5

Scenario 4

Group Source RP Peer
G2 S2 RP2 R3 RPF-Fail
G3 S3 RP3 R3
G4 S4 RP4 R3 RPF-Fail
G5 S5 RP5 R3

ip msdp default-peer Router 3
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
Figure 31-11. Accepting Source-active Messages with Limit the Source-active Messages from a Peer

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify the forwarding-peer and originating-RP from which all active sources are accepted without regard for 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.</td>
<td>ip msdp default-peer ip-address list</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Task</td>
<td>Command Syntax</td>
<td>Command Mode</td>
</tr>
<tr>
<td>OPTIONAL: Store sources that are received after the limit is reached in the rejected SA cache.</td>
<td>ip msdp cache-rejected-sa</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Set the upper limit for the number of sources allowed from an MSDP peer. The default limit is 100K.</td>
<td>ip msdp peer peer-address sa-limit</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPTIONAL: Cache sources that are denied by the redistribute list in the rejected SA cache.</td>
<td>ip msdp cache-rejected-sa</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Prevent the system from caching local SA entries based on source and group using an extended ACL.</td>
<td>ip msdp redistribute list</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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 31-12. 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
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           Distribute
```
Prevent MSDP from Caching a Remote Source

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPTIONAL: Cache sources that are denied by the SA filter in the rejected SA cache.</td>
<td>ip msdp cache-rejected-sa</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Prevent the system from caching remote sources learned from a specific peer based on source and group.</td>
<td>ip msdp sa-filter list out peer list ext-acl</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

In Figure 31-14, 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 31-13. Preventing MSDP from Advertising a Local Source](image-url)
Prevent MSDP from Advertising a Local Source

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevent an RP from advertising a source in the SA cache.</td>
<td>ip msdp sa-filter list in peer ext-acl</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

In Figure 31-14, 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 31-14. 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
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     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 31-14).
Log Changes in Peership States

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log peership state changes</td>
<td>ip msdp log-adjacency-changes</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terminate the TCP connection with a peer.</td>
<td>ip msdp shutdown</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

Once the relationship is terminated, the peering state of the terminator is SHUTDOWN, while the peering state of the peer is INACTIVE.

Figure 31-15. 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

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reset the TCP connection to the peer and clear all peer statistics.</td>
<td><code>clear ip msdp peer peer-address</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

Figure 31-16. 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
```
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.
To configure Anycast RP:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>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.</td>
<td>interface loopback</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Make this address the RP for the group.</td>
<td>ip pim rp-address</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>3</td>
<td>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.</td>
<td>interface loopback</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>4</td>
<td>Peer each RP with every other RP using MSDP, specifying the unique loopback address as the connect-source.</td>
<td>ip msdp peer</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>5</td>
<td>Advertise the network of each of the unique loopback addresses throughout the network.</td>
<td>network</td>
<td>ROUTER OSPF</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create a mesh group.</td>
<td>ip msdp mesh-group</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use the address of another interface as the originator-id instead of the RP address.</td>
<td>ip msdp originator-id</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
Figure 31-19. 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
!
routing 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
```
Multicast Source Discovery Protocol (MSDP)

Figure 31-20. R2 Configuration for MSDP with Anycast RP

```conf
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 31-21. 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 31-5, Figure 31-4, Figure 31-5, Figure 31-6.

Figure 31-22. MSDP Sample Configuration: R1 Running-config

```plaintext
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 31-23. 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
```
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
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
Multiple Spanning Tree Protocol (MSTP)

Multiple Spanning Tree Protocol (MSTP) is supported on platforms: E C S Z S4810

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 the following illustration, 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 demonstrates how you can use MSTP to achieve load balancing.

Figure 32-1. MSTP with Three VLANs Mapped to Two Spanning Tree Instances
FTOS supports three other variations of Spanning Tree, as shown in Table 44.

Table 32-1. FTOS Supported Spanning Tree Protocols

<table>
<thead>
<tr>
<th>Dell Force10 Term</th>
<th>IEEE Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanning Tree Protocol</td>
<td>802.1d</td>
</tr>
<tr>
<td>Rapid Spanning Tree Protocol</td>
<td>802.1w</td>
</tr>
<tr>
<td>Multiple Spanning Tree Protocol</td>
<td>802.1s</td>
</tr>
<tr>
<td>Per-VLAN Spanning Tree Plus</td>
<td>Third Party</td>
</tr>
</tbody>
</table>

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 1004.
2. Place the interfaces in VLANs.
3. Enable Multiple Spanning Tree Protocol. See page 693.
4. Create Multiple Spanning Tree Instances, and map VLANs to them. See page 694.

Related Configuration Tasks

- Create Multiple Spanning Tree Instances on page 694
- Add and Remove Interfaces on page 693
- Influence MSTP Root Selection on page 695
- Interoperate with Non-FTOS Bridges on page 695
- Modify Global Parameters on page 696
- Modify Interface Parameters on page 697
- Configure an EdgePort on page 698
- Flush MAC Addresses after a Topology Change on page 699
- Debugging and Verifying MSTP Configuration on page 704
Enable Multiple Spanning Tree Globally

MSTP is not enabled by default. To enable MSTP:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter PROTOCOL MSTP mode.</td>
<td>protocol spanning-tree mstp</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Enable MSTP.</td>
<td>no disable</td>
<td>PROTOCOL MSTP</td>
</tr>
</tbody>
</table>

Verify that MSTP is enabled using the `show config` command from PROTOCOL MSTP mode, as shown in Figure 32-2.

**Figure 32-2. 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 1008 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 32-3.

**Figure 32-3. 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 32-6.

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 32-4.

**Figure 32-4. 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assign a number as the bridge priority. A lower number increases the probability that the bridge becomes the root bridge.</td>
<td><code>msti instance bridge-priority priority</code></td>
<td>PROTOCOL MSTP</td>
</tr>
<tr>
<td><strong>Range</strong>: 0 to 61440, in increments of 4096</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Default</strong>: 32768</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The simple configuration Figure 32-1 by default yields the same forwarding path for both MSTIs. Figure 32-5, 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 32-5.

**Figure 32-5. 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change the region name.</td>
<td>name name</td>
<td>PROTOCOL MSTP</td>
</tr>
<tr>
<td>Change the region revision number.</td>
<td>revision number</td>
<td>PROTOCOL MSTP</td>
</tr>
<tr>
<td>• Range: 0 to 65535</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Default: 0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

View the current region name and revision using the command `show spanning-tree mst configuration` from EXEC Privilege mode, as shown in **Figure 32-6**.

**Figure 32-6. 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change the forward-delay parameter.</td>
<td>forward-delay seconds</td>
<td>PROTOCOL MSTP</td>
</tr>
<tr>
<td>• Range: 4 to 30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Default: 15 seconds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change the hello-time parameter.</td>
<td>hello-time seconds</td>
<td>PROTOCOL MSTP</td>
</tr>
<tr>
<td>Note: With large configurations (especially those with more ports) Dell Force10 recommends that you increase the hello-time.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range: 1 to 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Default: 2 seconds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change the max-age parameter.</td>
<td>max-age seconds</td>
<td>PROTOCOL MSTP</td>
</tr>
<tr>
<td>Range: 6 to 40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Default: 20 seconds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change the max-hops parameter.</td>
<td>max-hops number</td>
<td>PROTOCOL MSTP</td>
</tr>
<tr>
<td>Range: 1 to 40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Default: 20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

View the current values for MSTP parameters using the `show running-config spanning-tree mstp` command from EXEC privilege mode.

**Figure 32-7. 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 32-2 lists the default values for port cost by interface.

<table>
<thead>
<tr>
<th>Port Cost</th>
<th>Default Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>100-Mb/s Ethernet interfaces</td>
<td>200000</td>
</tr>
<tr>
<td>1-Gigabit Ethernet interfaces</td>
<td>20000</td>
</tr>
<tr>
<td>10-Gigabit Ethernet interfaces</td>
<td>2000</td>
</tr>
<tr>
<td>Port Channel with 100 Mb/s Ethernet interfaces</td>
<td>180000</td>
</tr>
<tr>
<td>Port Channel with 1-Gigabit Ethernet interfaces</td>
<td>18000</td>
</tr>
<tr>
<td>Port Channel with 10-Gigabit Ethernet interfaces</td>
<td>1800</td>
</tr>
</tbody>
</table>

To change the port cost or priority of an interface:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change the port cost of an interface.</td>
<td>spanning-tree msti number cost cost</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>Range: 0 to 200000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Default: see Table 32-2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change the port priority of an interface.</td>
<td>spanning-tree msti number priority priority</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>Range: 0 to 240, in increments of 16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Default: 128</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

View the current values for these interface parameters using the command show config from INTERFACE mode. See Figure 32-8.

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 interface, use the following command:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable EdgePort on an interface.</td>
<td>spanning-tree mstp edge-port</td>
<td>INTERFACE</td>
</tr>
<tr>
<td></td>
<td>[bpduguard</td>
<td></td>
</tr>
<tr>
<td></td>
<td>shutdown-on-violation]</td>
<td></td>
</tr>
</tbody>
</table>

Verify that EdgePort is enabled on a port using the command `show config` from the INTERFACE mode, as shown in Figure 32-8.

**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 disabled 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 32-8. 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
FTOS(conf-if-gi-3/41)#
```

**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.
MSTP Sample Configurations

The running-configurations in Figure 32-10, Figure 32-11, and Figure 32-11 support the topology shown in Figure 32-9. The configurations are from FTOS systems. An S50 system using SFTOS, configured as shown Figure 32-13, could be substituted for an FTOS router in this sample following topology and MSTP would function as designed.

Figure 32-9. MSTP with Three VLANs Mapped to Two Spanning Tree Instances
Figure 32-10. Router 1 Running-configuration

```
protocol spanning-tree mstp
  no disable
  name Tahiti
  revision 123
  MSTI 1 VLAN 100
  MSTI 2 VLAN 200,300

! interface GigabitEthernet 1/21
  no ip address
  switchport
  no shutdown

! interface GigabitEthernet 1/31
  no ip address
  switchport
  no shutdown

! interface Vlan 100
  no ip address
  tagged GigabitEthernet 1/21,31
  no shutdown

! interface Vlan 200
  no ip address
  tagged GigabitEthernet 1/21,31
  no shutdown

! interface Vlan 300
  no ip address
  tagged GigabitEthernet 1/21,31
  no shutdown
```
Figure 32-11. Router 2 Running-configuration

```
protocol spanning-tree mstp
  no disable
  name Tahiti
  revision 123
  MSTI 1 VLAN 100
  MSTI 2 VLAN 200,300
!
interface GigabitEthernet 2/11
  no ip address
  switchport
  no shutdown
!
interface GigabitEthernet 2/31
  no ip address
  switchport
  no shutdown
!
interface Vlan 100
  no ip address
  tagged GigabitEthernet 2/11,31
  no shutdown
!
interface Vlan 200
  no ip address
  tagged GigabitEthernet 2/11,31
  no shutdown
!
interface Vlan 300
  no ip address
  tagged GigabitEthernet 2/11,31
  no shutdown
```

- Set Region Name and Revision
- Map MSTP Instances to VLANs
- Enable MSTP globally
- Assign Layer-2 interfaces to MSTP topology
- Create VLANs mapped to MSTP Instances
- Tag interfaces to VLANs
Multiple Spanning Tree Protocol (MSTP) | 703

Figure 32-12. Router 3 Running-configuration

```
protocol spanning-tree mstp
   no disable
   name Tahiti
   revision 123
   MSTI 1 VLAN 100
   MSTI 2 VLAN 200,300

! interface GigabitEthernet 3/11
   no ip address
   switchport
   no shutdown

! interface GigabitEthernet 3/21
   no ip address
   switchport
   no shutdown

! interface Vlan 100
   no ip address
   tagged GigabitEthernet 3/11,21
   no shutdown

! interface Vlan 200
   no ip address
   tagged GigabitEthernet 3/11,21
   no shutdown

! interface Vlan 300
   no ip address
   tagged GigabitEthernet 3/11,21
   no shutdown
```

- Enable MSTP globally
- Set Region Name and Revision
- Map MSTP Instances to VLANs
- Assign Layer-2 interfaces to MSTP topology
- Create VLANs mapped to MSTP Instances
- Tag interfaces to VLANs
Figure 32-13. SFTOS Example Running-Configuration

```plaintext
spanning-tree
spanning-tree configuration name Tahiti
spanning-tree configuration revision 123
spanning-tree MSTi instance 1
spanning-tree MSTi vlan 1 100
spanning-tree MSTi instance 2
spanning-tree MSTi vlan 2 200
spanning-tree MSTi vlan 2 300
interface 1/0/31
   no shutdown
   spanning-tree port mode enable
   switchport protected 0
   exit
interface 1/0/32
   no shutdown
   spanning-tree port mode enable
   switchport protected 0
   exit
interface vlan 100
tagged 1/0/31
tagged 1/0/32
exit
interface vlan 200
tagged 1/0/31
tagged 1/0/32
exit
interface vlan 300
tagged 1/0/31
tagged 1/0/32
exit
```

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`.

Set Region Name and Revision
Map MSTP Instances to VLANs
Assign Layer-2 interfaces to MSTP topology
Create VLANs mapped to MSTP Instances
Tag interfaces to VLANs
Enable MSTP globally

---

704 | Multiple Spanning Tree Protocol (MSTP)
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 32-15).

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 32-16).

Key items to look for in the debug report:

- MSTP flags indicate communication received from the same region.
  - In Figure 32-16, 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 32-17)? 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 32-15. Sample Output for show running-configuration spanning-tree mstp command

```plaintext
FTOS#show run spanning-tree mstp

protocol spanning-tree mstp
name Tahiti
revision 123
MSTI 1 VLAN 100
MSTI 2 VLAN 200,300
```

Figure 32-16. Displaying BPDUs and Events - Debug Log of Successful MSTP Configuration

```plaintext
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 Pri: 32768/128, Rem Hops: 20
INST 2: Flags: 0x6e, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
Brg/Port Pri: 32768/128, Rem Hops: 20

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.e8d5.cbbd
4w0d4h : INST 1: Flags: 0x6e, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
Brg/Port Pri: 32768/128, Rem Hops: 20
INST 2: Flags: 0x6e, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
Brg/Port Pri: 32768/128, Rem Hops: 20
```

Figure 32-17. Displaying BPDUs and Events - Debug Log of Unsuccessful MSTP Configuration

```plaintext
4w0d4h : MSTP: Received BPDU on Gi 2/21:
ProtId: 0, Ver: 3, Bpdu Type: MSTP, Flags 0x78
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 Pri: 32768/128, Rem Hops: 19
INST 2: Flags: 0x78, Reg Root: 32768:0001.e806.953e, Int Root Cost: 0
Brg/Port Pri: 32768/128, Rem Hops: 19
```

**Figure 32-17.** Displaying BPDUs and Events - Debug Log of Unsuccessful MSTP Configuration

- **Same Region:** Indicates MSTP routers are in the (single) region.
- **Different Region:** Indicates MSTP routers are in different regions and are not communicating with each other.
Multicast Features

Multicast Features are supported on platforms: E C S S4810

This chapter contains the following sections:

- Enable IP Multicast on page 707
- Multicast with ECMP on page 708
- First Packet Forwarding for Lossless Multicast on page 709
- Multicast Policies on page 710
- Multicast Traceroute on page 717

FTOS supports the following multicast protocols:

- PIM Sparse-Mode (PIM-SM) on page 775
- Internet Group Management Protocol (IGMP) on page 457
- Multicast Source Discovery Protocol (MSDP) on page 663

Implementation Information

- Multicast is supported on secondary IP addresses on the S4810 platform.

Enable IP Multicast

Enable IP Multicast is supported on platforms C E S

Prior to enabling any multicast protocols, you must enable multicast routing.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable multicast routing.</td>
<td>ip multicast-routing</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
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 33-1, 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.

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.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Ethernet Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSPF</td>
<td>01:00:5e:00:00:05</td>
</tr>
<tr>
<td></td>
<td>01:00:5e:00:00:06</td>
</tr>
<tr>
<td>RIP</td>
<td>01:00:5e:00:00:09</td>
</tr>
<tr>
<td>NTP</td>
<td>01:00:5e:00:01:01</td>
</tr>
<tr>
<td>VRRP</td>
<td>01:00:5e:00:00:12</td>
</tr>
<tr>
<td>PIM-SM</td>
<td>01:00:5e:00:00:0d</td>
</tr>
</tbody>
</table>

- 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.

**First Packet Forwarding for Lossless Multicast**

Beginning with FTOS version 7.8.1.0 for the E-Series TeraScale, version 8.2.1.0 for E-Series ExaScale, and version 8.3.1.0 on all other FTOS platforms, all initial multicast packets are forwarded to receivers to achieve lossless multicast.

In previous versions, when the Dell Force10 system is an RP, all initial packets are dropped until PIM creates an (S,G) entry. When the system is an RP and a Source DR, these initial packet drops represent a loss of native data, and when the system is an RP only, the initial packets drops represent a loss of register packets.

Both scenarios might be unacceptable depending on the multicast application. Beginning with the FTOS versions above, when the Dell Force10 system is the RP, and has receivers for a group G, it forwards all initial multicast packets for the group based on the (*,G) entry rather than discarding them until the (S,G) entry is created, making Dell Force10 systems suitable for applications sensitive to multicast packet loss.

**Note:** When a source begins sending traffic, the Source DR forwards the initial packets to the RP as encapsulated registered packets. These packets are forwarded via the soft path at a maximum rate of 70 packets/second. Incoming packets beyond this rate are dropped.
Multicast Policies

FTOS offers parallel Multicast features for IPv4 and IPv6.

- IPv4 Multicast Policies on page 710
- IPv6 Multicast Policies on page 715

IPv4 Multicast Policies

- Limit the Number of Multicast Routes on page 710
- Prevent a Host from Joining a Group on page 711
- Rate Limit IGMP Join Requests on page 713
- Prevent a PIM Router from Forming an Adjacency on page 713
- Prevent a Source from Registering with the RP on page 713
- Prevent a PIM Router from Processing a Join on page 715

Limit the Number of Multicast Routes

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limit the total number of multicast routes on the system.</td>
<td>ip multicast-limit</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Range: 1-50000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Default: 15000</td>
<td></td>
</tr>
</tbody>
</table>

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.
```
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 33-2), 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 33-2, 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 33-2. Preventing a Host from Joining a Group

```
interface Vlan 300
ip pim sparse-mode
ip address 10.11.3.1/24 untagged GigabitEthernet 1/1
no shutdown

interface GigabitEthernet 1/2
ip pim sparse-mode
ip address 10.11.12.1/24 no shutdown

interface GigabitEthernet 1/3
ip pim sparse-mode
ip address 10.11.13.1/24 no shutdown

interface GigabitEthernet 2/1
ip pim sparse-mode
ip address 10.11.1.1/24 no shutdown

interface GigabitEthernet 2/11
ip pim sparse-mode
ip address 10.11.12.2/24 no shutdown

interface GigabitEthernet 3/1
ip pim sparse-mode
ip address 10.11.13.2/24 no shutdown

interface GigabitEthernet 3/2
ip pim sparse-mode
ip address 10.11.23.2/24 no shutdown

interface GigabitEthernet 2/3
ip pim sparse-mode
ip address 10.11.23.1/24 no shutdown

ip multicast-routing
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
router rip
network 10.0.0.0

ip igmp snooping enable

R1(conf-if-vl-300)# 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 ... for MSDP Advertisement
K - Ack-Pending State
Timers: Uptime/Expires
Interface state: Interface, next-Hop, State/Mode

(*, 239.0.0.1), uptime 00:00:07, expires 00:00:00, RP 10.11.12.2, flags: S
Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
Outgoing interface list:
Vlan 300  Forward/Sparse   00:00:07/Never

(*, 239.0.0.2), uptime 00:01:10, expires 00:00:00, RP 10.11.12.2, flags: S
Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
Outgoing interface list:
Vlan 300  Forward/Sparse   00:01:10/Never

R1(conf)#do show run acl
!ip access-list extended igmpjoinfilR2G2
seq 5 permit ip any host 239.0.0.1
239.0.0.2
```

Source 1
10.11.1.2

Source 2
10.11.5.2

Receiver 1
10.11.3.2

Receiver 2
10.11.4.2

Group: 239.0.0.1, 239.0.0.2

interface GigabitEthernet 2/11
ip pim sparse-mode
ip address 10.11.3.1/24
no shutdown

interface GigabitEthernet 3/11
ip pim sparse-mode
ip address 10.11.5.1/24
no shutdown
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 33-3, 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 33-3. Preventing a Source from Transmitting to a Group

```
interface Vlan 300
ip pim sparse-mode
ip address 10.11.3.1/24 untagged GigabitEthernet 1/1
no shutdown

interface GigabitEthernet 1/2
ip pim sparse-mode
ip address 10.11.12.1/24
no shutdown

interface GigabitEthernet 1/3
ip pim sparse-mode
ip address 10.11.13.1/24
no shutdown

interface GigabitEthernet 2/1
ip pim sparse-mode
ip address 10.11.12.2/24
no shutdown

interface GigabitEthernet 3/1
ip pim sparse-mode
ip address 10.11.13.2/24
no shutdown

interface GigabitEthernet 3/2
ip pim sparse-mode
ip address 10.11.23.2/24
no shutdown

interface GigabitEthernet 2/3
ip pim sparse-mode
ip address 10.11.23.1/24
no shutdown

ip multicast-routing
ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
router rip
network 10.0.0.0

ip igmp snooping enable
```

```
R1(config-line-console)#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

(*, 239.0.0.1), uptime 00:00:43, expires 00:00:00, RP 10.11.12.2, flags: SCJ
Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
Outgoing interface list:
  Vlan 400  Forward/Sparse  00:00:43/Never

(10.11.1.2, 239.0.0.1), uptime 00:00:17, expires 00:03:17, flags: CT
Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
Outgoing interface list:
  Vlan 400  Forward/Sparse  00:00:43/Never

(10.11.5.2, 239.0.0.1), uptime 00:00:23, expires 00:03:17, flags: CT
Incoming interface: GigabitEthernet 1/31, RPF neighbor 10.11.13.2
Outgoing interface list:
  Vlan 400  Forward/Sparse  00:00:43/Never

(*, 239.0.0.2), uptime 00:00:40, expires 00:00:00, RP 10.11.12.2, flags: SCJ
Incoming interface: GigabitEthernet 1/21, RPF neighbor 10.11.12.2
Outgoing interface list:
  Vlan 300  Forward/Sparse  00:00:40/Never

(10.11.5.2, 239.0.0.2), uptime 00:00:33, expires 00:03:07, flags: CT
Incoming interface: GigabitEthernet 1/31, RPF neighbor 10.11.13.2
Outgoing interface list:
  Vlan 400  Forward/Sparse  00:00:40/Never
```

```
R2(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

(10.11.5.2, 239.0.0.1), uptime 00:05:21, expires 00:02:46, flags: FT
Incoming interface: GigabitEthernet 3/1, RPF neighbor 0.0.0.0
Outgoing interface list:
  GigabitEthernet 3/11  Forward/Sparse  00:00:18/00:03:12

(10.11.5.2, 239.0.0.2), uptime 00:05:08, expires 00:03:26, flags: FT
Incoming interface: GigabitEthernet 3/1, RPF neighbor 0.0.0.0
Outgoing interface list:

(*, 239.0.0.1), uptime 00:03:00, expires 00:02:32, RP 10.11.12.2, flags: SF
Incoming interface: Null, RPF neighbor 0.0.0.0
Outgoing interface list:
  GigabitEthernet 2/11  Forward/Sparse  00:03:00/00:02:32

(10.11.1.2, 239.0.0.1), uptime 00:00:44, expires 00:02:56, flags: FT
Incoming interface: GigabitEthernet 2/1, RPF neighbor 0.0.0.0
Outgoing interface list:
  GigabitEthernet 2/11  Forward/Sparse  00:00:13/00:03:17

(10.11.5.2, 239.0.0.1), uptime 00:00:26, expires 00:03:06, flags: P
Incoming interface: GigabitEthernet 2/31, RPF neighbor 10.11.23.2
Outgoing interface list:
```

```
R3(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

(10.11.5.2, 239.0.0.1), uptime 00:03:00, expires 00:02:32, RP 10.11.12.2, flags: SF
Incoming interface: Null, RPF neighbor 0.0.0.0
Outgoing interface list:
  GigabitEthernet 2/11  Forward/Sparse  00:03:00/00:02:32

(10.11.1.2, 239.0.0.1), uptime 00:00:44, expires 00:02:56, flags: FT
Incoming interface: GigabitEthernet 2/1, RPF neighbor 0.0.0.0
Outgoing interface list:
  GigabitEthernet 2/11  Forward/Sparse  00:00:13/00:03:17

(*, 239.0.0.2), uptime 00:00:21, expires 00:03:28, flags: FT
Incoming interface: GigabitEthernet 2/1, RPF neighbor 0.0.0.0
Outgoing interface list:
  GigabitEthernet 2/11  Forward/Sparse  00:00:21/00:03:28
```

```
R3(conf)#do show run acl
!ip access-list extended regfilS1G2
seq 5  permit ip host 10.11.5.2 host 239.0.0.1
```

```
R3(conf)#do show run pim
!ip pim rp-address 10.11.12.2 group-address 224.0.0.0/4
ip pim register-filter regfilS1G2
```
```
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.

**Note:** Dell Force10 recommends that you do not use the `ip pim join-filter` command on an interface between a source and the RP router. Use of this command in this scenario could cause problems with the PIM-SM source registration process resulting in excessive traffic being sent to the CPU of both the RP and PIM DR of the source.

Excessive traffic is generated when the join process from the RP back to the source is blocked due to a new source group being permitted in the join-filter. This results in the new source becoming stuck in registering on the DR and the continuous generation of UDP-encapsulated registration messages between the DR and RP routers which are being sent to the CPU.

**IPv6 Multicast Policies**

IPv6 Multicast Policies is available only on platform: E

- Limit the Number of IPv6 Multicast Routes on page 715
- Prevent an IPv6 Neighbor from Forming an Adjacency on page 716
- Prevent an IPv6 Source from Registering with the RP on page 716
- Prevent an IPv6 PIM Router from Processing an IPv6 Join on page 716

**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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limit the total number of IPv6 multicast routes on the system.</td>
<td><code>ipv6 multicast-limit</code> Range: 1-50000 Default: 15000</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
### Prevent an IPv6 Neighbor from Forming an Adjacency

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevent a router from participating in PIM.</td>
<td><code>ipv6 pim neighbor-filter access-list</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

```bash
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

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configured on the source DR, prevent the source DR from sending register packets to the RP for specific sources and groups.</td>
<td><code>ipv6 pim register-filter access-list</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

```bash
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

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
<td>`ipv6 pim join-filter access-list [in</td>
<td>out]`</td>
</tr>
</tbody>
</table>

```bash
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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Print the network path that a multicast packet takes from a multicast source to receiver, for a particular group.</td>
<td><code>mtrace multicast-source-address multicast-receiver-address multicast-group-address</code></td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

Figure 33-4. 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
```
Open Shortest Path First (OSPFv2 and OSPFv3)

Open Shortest Path First version 2 (OSPF for IPv4) is supported on platforms CEZ.
Open Shortest Path First version 3 (OSPF for IPv6) is supported on platforms CEZ.

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 Force10 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:

- Protocol Overview
- Implementing OSPF with FTOS
  - Graceful Restart
  - Fast Convergence (OSPFv2, IPv4 only)
  - Multi-Process OSPF (OSPFv2, IPv4 only)
  - RFC-2328 Compliant OSPF Flooding
  - OSPF ACK Packing
  - OSPF Adjacency with Cisco Routers
- Configuration Information
  - Configuration Task List for OSPFv2 (OSPF for IPv4)
  - Configuration Task List for OSPFv3 (OSPF for IPv6)
- Sample Configurations for OSPFv2

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.
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 a default area number. All other areas can have their Area ID assigned in the configuration.

**Figure 34-1** 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 not 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, Dell Force10 recommends that the router ID and the router’s IP address reflect each other to make troubleshooting easier.

Figure 34-2 gives some examples of the different router designations.
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 34-2 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 to which it connects.

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 34-2.

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 as 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 Dell Force10 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 to which this link connects.

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 where 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; 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.
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). The S-Series supports up to 16 processes simultaneously. The Z-Series supports up to three OSPF processes simultaneously.

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 S Z S48T0" platforms in Helper and Restart modes.

Graceful Restart for OSPFv3 is supported on "E T Z S48T0" 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 helper-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 LSAs 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 S Z S4810 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 S50 and S25 support up to 4 OSPFv2 processes.
- The S55, S60, and S4810 support up to 16 OSPFv2 processes.
• The Z9000 supports up to 16 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 adjacency, 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 OSPFVv2 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.) hen 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:
Use `show ip ospf` to confirm that RFC 2328-compliant OSPF flooding is enabled, as shown below.

### 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 is non-configurable.

### OSPF Adjacency with Cisco Routers

To establish an OSPF adjacency between Dell Force10 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 34-6. Command Example: ip ospf intervals**

```bash
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
```

**Figure 34-7. OSPF Configuration with intervals set**

```bash
FTOS (conf-if-gi-2/2)#ip ospf dead-interval 20
FTOS (conf-if-gi-2/2)#do show ip interface
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
  Dead Interval Set at 4x
  Hello Interval
    Set at 4x
  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)
```

For more information regarding this functionality or for assistance, go to [http://support.dell.com/force10](http://support.dell.com/force10).

## Configuration Information

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 Z 54810
```

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 OSPF commands, refer to the OSPF section in the *FTOS Command Line Interface Guide*.

**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 four OSPFv2 process IDs, you must have four interfaces with Layer 3 enabled.

Use these commands on one of the interfaces to enable OSPFv2 routing.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ip address ip-address mask</td>
<td>CONFIG-INTERFACE</td>
<td>Assign an IP address to an interface. Format: A.B.C.D/M</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>If using a Loopback interface, refer to Loopback Interfaces.</td>
</tr>
<tr>
<td>2</td>
<td>no shutdown</td>
<td>CONFIG-INTERFACE</td>
<td>Enable the interface.</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>router ospf process-id</td>
<td>CONFIGURATION</td>
<td>Enable the OSPFv2 process globally. Range: 0-65535</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vrf name: Enter the VRF keyword and instance name to tie the OSPF process to the VRF. All network commands under this OSPF instance are subsequently tied to the VRF instance.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>router-id ip address</td>
<td>CONFIG-ROUTER-OSPF-id</td>
<td>Assign the Router ID for the OSPFv2 process. IP Address: A.B.C.D</td>
</tr>
</tbody>
</table>

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 to view the current OSPFv2 status.

**Figure 34-8. 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 SPF's 10 secs
Number of area in this router is 0, normal 0 stub 0 nssa 0
```
Enable Multi-Process OSPF

Multi-Process OSPF allows multiple OSPFv2 processes on a single router. For more information, see Multi-Process OSPF (OSPFv2, IPv4 only).

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ip address ip-address mask</td>
<td>CONFIG-INTERFACE</td>
<td>Assign an IP address to an interface. Format: A.B.C.D/M</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>If using a Loopback interface, refer to Loopback Interfaces.</td>
</tr>
<tr>
<td>2</td>
<td>no shutdown</td>
<td>CONFIG-INTERFACE</td>
<td>Enable the interface.</td>
</tr>
</tbody>
</table>

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.

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>router-id ip address</td>
<td>CONFIG-ROUTER-OSPF-id</td>
<td>Assign the Router ID for the OSPFv2 process. IP Address: A.B.C.D</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>network ip-address mask area area-id</td>
<td>CONFIG-ROUTER-OSPF-id</td>
<td>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</td>
</tr>
</tbody>
</table>
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 34-9 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 34-9. Configuring an OSPF Area Example**

Dell Force10 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 to view the interfaces currently active and the areas assigned to the interfaces.
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 34-11 gives an example of the **show ip ospf process-id interface** command with a Loopback interface.
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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>show ip ospf process-id [vrf vrf name] database database-summary</code></td>
<td>EXEC Privilege</td>
<td>Review all areas after they were configured to determine which areas are NOT receiving type 5 LSAs. <code>vrf name</code>: Show only the OSPF information tied to the VRF process.</td>
</tr>
<tr>
<td>2</td>
<td>configure</td>
<td>EXEC Privilege</td>
<td>Enter the CONFIGURATION mode.</td>
</tr>
<tr>
<td>3</td>
<td><code>router ospf process-id [vrf {vrf name}]</code></td>
<td>CONFIGURATION</td>
<td>Enter the ROUTER OSPF mode. Process ID is the ID assigned when configuring OSPFv2 globally. <code>vrf name</code>: Enter the VRF keyword 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.</td>
</tr>
<tr>
<td>4</td>
<td><code>area area-id stub [no-summary]</code></td>
<td>CONFIG-ROUTER-OSPF-id</td>
<td>Configure the area as a stub area. Use the <code>no-summary</code> 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.</td>
</tr>
</tbody>
</table>

Use the `show ip ospf database process-id database-summary` command syntax in the EXEC Privilege mode to view which LSAs are transmitted.

**Figure 34-12. 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)

<table>
<thead>
<tr>
<th>Area ID</th>
<th>Router</th>
<th>Network</th>
<th>S-Net</th>
<th>S-ASBR</th>
<th>Type-7</th>
<th>Subtotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.2.2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3.3.3.3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>max-metric router-lsa</td>
<td>ROUTER OSPF</td>
<td>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.</td>
</tr>
<tr>
<td>on-startup {announce-time</td>
<td>wait-for-bgp [wait-time]}</td>
<td></td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>passive-interface</td>
<td>CONFIG-ROUTER-</td>
<td>Specify whether all or some of the interfaces will be passive. Default enabled passive interfaces on ALL interfaces in the OSPF process.</td>
</tr>
<tr>
<td></td>
<td>OSPF-id</td>
<td>Entering the physical interface type, slot, and number enable passive interface on only the identified interface.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- For a Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/port information (e.g., passive-interface gi 2/1).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- For a port channel, enter the keyword port-channel followed by a number from 1 to 255 for TeraScale and ExaScale.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- For a SONET interface, enter the keyword sonet followed by the slot/port information (e.g., passive-interface so 2/2).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- For a 10-Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information (e.g., passive-interface ten 2/3).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- For a VLAN, enter the keyword vlan followed by a number from 1 to 4094 (e.g., passive-interface vlan 2222).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>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.</td>
</tr>
</tbody>
</table>

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 adds the words “passive interface” to indicate that hello packets are not transmitted on that interface.
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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>fast-convergence {number}</td>
<td>CONFIG-ROUTER-OSPF-id</td>
<td>Enable OSPF fast-convergence and specify the convergence level.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Parameter: 1-4</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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 34-14 shows the convergence settings when fast-convergence is enabled and Figure 34-15 shows settings when fast-convergence is disabled. These displays appear with the `show ip ospf` command.

**Figure 34-14. 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
```

**Figure 34-15. 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
```

**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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip ospf cost</td>
<td>CONFIG-INTERFACE</td>
<td>Change the cost associated with OSPF traffic on the interface.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cost: 1 to 65535 (default depends on the interface speed).</td>
</tr>
<tr>
<td>ip ospf dead-interval seconds</td>
<td>CONFIG-INTERFACE</td>
<td>Change the time interval the router waits before declaring a neighbor dead. Configure Seconds range: 1 to 65535. Default: 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.</td>
</tr>
<tr>
<td>ip ospf hello-interval seconds</td>
<td>CONFIG-INTERFACE</td>
<td>Change the time interval between hello-packet transmission.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Seconds range: from 1 to 65535 (default is 10 seconds).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The hello interval must be the same on all routers in the OSPF network.</td>
</tr>
<tr>
<td>ip ospf message-digest-key keyid md5 key</td>
<td>CONFIG-INTERFACE</td>
<td>Use the MD5 algorithm to produce a message digest/key, which is sent instead of the key. Keyid range: 1 to 255 Key: a character string</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Note: 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.</td>
</tr>
<tr>
<td>ip ospf priority number</td>
<td>CONFIG-INTERFACE</td>
<td>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).</td>
</tr>
<tr>
<td>ip ospf retransmit-interval seconds</td>
<td>CONFIG-INTERFACE</td>
<td>Change the retransmission interval between LSAs. Seconds range: 1 to 65535 (default is 5 seconds). The retransmit interval must be the same on all routers in the OSPF network.</td>
</tr>
<tr>
<td>ip ospf transmit-delay seconds</td>
<td>CONFIG-INTERFACE</td>
<td>Change the wait period between link state update packets sent out the interface. Seconds range: 1 to 65535 (default is 1 second). The transmit delay must be the same on all routers in the OSPF network.</td>
</tr>
</tbody>
</table>

Use the `show config` command in CONFIGURATION INTERFACE mode (Figure 34-16) to view interface configurations. Use the `show ip ospf interface` command in EXEC mode to view interface status in the OSPF process.
Enable OSPFv2 authentication

Use the following commands in CONFIGURATION INTERFACE mode to enable or change various OSPF authentication parameters:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ip ospf authentication-key key</code></td>
<td>CONFIG-INTERFACE</td>
<td>Set clear text authentication scheme on the interface. Configure a <code>key</code> that is a text string no longer than eight characters. All neighboring routers must share the same password to exchange OSPF information.</td>
</tr>
<tr>
<td><code>ip ospf auth-change-wait-time seconds</code></td>
<td>CONFIG-INTERFACE</td>
<td>Set the authentication change wait time in <code>seconds</code> 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.</td>
</tr>
</tbody>
</table>

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 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.
Note: By default, OSPFv2 graceful restart is disabled.

You enable OSPFv2 graceful restart in CONFIGURATION ROUTER OSPF mode.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>graceful-restart</td>
<td>CONFIG-ROUTER-OSPF-id</td>
<td>Enable OSPFv2 graceful-restart globally and set the grace period.</td>
</tr>
<tr>
<td>grace-period seconds</td>
<td></td>
<td>Seconds range: between 40 and 3000.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>This is the period of time that an OSPFv2 router’s neighbors will advertise it as fully</td>
</tr>
<tr>
<td></td>
<td></td>
<td>adjacent, regardless of the synchronization state, during a graceful restart.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OSPFv2 terminates this process when the grace period ends.</td>
</tr>
<tr>
<td>graceful-restart</td>
<td>CONFIG-ROUTER-OSPF-id</td>
<td>Enter the Router ID of the OSPFv2 helper router from which the router does not accept</td>
</tr>
<tr>
<td>helper-reject router-id</td>
<td></td>
<td>graceful-restart assistance. This applies to the specified router only.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IP Address: A.B.C.D</td>
</tr>
<tr>
<td>graceful-restart</td>
<td>CONFIG-ROUTER-OSPF-id</td>
<td>Specify the operating mode in which graceful-restart functions.</td>
</tr>
<tr>
<td>mode</td>
<td></td>
<td>FTOS supports the following options:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Planned-only. The OSPFv2 router supports graceful-restart for planned restarts only.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A planned restart is when the user manually enters a fail-over command to force the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>primary RPM over to the secondary RPM. During a planned restart, OSPF sends out a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grace LSA before the system switches over to the secondary RPM. OSPF also is notified</td>
</tr>
<tr>
<td></td>
<td></td>
<td>that a planned restart is happening.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Unplanned-only. The OSPFv2 router supports graceful-restart for only unplanned</td>
</tr>
<tr>
<td></td>
<td></td>
<td>restarts. During an unplanned restart, OSPF sends out a Grace LSA once the secondary</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RPM comes online.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>By default, OSPFv2 supports both planned and unplanned restarts. Selecting one or the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>other mode restricts OSPFv2 to the single selected mode.</td>
</tr>
<tr>
<td>graceful-restart</td>
<td>CONFIG-ROUTER-OSPF-id</td>
<td>Configure the graceful restart role or roles that this OSPFv2 router performs.</td>
</tr>
<tr>
<td>role</td>
<td></td>
<td>FTOS supports the following options:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Helper-only. The OSPFv2 router supports graceful-restart only as a helper router.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Restart-only. The OSPFv2 router supports graceful-restart only during unplanned restarts.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>By default, OSPFv2 supports both restarting and helper roles. Selecting one or the other</td>
</tr>
<tr>
<td></td>
<td></td>
<td>role restricts OSPFv2 to the single selected role.</td>
</tr>
</tbody>
</table>

When you configure a graceful restart on an OSPFv2 router, the `show run ospf` command (Figure 34-17) displays information similar to the following.
Use the following command to disable OSPFv2 graceful-restart after you have enabled it.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>no graceful-restart grace-period</td>
<td>CONFIG-ROUTER-OSPF-id</td>
<td>Disable OSPFv2 graceful-restart. Returns OSPF graceful-restart to its default state.</td>
</tr>
</tbody>
</table>

For more information on OSPF graceful restart, refer to the *FTOS Command Line Interface Reference Guide*.

**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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>area area-id virtual-link router-id [hello-interval seconds</td>
<td>retransmit-interval seconds</td>
<td>transmit-delay seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 34-18) in the EXEC mode to view the virtual link.

**Figure 34-18. 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
```

**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. 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip prefix-list prefix-name</td>
<td>CONFIGURATION</td>
<td>Create a prefix list and assign it a unique name. You are in PREFIX LIST mode.</td>
</tr>
</tbody>
</table>
For configuration information on prefix lists, refer to the Access Control Lists (ACLs) 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

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>seq</code> sequence-number {deny</td>
<td>permit} ip-prefix [ge min-prefix-length] [le max-prefix-length]</td>
<td>CONFIG- PREFIX LIST</td>
</tr>
</tbody>
</table>

For configuration information on prefix lists, refer to the Access Control Lists (ACLs) 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

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>distribute-list prefix-list-name in [interface]</code></td>
<td>CONFIG-ROUTER-OSPF-id</td>
<td>Apply a configured prefix list to incoming OSPF routes.</td>
</tr>
<tr>
<td>`distribute-list prefix-list-name out [connected</td>
<td>isis</td>
<td>rip</td>
</tr>
</tbody>
</table>

### 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

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>`redistribute {bgp</td>
<td>connected</td>
<td>isis</td>
</tr>
</tbody>
</table>
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 34-19. 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>show running-config ospf</code></td>
<td>EXEC Privilege</td>
<td>View the summary of all OSPF process IDs enables on the router.</td>
</tr>
</tbody>
</table>

**Figure 34-20. 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
```

Use the following commands in EXEC Privilege mode to get general route and links status information.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>show ip route summary</code></td>
<td>EXEC Privilege</td>
<td>View the summary information of the IP routes.</td>
</tr>
<tr>
<td><code>show ip ospf database</code></td>
<td>EXEC Privilege</td>
<td>View the summary information for the OSPF database.</td>
</tr>
</tbody>
</table>

Use the following command in EXEC Privilege mode to view the OSPFv2 configuration for a neighboring router:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>show ip ospf neighbor</code></td>
<td>EXEC Privilege</td>
<td>View the configuration of OSPF neighbors.</td>
</tr>
</tbody>
</table>
Use the following command in EXEC Privilege mode to configure the debugging options of an OSPFv2 process:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
</table>
| 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>show ip ospf database database-summary</td>
<td>EXEC Privilege</td>
<td>View a summary of OSPFv2 database information.</td>
</tr>
</tbody>
</table>

Figure 34-21. 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)

<table>
<thead>
<tr>
<th>Area ID</th>
<th>Router Net</th>
<th>S-Net</th>
<th>S-ASBR Type7</th>
<th>Type9</th>
<th>Type10</th>
<th>Total</th>
<th>ChSum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>4</td>
<td>3000</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3008</td>
</tr>
</tbody>
</table>
```
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.

![Basic OSPFv2 Router Topology](image)

**Figure 34-22. Basic topology and CLI commands for OSPFv2**

```
routing 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.

The OSPFv3 `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 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

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipv6 unicast routing</td>
<td>CONFIGURATION</td>
<td>Enables IPv6 unicast routing globally.</td>
</tr>
</tbody>
</table>

Assign IPv6 addresses on an interface

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipv6 address ipv6 address</td>
<td>CONF-INT-type slot/port</td>
<td>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</td>
</tr>
<tr>
<td>no shutdown</td>
<td>CONF-INT-type slot/port</td>
<td>Bring the interface up.</td>
</tr>
</tbody>
</table>

Assign Area ID on interface

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipv6 ospf process-id area area-id</td>
<td>CONF-INT-type slot/port</td>
<td>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.</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipv6 router ospf (process ID))</td>
<td>CONFIGURATION</td>
<td>Enable the OSPFv3 process globally and enter OSPFv3 mode. Range: 0-65535</td>
</tr>
<tr>
<td>router-id (number)</td>
<td>CONF-IPV6-ROUTER-OSPF</td>
<td>Assign the Router ID for this OSPFv3 process (number): IPv4 address Format: A.B.C.D</td>
</tr>
</tbody>
</table>

**Note:** The router-id for an OSPFv3 router is entered as an IPv4 IP address.

## Configure stub areas

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>area area-id stub [no-summary]</td>
<td>CONF-IPV6-ROUTER-OSPF</td>
<td>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.</td>
</tr>
</tbody>
</table>
## 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>passive-interface</td>
<td>CONF-IPV6-ROUTER-OSPF</td>
<td>Specify whether some or all some of the interfaces will be passive. Interface identifies the specific interface that will be passive.</td>
</tr>
<tr>
<td>{type slot/port}</td>
<td></td>
<td>• For a Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/port information (e.g. passive-interface gi 2/1).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 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).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10-Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information (e.g. passive-interface ten 2/3).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 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.</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>redistribute {bgp</td>
<td>connected</td>
<td>CONF-IPV6-ROUTER-OSPF Specify which routes will be redistributed into OSPF process. Configure the following required and optional parameters:</td>
</tr>
<tr>
<td>static} [metric metric-value</td>
<td></td>
<td>• <code>bgp, connected, or static</code>: enter one of the keyword to redistribute those routes.</td>
</tr>
<tr>
<td>metric-type type-value}] [route-map</td>
<td></td>
<td>• <code>metric metric-value</code> range: 0 to 4294967295.</td>
</tr>
<tr>
<td>map-name] [tag tag-value]</td>
<td></td>
<td>• <code>metric-type metric-type</code>: 1 for OSPFv3 external route type 1 OR 2 for OSPFv3 external route type 2.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>route-map map-name</code>: enter a name of a configured route map.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>tag tag-value</code> range: 0 to 4294967295.</td>
</tr>
</tbody>
</table>

Configure a default route

Configure FTOS to generate a default external route into the OSPFv3 routing domain.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>default-information originate [always [metric metric-value] [metric-type type-value]] [route-map map-name]</td>
<td>CONF-IPV6-ROUTER-OSPF Specify the information for the default route. Configure the following required and optional parameters:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>always</code>: indicate that default route information must always be advertised</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>metric metric-value</code> range: 0 to 4294967295.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>metric-type metric-type</code>: 1 for OSPFv3 external route type 1 OR 2 for OSPFv3 external route type 2.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>route-map map-name</code>: enter a name of a configured route map.</td>
</tr>
</tbody>
</table>
Enable OSPFv3 graceful restart

Graceful Restart for OSPFv3 is supported on platforms E Z S4810. Refer to Graceful Restart 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.)

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>graceful-restart</td>
<td>CONF-IPV6-ROUTE</td>
<td>Enable OSPFv3 graceful restart globally by setting the grace period</td>
</tr>
<tr>
<td>grace-period seconds</td>
<td>R-OSPF</td>
<td>(in seconds). Valid values are from 40 to 1800 seconds.</td>
</tr>
<tr>
<td>ipv6 ospf graceful-restart</td>
<td>INTERFACE</td>
<td>Configure an OSPFv3 interface to not act upon the Grace LSAs that</td>
</tr>
<tr>
<td>helper-reject</td>
<td></td>
<td>it receives from a restarting OSPFv3 neighbor.</td>
</tr>
<tr>
<td>graceful-restart mode</td>
<td>CONF-IPV6-ROUTE</td>
<td>Specify the operating mode and type of events that trigger a graceful</td>
</tr>
<tr>
<td>[planned-only</td>
<td></td>
<td>Restart:</td>
</tr>
<tr>
<td>unplanned-only]</td>
<td>R-OSPF</td>
<td>• <strong>Planned-only.</strong> The OSPFv3 router supports graceful restart only</td>
</tr>
<tr>
<td></td>
<td></td>
<td>for planned restarts. A planned restart is when you manually enter</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a redundancy force-failover rpm command to force the primary RPM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>over to the secondary RPM. During a planned restart, OSPFv3 sends</td>
</tr>
<tr>
<td></td>
<td></td>
<td>out a Grace LSA before the system switches over to the secondary RPM.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OSPFv3 is notified that a planned restart is happening.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>Unplanned-only.</strong> The OSPFv3 router supports graceful-restart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>only for unplanned restarts. During an unplanned restart, OSPFv3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sends out a Grace LSA once the secondary RPM comes online.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Default:</strong> Both planned and unplanned restarts trigger an OSPFv3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>graceful restart. Selecting one or the other mode restricts OSPFv3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to the single selected mode.</td>
</tr>
</tbody>
</table>

To disable OSPFv3 graceful restart when it is enabled, enter the following command:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>no graceful-restart grace-period</td>
<td>CONF-IPV6-ROUTE</td>
<td>Disable OSPFv3 graceful-restart.</td>
</tr>
<tr>
<td></td>
<td>R-OSPF</td>
<td></td>
</tr>
</tbody>
</table>
To display information on the use and configuration of OSPFv3 graceful restart, enter any of the following commands:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>show run ospf</td>
<td>EXEC Privilege</td>
<td>Display the graceful-restart configuration for OSPFv2 and OSPFv3 (Figure 34-23).</td>
</tr>
<tr>
<td>show ipv6 ospf database grace-lsa</td>
<td>EXEC Privilege</td>
<td>Display the Type-11 Grace LSAs sent and received on an OSPFv3 router (Figure 34-24).</td>
</tr>
<tr>
<td>show ipv6 ospf database database-summary</td>
<td>EXEC Privilege</td>
<td>Display the currently configured OSPFv3 parameters for graceful restart (Figure 34-25).</td>
</tr>
</tbody>
</table>

**Figure 34-23. 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
```
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

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
OSPFv3 Authentication Using IPsec

OSPFv3 Authentication Using IPsec is supported only on platforms: E Z S A T O.

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)).

To configure IPsec authentication on an interface, enter the following command:

```plaintext
ipv6 ospf authentication {null | ipsec spi number {MD5 | SHA1} [key-encryption-type] key}
```

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipv6 ospf authentication {null</td>
<td>INTERFACE</td>
<td>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</td>
</tr>
<tr>
<td>ipsec spi number {MD5</td>
<td>SHA1} [key-encryption-type] key}</td>
<td></td>
</tr>
</tbody>
</table>

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)).

To configure IPsec encryption on an interface, enter the following command:

```
ipv6 ospf encryption {null | ipsec}
  spi number esp encryption-algorithm
      [key-encryption-type] key
  authentication-algorithm
      [key-authentication-type] key
```

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipv6 ospf encryption {null</td>
<td>ipsec</td>
<td>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: 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: 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.</td>
</tr>
</tbody>
</table>

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)).

To configure IPsec authentication for an OSPFv3 area, enter the following command in global configuration mode:

```
area-id authentication ipsec spi number [MD5 | SHA1] [key-encryption-type] key
```

**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 OSPFv3 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)).

To configure IPsec encryption in an OSPFv3 area, enter the following command in global configuration mode:

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 OSPFv3 link.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
</table>
| `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: 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: 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. |
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:

```
show crypto ipsec policy [name name]
```

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>show crypto ipsec policy</code></td>
<td>EXEC Privilege</td>
<td>Display the AH and ESP parameters configured in IPsec security policies, including the SPI number, key, and algorithms used.</td>
</tr>
<tr>
<td><code>[name name]</code></td>
<td></td>
<td><code>name</code> displays configuration details about a specified policy.</td>
</tr>
</tbody>
</table>
Figure 34-26. 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 : 123456789a123456789b123456789c123456789d123456789
Outbound ESP Cipher Key : 123456789a123456789b123456789c123456789d123456789
Transform set : esp-3des esp-md5-hmac

In this encryption policy, the keys are not encrypted.

In this authentication policy, the keys are encrypted.

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

In this encryption policy, the keys are encrypted.

In this authentication policy, the keys are encrypted.

Crypto IPSec client security policy data

Policy name : OSPFv3-0-501
Policy refcount : 1
Inbound ESP Auth Key : bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba8ed8bb5efe91e97eb7c030808825fb5
Outbound ESP Auth Key : bbdd96e6eb4828e2e27bc3f9ff541e43faa759c9ef5706ba8ed8bb5efe91e97eb7c030808825fb5
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 encrypted.
To display the IPsec security associations (SAs) used on OSPFv3 interfaces, enter the following command:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>show crypto ipsec sa ipv6</code></td>
<td>EXEC Privilege</td>
<td>Displays security associations set up for OSPFv3 links in IPsec authentication and encryption policies on the router.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>To display information on the SAs used on a specific interface, enter <code>interface interface</code>, where <code>interface</code> is one of the following</td>
</tr>
<tr>
<td></td>
<td></td>
<td>values:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>For a 1-Gigabit Ethernet interface, enter <code>GigabitEthernet slot/port</code>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>For a Port Channel interface, enter <code>port-channel number</code>. Valid port-channel numbers (on an E-Series TeraScale):</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 to 255.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>For a 10-Gigabit Ethernet interface, enter <code>TenGigabitEthernet slot/port</code>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>For a VLAN interface, enter <code>vlan vlan-id</code>. Valid VLAN IDs: 1 to 4094.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Figure 34-27. 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)
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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>show ipv6 route summary</td>
<td>EXEC Privilege</td>
<td>View the summary information of the IPv6 routes.</td>
</tr>
<tr>
<td>show ipv6 ospf database</td>
<td>EXEC Privilege</td>
<td>View the summary information for the OSPFv3 database.</td>
</tr>
</tbody>
</table>

Use the following command in EXEC Privilege mode to view the OSPF configuration for a neighboring router:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>show ipv6 ospf neighbor</td>
<td>EXEC Privilege</td>
<td>View the configuration of OSPFv3 neighbors.</td>
</tr>
</tbody>
</table>
Use the following command in EXEC Privilege mode to configure the debugging options of an OSPFv3 process:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>`debug ipv6 ospf [event</td>
<td>packet] {type slot/port}`</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- <strong>event</strong>: View OSPF event messages.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- <strong>packet</strong>: View OSPF packets.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- For a Gigabit Ethernet interface, enter the keyword <strong>GigabitEthernet</strong> followed by the slot/port information (e.g., <code>passive-interface gi 2/1</code>).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- For a port channel, enter the keyword <strong>port-channel</strong> followed by a number from 1 to 255 for TeraScale and ExaScale.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- For a 10-Gigabit Ethernet interface, enter the keyword <strong>TenGigabitEthernet</strong> followed by the slot/port information (e.g., <code>passive-interface ten 2/3</code>).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- For a VLAN, enter the keyword <strong>vlan</strong> followed by a number from 1 to 4094 (e.g., <code>passive-interface vlan 2222</code>).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANs.</td>
</tr>
</tbody>
</table>
PIM Sparse-Mode (PIM-SM)

PIM Sparse-Mode (PIM-SM) is supported on platforms: E C S Z 54810

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 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.
- PIM-SM on VLAN interfaces is supported on the E-Series on TeraScale platforms only.
- IPv6 Multicast is not supported on SONET interfaces.

Protocol Overview

PIM-SM initially uses unidirectional shared trees to forward multicast traffic; that is, all multicast traffic must flow only from the Rendezvous Point (RP) to the receivers. Once a receiver receives traffic from the RP, PM-SM switches to shortest path trees (SPT) to forward multicast traffic. Every 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 multicast routing using the command `ip multicast-routing` from CONFIGURATION mode.
2. Select a Rendezvous Point.

**Related Configuration Tasks**

- Configurable S,G Expiry Timers on page 779
- Configure a Static Rendezvous Point on page 780
- Configure a Designated Router on page 781
- Create Multicast Boundaries and Domains on page 782
- PIM-SM Graceful Restart on page 782
- Monitoring PIM on page 783
Enable PIM-SM

You must enable PIM-SM on each participating interface:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enable multicast routing on the system.</td>
<td>ip multicast-routing</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Enable PIM-Sparse Mode</td>
<td>ip pim sparse-mode</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

Display which interfaces are enabled with PIM-SM using the command `show ip pim interface` from EXEC Privilege mode, as shown in Figure 35-1.

**Figure 35-1. 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 pim neighbor` from EXEC Privilege mode, as shown in Figure 35-2.

**Figure 35-2. 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 pim tib` from EXEC privilege mode, as shown in Figure 35-3.
Figure 35-3. 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: GigabitEtherent 4/12, RPF neighbor 10.87.3.5
  Outgoing interface list:
    GigabitEtherent 4/11
    GigabitEtherent 7/13

(10.87.31.5, 192.1.2.1), uptime 00:01:24, expires 00:02:26, flags: FT
  Incoming interface: GigabitEtherent 7/11, RPF neighbor 0.0.0.0
  Outgoing interface list:
    GigabitEtherent 4/11
    GigabitEtherent 4/12
    GigabitEtherent 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable global expiry timer for S, G entries</td>
<td>ip pim sparse-mode</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Range 211-86400 seconds</td>
<td>sg-expiry-timer seconds</td>
<td></td>
</tr>
<tr>
<td>Default: 210</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Configure the expiry time for a particular (S,G) entry:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create an Extended ACL</td>
<td>ip access-list extended access-list-name</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Specify the source and group to which the timer will be applied using extended ACLs with permit rules only.</td>
<td>[seq sequence-number] permit ip source-address/mask</td>
<td>CONFIG-EXT-NAACL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>{any</td>
<td>host source-address}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>{destination-address/mask</td>
<td>any</td>
</tr>
</tbody>
</table>
Figure 35-4. Configuring an (S,G) Expiry Time

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 35-5.

#### Figure 35-5. 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 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 pim rp` from EXEC privilege mode, as shown in Figure 35-6.

Figure 35-6. 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 35-7. 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
```

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.

- Assign a DR priority value using the command `ip pim dr-priority priority-value` from INTERFACE mode.
- Change the interval at which a router sends hello messages using the command `ip pim query-interval seconds` from INTERFACE mode.
- Display the current value of these parameter using the command `show ip pim interface` EXEC Privilege mode.
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, use the `ip 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 pim rp-mapping` command.

PIM-SM Graceful Restart

PIM-SM Graceful Restart is supported only on platform 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.

### 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.
Port Monitoring

Port Monitoring is supported on platforms: E C S Z S4810

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 on page 785
- Port Monitoring on E-Series on page 786
- Port Monitoring on C-Series and S-Series on page 787
- Configuring Port Monitoring on page 790
- Flow-based Monitoring on page 791

Important Points to Remember

- On the E-Series, Port Monitoring is supported on TeraScale and ExaScale platforms.
- 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 36-1 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 36-1. Maximum Number of Monitoring Sessions per System

<table>
<thead>
<tr>
<th>System</th>
<th>Maximum Sessions</th>
<th>System</th>
<th>Maximum Sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>C150</td>
<td>∞ (Note)</td>
<td>E1200/E1200i (TeraScale)</td>
<td>28</td>
</tr>
<tr>
<td>C300</td>
<td>∞ (Note)</td>
<td>E1200i (ExaScale)</td>
<td>∞</td>
</tr>
<tr>
<td>S50V, S50N</td>
<td>∞ (Note)</td>
<td>E600/E600i (TeraScale)</td>
<td>14</td>
</tr>
<tr>
<td>S25P</td>
<td>∞ (Note)</td>
<td>E600i (ExaScale)</td>
<td>∞</td>
</tr>
<tr>
<td>S55</td>
<td>∞ (Note)</td>
<td>E300</td>
<td>6</td>
</tr>
<tr>
<td>S60</td>
<td>∞ (Note)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4810</td>
<td>∞ (Note)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 36-1 illustrates a possible port monitoring configuration on the E-Series.

**Figure 36-1. Port Monitoring Configurations on the E-Series**

<table>
<thead>
<tr>
<th>Line Card 0</th>
<th>Line Card 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Port-Pipe 0</td>
<td>Port-Pipe 1</td>
</tr>
<tr>
<td>Monitor Session 0</td>
<td>MD  →  MG</td>
</tr>
<tr>
<td>Monitor Session 1</td>
<td>MD  →  MG</td>
</tr>
<tr>
<td>Monitor Session 2</td>
<td>MD  →  MG</td>
</tr>
<tr>
<td>Monitor Session 3</td>
<td>MD  →  MG</td>
</tr>
</tbody>
</table>

**E-Series ExaScale**

FTOS on E-Series ExaScale supports a single destination (MG) port monitoring 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 used 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).

<table>
<thead>
<tr>
<th>SessionID</th>
<th>Source</th>
<th>Destination</th>
<th>Direction</th>
<th>Mode</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Gi 0/13</td>
<td>Gi 0/1</td>
<td>rx</td>
<td>interface</td>
<td>Port-based</td>
</tr>
<tr>
<td>10</td>
<td>Gi 0/14</td>
<td>Gi 0/2</td>
<td>rx</td>
<td>interface</td>
<td>Port-based</td>
</tr>
<tr>
<td>20</td>
<td>Gi 0/15</td>
<td>Gi 0/3</td>
<td>rx</td>
<td>interface</td>
<td>Port-based</td>
</tr>
<tr>
<td>30</td>
<td>Gi 0/16</td>
<td>Gi 0/37</td>
<td>rx</td>
<td>interface</td>
<td>Port-based</td>
</tr>
</tbody>
</table>

Figure 36-2. Number of Monitoring Ports on the C-Series and S-Series

In Figure 36-2, 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 36-3.

<table>
<thead>
<tr>
<th>SessionID</th>
<th>Source</th>
<th>Destination</th>
<th>Direction</th>
<th>Mode</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Gi 0/13</td>
<td>Gi 0/1</td>
<td>rx</td>
<td>interface</td>
<td>Port-based</td>
</tr>
<tr>
<td>10</td>
<td>Gi 0/14</td>
<td>Gi 0/2</td>
<td>rx</td>
<td>interface</td>
<td>Port-based</td>
</tr>
<tr>
<td>20</td>
<td>Gi 0/15</td>
<td>Gi 0/3</td>
<td>rx</td>
<td>interface</td>
<td>Port-based</td>
</tr>
<tr>
<td>30</td>
<td>Gi 0/16</td>
<td>Gi 0/37</td>
<td>rx</td>
<td>interface</td>
<td>Port-based</td>
</tr>
<tr>
<td>300</td>
<td>Gi 0/17</td>
<td>Gi 0/1</td>
<td>tx</td>
<td>interface</td>
<td>Port-based</td>
</tr>
</tbody>
</table>

Figure 36-3. Number of Monitoring Ports on the C-Series and S-Series

In Figure 36-4, 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.
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 36-1 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 36-5. 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Verify that the intended monitoring port has no configuration other than no shutdown, as shown in Figure 36-6.</td>
<td>show interface</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>2</td>
<td>Create a monitoring session using the command monitor session from CONFIGURATION mode, as shown in Figure 36-6.</td>
<td>monitor session</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>3</td>
<td>Specify the source and destination port and direction of traffic, as shown in Figure 36-6.</td>
<td>source</td>
<td>MONITOR SESSION</td>
</tr>
</tbody>
</table>

Display monitor sessions using the command show monitor session from EXEC Privilege mode, as shown in Figure 36-6.

**Figure 36-6. 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

<table>
<thead>
<tr>
<th>SessionID</th>
<th>Source</th>
<th>Destination</th>
<th>Direction</th>
<th>Mode</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Gi 1/1</td>
<td>Gi 1/2</td>
<td>rx</td>
<td>interface</td>
<td>Port-based</td>
</tr>
</tbody>
</table>
```

In Figure 36-7, 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 36-7. Port Monitoring Example**
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Enable flow-based monitoring for a monitoring session.</td>
<td>flow-based enable</td>
<td>MONITOR SESSION</td>
</tr>
<tr>
<td>5</td>
<td>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 7, Access Control Lists (ACLs).</td>
<td>ip access-list</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>6</td>
<td>Apply the ACL to the monitored port. See Chapter 7, Access Control Lists (ACLs).</td>
<td>ip access-group access-list</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

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 36-8.

FTOS(conf-if-gi-1/2)#show config
!
interface GigabitEthernet 1/2
no ip address
no shutdown
FTOS(conf)#monitor session 0
FTOS(conf-mon-sess-0)#source gig 1/1 destination gig 1/2 direction rx

Port Monitoring 001
Figure 36-8. Configuring Flow-based Monitoring

```
FTOS(conf)#monitor session 0
FTOS(conf-mon-sess-0)#flow-based enable
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 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 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

<table>
<thead>
<tr>
<th>SessionID</th>
<th>Source</th>
<th>Destination</th>
<th>Direction</th>
<th>Mode</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Gi 1/1</td>
<td>Gi 1/2</td>
<td>rx</td>
<td>interface</td>
<td>Flow-based</td>
</tr>
</tbody>
</table>
```
The **Private VLANs (PVLAN)** feature is supported on platforms C S Z S4810.

For syntax details on the commands discussed in this chapter, see the Private VLANs Commands chapter in the *FTOS Command Line Reference*.

This chapter contains the following major sections:

- Private VLAN Concepts on page 793
- Private VLAN Commands on page 795
- Private VLAN Configuration Task List on page 796
- Private VLAN Configuration Example on page 799
- Inspecting the Private VLAN Configuration on page 800

Private VLANs extend the FTOS security suite by providing Layer 2 isolation between ports within the same VLAN. A private VLAN partitions a traditional VLAN into subdomains identified by a primary and secondary VLAN pair. Private VLANs block all traffic to isolated ports except traffic from promiscuous ports. Traffic received from an isolated port is forwarded only to promiscuous ports or trunk ports.

Example uses of PVLANs:

- 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.
- 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.

In more detail, community VLANs are especially useful in the service provider environment, because, multiple customers are likely to maintain servers that must be strictly separated in customer-specific groups. A set of servers owned by a customer could comprise a community VLAN, so that those servers could communicate with each other, and would be isolated from other customers. Another customer might have another set of servers in another community VLAN. Another customer might want an isolated VLAN, which is has one or more ports that are also isolated from each other.

## Private VLAN Concepts

The VLAN types in a private VLAN (PVLAN) include:

**Community VLAN** — A community VLAN is a type of secondary VLAN in a primary VLAN:
Private VLANs (PVLAN)

- Ports in a community VLAN can communicate with each other.
- Ports in a community VLAN can communicate with all promiscuous ports in the primary VLAN.
- A community VLAN can only contain ports configured as **host**.

**Isolated VLAN** — An *isolated VLAN* is a type of secondary VLAN in a primary VLAN:

- Ports in an isolated VLAN cannot talk directly to each other.
- Ports in an isolated VLAN can only communicate with promiscuous ports in the primary VLAN.
- An isolated VLAN can only contain ports configured as **host**.

**Primary VLAN** — A *primary VLAN* is the base VLAN of a private VLAN:

- A switch can have one or more primary VLANs, and it can have none.
- A primary VLAN has one or more secondary VLANs.
- A primary VLAN and each of its secondary VLANs decrement the available number of VLAN IDs in the switch.
- A primary VLAN has one or more promiscuous ports.
- A primary VLAN might have one or more trunk ports, or none.

**Secondary VLAN** — A *secondary VLAN* is a subdomain of the primary VLAN. There are two types of secondary VLAN — community VLAN and isolated VLAN.

PVLAN port types:

- **Community port**: A *community port* is, by definition, a port that belongs to a community VLAN and is allowed to communicate with other ports in the same community VLAN and with promiscuous ports.
- **Host port**: A *host port*, in the context of a private VLAN, is a port in a secondary VLAN:
  - The port must first be assigned that role in INTERFACE mode.
  - A port assigned the host role cannot be added to a regular VLAN.
- **Isolated port**: An *isolated port* is, by definition, a port that, in Layer 2, can only communicate with promiscuous ports that are in the same PVLAN.
- **Promiscuous port**: A *promiscuous port* is, by definition, a port that is allowed to communicate with any other port type in the PVLAN:
  - A promiscuous port can be part of more than one primary VLAN.
  - A promiscuous port cannot be added to a regular VLAN.
- **Trunk port**: A *trunk port*, by definition, carries traffic between switches:
  - A trunk port in a PVLAN is always tagged.
  - Primary or secondary VLAN traffic is carried by the trunk port in tagged mode. The tag on the packet helps identify the VLAN to which the packet belongs.
  - A trunk port can also belong to a regular VLAN (non-private VLAN).

Each of the port types can be any type of physical Ethernet port, including port channels (LAGs). For details on port channels, see *Port Channel Interfaces on page 480* in Chapter 23, Interfaces.

For an introduction to VLANs, see Chapter 29, Layer 2.
### Private VLAN Commands

The commands dedicated to supporting the Private VLANs feature are:

**Table 37-1. Private VLAN Commands**

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable/disable Layer 3 communication between secondary VLANs.</td>
<td>[no] ip local-proxy-arp</td>
<td>INTERFACE VLAN</td>
</tr>
<tr>
<td><strong>Note:</strong> Even after ip-local-proxy-arp is disabled (no ip-local-proxy-arp) in a secondary VLAN, Layer 3 communication may happen between some secondary VLAN hosts, until the ARP timeout happens on those secondary VLAN hosts.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set the mode of the selected VLAN to community, isolated, or primary.</td>
<td>[no] private-vlan mode {community</td>
<td>isolated</td>
</tr>
<tr>
<td>Map secondary VLANs to the selected primary VLAN.</td>
<td>[no] private-vlan mapping secondary-vlan vlan-list</td>
<td>INTERFACE VLAN</td>
</tr>
<tr>
<td>Display type and status of PVLAN interfaces.</td>
<td>show interfaces private-vlan [interface interface]</td>
<td>EXEC \n EXEC Privilege</td>
</tr>
<tr>
<td>Display PVLANs and/or interfaces that are part of a PVLAN.</td>
<td>show vlan private-vlan [community</td>
<td>interface</td>
</tr>
<tr>
<td>Display primary-secondary VLAN mapping.</td>
<td>show vlan private-vlan mapping</td>
<td>EXEC \n EXEC Privilege</td>
</tr>
<tr>
<td>Set the PVLAN mode of the selected port.</td>
<td>switchport mode private-vlan {host</td>
<td>promiscuous</td>
</tr>
</tbody>
</table>

**Note:** Secondary VLANs are Layer 2 VLANs, so even if they are operationally down while primary VLANs are operationally up, Layer 3 traffic will still be transmitted across secondary VLANs.

The outputs of the following commands are augmented in FTOS 7.8.1.0 to provide PVLAN data:

- **show arp**: See the IP Routing Commands chapter in the *FTOS Command Line Reference*.
- **show vlan**: See the Layer 2 Commands chapter in the *FTOS Command Line Reference*. 
Private VLAN Configuration Task List

The following sections contain the procedures that configure a private VLAN:

• Creating PVLAN ports
• Creating a Primary VLAN on page 797
• Creating a Community VLAN on page 798
• Creating an Isolated VLAN on page 798

Creating PVLAN ports

Private VLAN ports are those that will be assigned to the private VLAN (PVLAN).

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interface interface</td>
<td>CONFIGURATION</td>
<td>Access the INTERFACE mode for the port that you want to assign to a PVLAN.</td>
</tr>
<tr>
<td>2</td>
<td>no shutdown</td>
<td>INTERFACE</td>
<td>Enable the port.</td>
</tr>
<tr>
<td>3</td>
<td>switchport</td>
<td>INTERFACE</td>
<td>Set the port in Layer 2 mode.</td>
</tr>
<tr>
<td>4</td>
<td>switchport mode private-vlan {host</td>
<td>promiscuous</td>
<td>trunk}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• host (port in isolated or community VLAN)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• promiscuous (intra-VLAN communication port)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• trunk (inter-switch PVLAN hub port)</td>
</tr>
</tbody>
</table>

Note: Interfaces that are configured as PVLAN ports cannot be added to regular VLANs. Conversely, “regular” ports (ports not configured as PVLAN ports) cannot be added to PVLANs.

Figure 37-1 shows the use of the switchport mode private-vlan command on a port and on a port channel:

Figure 37-1. Examples of switchport mode private-vlan Command

```
FTOS#conf
FTOS(conf)#interface GigabitEthernet 2/1
FTOS(conf-if-gi-2/1)#switchport mode private-vlan promiscuous

FTOS(conf)#interface GigabitEthernet 2/2
FTOS(conf-if-gi-2/2)#switchport mode private-vlan host

FTOS(conf)#interface GigabitEthernet 2/3
FTOS(conf-if-gi-2/3)#switchport mode private-vlan trunk

FTOS(conf)#interface port-channel 10
FTOS(conf-if-po-10)#switchport mode private-vlan promiscuous
```
Creating a Primary VLAN

A primary VLAN is a port-based VLAN that is specifically enabled as a primary VLAN to contain the promiscuous ports and PVLAN trunk ports for the private VLAN. A primary VLAN also contains a mapping to secondary VLANs, which are comprised of community VLANs and isolated VLANs.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interface vlan vlan-id</td>
<td>CONFIGURATION</td>
<td>Access the INTERFACE VLAN mode for the VLAN to which you want to assign the PVLAN interfaces.</td>
</tr>
<tr>
<td>2</td>
<td>no shutdown</td>
<td>INTERFACE VLAN</td>
<td>Enable the VLAN.</td>
</tr>
<tr>
<td>3</td>
<td>private-vlan mode primary</td>
<td>INTERFACE VLAN</td>
<td>Set the PVLAN mode of the selected VLAN to primary.</td>
</tr>
</tbody>
</table>
| 4    | private-vlan mapping secondary-vlan vlan-list | INTERFACE VLAN | Map secondary VLANs to the selected primary VLAN. The list of secondary VLANs can be:  
- Specified in comma-delimited (VLAN-ID,VLAN-ID) or hyphenated-range format (VLAN-ID-VLAN-ID).  
- Specified with this command even before they have been created.  
- Amended by specifying the new secondary VLAN to be added to the list. |
| 5    | tagged interface or untagged interface | INTERFACE VLAN | Add promiscuous ports as tagged or untagged interfaces. Add PVLAN trunk ports to the VLAN only as tagged interfaces. Interfaces can be entered singly or in range format, either comma-delimited (slot/port,port,port) or hyphenated (slot/port-port).  
Only promiscuous ports or PVLAN trunk ports can be added to the PVLAN (no host or regular ports). |
| 6    | ip address ip address | INTERFACE VLAN | (OPTIONAL) Assign an IP address to the VLAN. |
| 7    | ip local-proxy-arp | INTERFACE VLAN | (OPTIONAL) Enable/disable Layer 3 communication between secondary VLANs. |

Note: 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.
Creating a Community VLAN

A community VLAN is a secondary VLAN of the primary VLAN in a private VLAN. The ports in a community VLAN can talk to each other and with the promiscuous ports in the primary VLAN.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interface vlan vlan-id</td>
<td>CONFIGURATION</td>
<td>Access the INTERFACE VLAN mode for the VLAN that you want to make a community VLAN.</td>
</tr>
<tr>
<td>2</td>
<td>no shutdown</td>
<td>INTERFACE VLAN</td>
<td>Enable the VLAN.</td>
</tr>
<tr>
<td>3</td>
<td>private-vlan mode community</td>
<td>INTERFACE VLAN</td>
<td>Set the PVLAN mode of the selected VLAN to community.</td>
</tr>
<tr>
<td>4</td>
<td>tagged interface or untagged interface</td>
<td>INTERFACE VLAN</td>
<td>Add one or more host ports to the VLAN. The interfaces can be entered singly or in range format, either comma-delimited (slot/port,port,port) or hyphenated (slot/port-port). Only host (isolated) ports can be added to the VLAN.</td>
</tr>
</tbody>
</table>

Creating an Isolated VLAN

An isolated VLAN is a secondary VLAN of a primary VLAN. Its ports can only talk with the promiscuous ports in that primary VLAN.

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interface vlan vlan-id</td>
<td>CONFIGURATION</td>
<td>Access the INTERFACE VLAN mode for the VLAN that you want to make an isolated VLAN.</td>
</tr>
<tr>
<td>2</td>
<td>no shutdown</td>
<td>INTERFACE VLAN</td>
<td>Enable the VLAN.</td>
</tr>
<tr>
<td>3</td>
<td>private-vlan mode isolated</td>
<td>INTERFACE VLAN</td>
<td>Set the PVLAN mode of the selected VLAN to isolated.</td>
</tr>
<tr>
<td>4</td>
<td>tagged interface or untagged interface</td>
<td>INTERFACE VLAN</td>
<td>Add one or more host ports to the VLAN. The interfaces can be entered singly or in range format, either comma-delimited (slot/port,port,port) or hyphenated (slot/port-port). Only ports defined as host can be added to the VLAN.</td>
</tr>
</tbody>
</table>

Figure 37-2 shows the use of the PVLAN commands that are used in VLAN INTERFACE mode to configure the PVLAN member VLANs (primary, community, and isolated VLANs):
Private VLAN Configuration Example

The following configuration is based on the example diagram, above:

On C300-1:

- Gi 0/0 and Gi 23 are configured as promiscuous ports, assigned to the primary VLAN, VLAN 4000.
- Gi 0/25 is configured as a PVLAN trunk port, also assigned to the primary VLAN 4000.
- Gi 0/24 and Gi 0/47 are configured as host ports and assigned to the isolated VLAN, VLAN 4003.
- Gi 4/0 and Gi 23 are configured as host ports and assigned to the community VLAN, VLAN 4001.
- Gi 4/24 and Gi 4/47 are configured as host ports and assigned to community VLAN 4002.
The result is that:

- The ports in community VLAN 4001 can communicate directly with each other and with promiscuous ports.
- The ports in community VLAN 4002 can communicate directly with each other and with promiscuous ports.
- The ports in isolated VLAN 4003 can only communicate with the promiscuous ports in the primary VLAN 4000.
- All the ports in the secondary VLANs (both community and isolated VLANs) can only communicate with ports in the other secondary VLANs of that PVLAN over Layer 3, and only when the command `ip local-proxy-arp` is invoked in the primary VLAN.

**Note:** Even after `ip-local-proxy-arp` is disabled (no `ip-local-proxy-arp`) in a secondary VLAN, Layer 3 communication may happen between some secondary VLAN hosts, until the ARP timeout happens on those secondary VLAN hosts.

In parallel, on S50-1:

- Gi 0/3 is a promiscuous port and Gi 0/25 is a PVLAN trunk port, assigned to the primary VLAN 4000.
- Gi 0/4-6 are host ports. Gi 0/4 and Gi 0/5 are assigned to the community VLAN 4001, while Gi 0/6 is assigned to the isolated VLAN 4003.

The result is that:

- The S50V ports would have the same intra-switch communication characteristics as described above for the C300.
- For transmission between switches, tagged packets originating from host PVLAN ports in one secondary VLAN and destined for host PVLAN ports in the other switch travel through the promiscuous ports in the local VLAN 4000 and then through the trunk ports (0/25 in each switch).

### Inspecting the Private VLAN Configuration

The standard methods of inspecting configurations also apply in PVLANs:

- Within the INTERFACE and INTERFACE VLAN modes, use the `show config` command to display the specific interface configuration.
- Inspect the running-config, and, with the `grep` pipe option (`show running-config | grep string`), you can display a specific part of the running-config. Figure 37-8 shows the PVLAN parts of the running-config from the S50V switch in the topology diagram shown in Figure 37-3, above.
- You can also use one of three `show` commands that are specific to the Private VLAN feature:
  - `show interfaces private-vlan [interface interface]`: Display the type and status of the configured PVLAN interfaces. See the example output in the Security chapter of the FTOS Command Line Reference.
  - `show vlan private-vlan [community | interface | isolated | primary | primary_vlan | interface interface]`: Display the configured PVLANs or interfaces that are part of a PVLAN. Figure 37-4 shows the results of using the command without command options on the C300 switch in the topology diagram shown in Figure 37-3, above, while Figure 37-5 shows the results on the S50V.
• **show vlan private-vlan mapping**: Display the primary-secondary VLAN mapping. See the example output from the S50V, above, in Figure 37-6.

• Two **show** commands revised to display PVLAN data are:
  - **show arp**
  - **show vlan**: See revised output in Figure 37-7.

**Figure 37-4. show vlan private-vlan Example Output from C300**

<table>
<thead>
<tr>
<th>Primary</th>
<th>Secondary</th>
<th>Type</th>
<th>Active Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>4000</td>
<td>Primary</td>
<td>Yes</td>
<td>Gi 0/0,23,25</td>
</tr>
<tr>
<td>4001</td>
<td>Community</td>
<td>Yes</td>
<td>Gi 4/0,23</td>
</tr>
<tr>
<td>4002</td>
<td>Community</td>
<td>Yes</td>
<td>Gi 4/24,47</td>
</tr>
<tr>
<td>4003</td>
<td>Isolated</td>
<td>Yes</td>
<td>Gi 0/24,47</td>
</tr>
</tbody>
</table>

**Figure 37-5. show vlan private-vlan Example Output from S50V**

<table>
<thead>
<tr>
<th>Primary</th>
<th>Secondary</th>
<th>Type</th>
<th>Active Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>4000</td>
<td>Primary</td>
<td>Yes</td>
<td>Gi 0/3,25</td>
</tr>
<tr>
<td>4001</td>
<td>Community</td>
<td>Yes</td>
<td>Gi 0/4-5</td>
</tr>
<tr>
<td>4003</td>
<td>Isolated</td>
<td>Yes</td>
<td>Gi 0/6</td>
</tr>
</tbody>
</table>

**Figure 37-6. show vlan private-vlan mapping Example Output from S50V**

```
S50-1#show vlan private-vlan mapping
Private Vlan:
 Primary : 4000
 Isolated : 4003
 Community : 4001
```

In the following screenshot, note the addition of the PVLAN codes – P, I, and C – in the left column:

**Figure 37-7. show vlan Example Output from S50V**

```
S50V#show vlan
Codes: * - Default VLAN, G - GVRP VLANs, P - Primary, C - Community, I - Isolated
Q: U - Untagged, T - Tagged
   x - Dot1x untagged, X - Dot1x tagged
   G - GVRP tagged, M - Vlan-stack

<table>
<thead>
<tr>
<th>NUM</th>
<th>Status</th>
<th>Description</th>
<th>Q Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>Inactive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Inactive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Inactive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>200</td>
<td>Inactive primary VLAN in PVLAN</td>
<td>T Gi 0/19-20</td>
</tr>
<tr>
<td>I</td>
<td>201</td>
<td>Inactive isolated VLAN in VLAN 200</td>
<td>T Gi 0/21</td>
</tr>
</tbody>
</table>
```
Figure 37-8. Example running-config Output of PVLAN Configuration from S50V

```plaintext
interface GigabitEthernet 0/3
  no ip address
  switchport
  switchport mode private-vlan promiscuous
  no shutdown

interface GigabitEthernet 0/4
  no ip address
  switchport
  switchport mode private-vlan host
  no shutdown

interface GigabitEthernet 0/5
  no ip address
  switchport
  switchport mode private-vlan host
  no shutdown

interface GigabitEthernet 0/6
  no ip address
  switchport
  switchport mode private-vlan host
  no shutdown

interface GigabitEthernet 0/25
  no ip address
  switchport
  switchport mode private-vlan trunk
  no shutdown

interface Vlan 4000
  private-vlan mode primary
  private-vlan mapping secondary-vlan 4001-4003
  no ip address
  tagged GigabitEthernet 0/3,25
  no shutdown

interface Vlan 4001
  private-vlan mode community
```
Per-VLAN Spanning Tree Plus (PVST+)

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 50, Spanning Tree Protocol (STP).

Figure 38-1. Per-VLAN Spanning Tree

FTOS supports three other variations of Spanning Tree, as shown in Table 38-1.

Table 38-1. FTOS Supported Spanning Tree Protocols

<table>
<thead>
<tr>
<th>Dell Force10 Term</th>
<th>IEEE Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanning Tree Protocol (STP)</td>
<td>802.1d</td>
</tr>
<tr>
<td>Rapid Spanning Tree Protocol (RSTP)</td>
<td>802.1w</td>
</tr>
</tbody>
</table>
Table 38-1. FTOS Supported Spanning Tree Protocols

<table>
<thead>
<tr>
<th>Dell Force10 Term</th>
<th>IEEE Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple Spanning Tree Protocol (MSTP)</td>
<td>802.1s</td>
</tr>
<tr>
<td>Per-VLAN Spanning Tree Plus (PVST+)</td>
<td>Third Party</td>
</tr>
</tbody>
</table>

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 38-2). 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 285.
- On the C-Series and S-Series, you can enable PVST+ on 254 VLANs. To set up VLANs, see Chapter 54, Virtual LANs (VLAN).

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 805.
4. Optionally, for load balancing, select a non-default bridge-priority for a VLAN. See page 805.

Related Configuration Tasks

- Modify Global PVST+ Parameters on page 807
- Modify Interface PVST+ Parameters on page 808
- Configure an EdgePort on page 809
- Flush MAC Addresses after a Topology Change on page 699
- Preventing Network Disruptions with BPDU Guard on page 1011
- SNMP Traps for Root Elections and Topology Changes on page 1016
- Configuring Spanning Trees as Hitless on page 1017
- PVST+ in Multi-vendor Networks on page 810
- PVST+ Extended System ID on page 810
- PVST+ Sample Configurations on page 811
Enable PVST+

When you enable PVST+, FTOS instantiates STP on each active VLAN. To enable PVST+ globally:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter PVST context.</td>
<td>protocol spanning-tree pvst</td>
<td>PROTOCOL PVST</td>
</tr>
<tr>
<td>2</td>
<td>Enable PVST+.</td>
<td>no disable</td>
<td>PROTOCOL PVST</td>
</tr>
</tbody>
</table>

Disable PVST+

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disable PVST+ globally.</td>
<td>disable</td>
<td>PROTOCOL PVST</td>
</tr>
<tr>
<td>Disable PVST+ on an interface, or remove a PVST+ parameter configuration.</td>
<td>no spanning-tree pvst</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

Display your PVST+ configuration by entering the command show config from PROTOCOL PVST context, as shown in fig.

Figure 38-2. Display the PVST+ Configuration

```
FTOS_E600(conf-pvst)#show config verbose
!
protocol spanning-tree pvst
no disable
vlan 100 bridge-priority 4096
```

Influence PVST+ Root Selection

In Figure 38-1, all VLANs use the same forwarding topology because R2 is elected the root, and all GigabitEthernet ports have the same cost. Figure 38-3 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.
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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assign a bridge priority.</td>
<td><code>vlan bridge-priority</code></td>
<td>PROTOCOL PVST</td>
</tr>
<tr>
<td>Range: 0 to 61440</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Default: 32768</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Display the PVST+ forwarding topology by entering the command `show spanning-tree pvst [vlan vlan-id]` from EXEC Privilege mode, as shown in Figure 38-4.
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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change the forward-delay parameter.</td>
<td>vlan forward-delay</td>
<td>PROTOCOL PVST</td>
</tr>
<tr>
<td>• Range: 4 to 30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Default: 15 seconds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change the hello-time parameter.</td>
<td>vlan hello-time</td>
<td>PROTOCOL PVST</td>
</tr>
<tr>
<td>Note: With large configurations (especially those with more ports) Dell Force10 recommends that you increase the hello-time.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range: 1 to 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Default: 2 seconds</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 38-4. Display the PVST+ Forwarding Topology

```plaintext
FTOS_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
BPU 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
```
The values for global PVST+ parameters are given in the output of the command `show spanning-tree pvst`, as shown in Figure 38-4.

### 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 38-2 lists the default values for port cost by interface.

#### Table 38-2.  PVST+ Default Port Cost Values

<table>
<thead>
<tr>
<th>Port Cost</th>
<th>Default Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>100-Mb/s Ethernet interfaces</td>
<td>200000</td>
</tr>
<tr>
<td>1-Gigabit Ethernet interfaces</td>
<td>20000</td>
</tr>
<tr>
<td>10-Gigabit Ethernet interfaces</td>
<td>2000</td>
</tr>
<tr>
<td>Port Channel with 100 Mb/s Ethernet interfaces</td>
<td>180000</td>
</tr>
<tr>
<td>Port Channel with 1-Gigabit Ethernet interfaces</td>
<td>18000</td>
</tr>
<tr>
<td>Port Channel with 10-Gigabit Ethernet interfaces</td>
<td>1800</td>
</tr>
</tbody>
</table>

**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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change the port cost of an interface.</td>
<td>spanning-tree pvst vlan cost</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>Range: 0 to 200000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Default: see Table 38-2.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The values for interface PVST+ parameters are given in the output of the command `show spanning-tree pvst`, as shown in Figure 38-4.

**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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable EdgePort on an interface.</td>
<td>spanning-tree pvst edge-port [bpduguard</td>
<td>INTERFACE</td>
</tr>
<tr>
<td></td>
<td>shutdown-on-violation]</td>
<td></td>
</tr>
</tbody>
</table>

The EdgePort status of each interface is given in the output of the command `show spanning-tree pvst`, as shown in Figure 38-4.
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

This command places the port in a default state. 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 38-5, 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 misconfigurations.

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 38-5.  PVST+ with Extend System ID
Augment the Bridge ID with the VLAN 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
```

### PVST+ Sample Configurations

**Figure 38-6, Figure 38-7, and Figure 38-8** provide the running configurations for the topology shown in **Figure 38-3**.
Figure 38-6. 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 38-7. PVST+ Sample Configuration: R2 Running-configuration

```plaintext
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 38-8. PVST+ Sample Configuration: R3 Running-configuration

```plaintext
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
```
Quality of Service (QoS)

Quality of Service (QoS) is supported on platforms: E C S Z (S4810)

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 39-1. FTOS Support for Port-based, Policy-based, and Multicast QoS Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Platform</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Port-based QoS Configurations</strong></td>
<td>E C S</td>
<td>Ingress + Egress</td>
</tr>
<tr>
<td>Set dot1p Priorities for Incoming Traffic</td>
<td>E C S</td>
<td>Ingress</td>
</tr>
<tr>
<td>Honor dot1p Priorities on Ingress Traffic</td>
<td>E C S</td>
<td></td>
</tr>
<tr>
<td>Configure Port-based Rate Policing</td>
<td>E C S</td>
<td></td>
</tr>
<tr>
<td>Configure Port-based Rate Limiting</td>
<td>E</td>
<td>Egress</td>
</tr>
<tr>
<td>Configure Port-based Rate Shaping</td>
<td>E C S</td>
<td></td>
</tr>
<tr>
<td><strong>Policy-based QoS Configurations</strong></td>
<td>E C S</td>
<td>Ingress + Egress</td>
</tr>
<tr>
<td>Classify Traffic</td>
<td>E C S</td>
<td>Ingress</td>
</tr>
<tr>
<td>Create a Layer 3 class map</td>
<td>E C S</td>
<td></td>
</tr>
<tr>
<td>Set DSCP values for egress packets based on</td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>flow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Create a Layer 2 class map</td>
<td>E C S</td>
<td></td>
</tr>
<tr>
<td>Create a QoS Policy</td>
<td>E C S</td>
<td>Ingress + Egress</td>
</tr>
<tr>
<td>Feature</td>
<td>Platform</td>
<td>Direction</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>Create an input QoS policy</td>
<td>CS</td>
<td>Ingress</td>
</tr>
<tr>
<td>Configure policy-based rate policing</td>
<td>CS</td>
<td></td>
</tr>
<tr>
<td>Set a DSCP value for egress packets</td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>Set a dot1p value for egress packets</td>
<td>CS</td>
<td></td>
</tr>
<tr>
<td>Create an output QoS policy</td>
<td>CS</td>
<td>Egress</td>
</tr>
<tr>
<td>Configure policy-based rate limiting</td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>Configure policy-based rate shaping</td>
<td>CS</td>
<td></td>
</tr>
<tr>
<td>Allocate bandwidth to queue</td>
<td>CS</td>
<td></td>
</tr>
<tr>
<td>Specify WRED drop precedence</td>
<td>54810</td>
<td>E</td>
</tr>
<tr>
<td>Create Policy Maps</td>
<td>CS</td>
<td>Ingress + Egress</td>
</tr>
<tr>
<td>Create Input Policy Maps</td>
<td>CS</td>
<td>Ingress</td>
</tr>
<tr>
<td>Honor DSCP values on ingress packets</td>
<td>CS</td>
<td></td>
</tr>
<tr>
<td>Honoring dot1p values on ingress packets</td>
<td>E CS</td>
<td></td>
</tr>
<tr>
<td>Create Output Policy Maps</td>
<td>CS</td>
<td>Egress</td>
</tr>
<tr>
<td>Specify an aggregate QoS policy</td>
<td>CS</td>
<td></td>
</tr>
<tr>
<td>QoS Rate Adjustment</td>
<td>CS</td>
<td></td>
</tr>
<tr>
<td>Strict-priority Queueing</td>
<td>CS</td>
<td></td>
</tr>
<tr>
<td>Weighted Random Early Detection</td>
<td>54810</td>
<td>E</td>
</tr>
<tr>
<td>Create WRED Profiles</td>
<td>54810</td>
<td>E</td>
</tr>
<tr>
<td>Pre-calculating Available QoS CAM Space</td>
<td>CS</td>
<td></td>
</tr>
</tbody>
</table>
Implementation Information

The 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:

**Note:** Egress rate shaping and ingress rate policing cannot be simultaneously used on the same VLAN.
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 39-2. FTOS places traffic marked with a priority in a queue based on Table 39-2. 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 39-2).

<table>
<thead>
<tr>
<th>dot1p</th>
<th>E-Series Queue Number</th>
<th>C-Series Queue Number</th>
<th>S-Series Queue Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

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 39-3. You can configure this feature on physical interfaces and port-channels, but you cannot configure it on 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 39-2).
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 832.

**Note:** You cannot configure service-policy input and service-class dynamic dot1p on the same interface.

**Figure 39-3. 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 dot1p 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 INTERFACE mode, as shown in **Figure 39-4**. If the interface is a member of a VLAN, you may specify the VLAN for which ingress packets are policed.

**Figure 39-4. 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#
```
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.

Rate limit egress traffic on an interface using the command rate limit from INTERFACE mode, as shown in Figure 39-6. If the interface is a member of a VLAN, you may specify the VLAN for which egress packets are rate limited.

Figure 39-6. 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 39-7.
Configure Port-based Rate Shaping

Configure Port-based Rate Limiting is supported only on platform C E S.

FTOS Behavior: On the C-Series and S-Series, rate shaping is effectively rate limiting because of its smaller buffer size. On the S4810, rate shaping on tagged ports is slightly greater than the configured rate and rate shaping on untagged ports is slightly less than configured rate.

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 39-8.
- Apply rate shaping to a queue using the command `rate-shape` from QoS Policy mode.

Figure 39-8. Applying Rate Shaping to Outgoing Traffic

```
FTOS(config)
FTOS(conf-if)#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 39-9.

Figure 39-9. 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 39-10.
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 39-10. 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 39-10.

Figure 39-10. Using the Order Keyword in ACLs

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.

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 39-10, 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 39-10. 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**

Set DSCP values for egress packets based on flow is supported only on platform E.

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 826) 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 39-11. 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 39-11. 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
```

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
<table>
<thead>
<tr>
<th>Cam</th>
<th>Port</th>
<th>Dscp</th>
<th>Proto</th>
<th>Tcp</th>
<th>Src</th>
<th>Dst</th>
<th>SrcIp</th>
<th>DstIp</th>
<th>DSCP</th>
<th>Queue</th>
</tr>
</thead>
<tbody>
<tr>
<td>20416</td>
<td>1</td>
<td>18</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>23.64.0.5/32</td>
<td>0.0.0.0/0</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>20417</td>
<td>1</td>
<td>18</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>23.64.0.2/32</td>
<td>0.0.0.0/0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>20418</td>
<td>1</td>
<td>0</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>23.64.0.3/32</td>
<td>0.0.0.0/0</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>20419</td>
<td>1</td>
<td>0</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>23.64.0.2/32</td>
<td>0.0.0.0/0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>20420</td>
<td>1</td>
<td>0</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>23.64.0.3/32</td>
<td>0.0.0.0/0</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>20421</td>
<td>1</td>
<td>10</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>23.64.0.2/32</td>
<td>0.0.0.0/0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>20422</td>
<td>1</td>
<td>0</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>0.0.0.0/0</td>
<td>0.0.0.0/0</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>24511</td>
<td>1</td>
<td>0</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>0.0.0.0/0</td>
<td>0.0.0.0/0</td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>
```

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
<table>
<thead>
<tr>
<th>Cam</th>
<th>Port</th>
<th>Dscp</th>
<th>Proto</th>
<th>Tcp</th>
<th>Src</th>
<th>Dst</th>
<th>SrcIp</th>
<th>DstIp</th>
<th>DSCP</th>
<th>Queue</th>
</tr>
</thead>
<tbody>
<tr>
<td>20416</td>
<td>1</td>
<td>18</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>23.64.0.5/32</td>
<td>0.0.0.0/0</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>20417</td>
<td>1</td>
<td>0</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>23.64.0.2/32</td>
<td>0.0.0.0/0</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>20418</td>
<td>1</td>
<td>10</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>0.0.0.0/0</td>
<td>0.0.0.0/0</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>20419</td>
<td>1</td>
<td>0</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>0.0.0.0/0</td>
<td>0.0.0.0/0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>24511</td>
<td>1</td>
<td>0</td>
<td>IP</td>
<td>0x0</td>
<td>0</td>
<td>0</td>
<td>0.0.0.0/0</td>
<td>0.0.0.0/0</td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>
```
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.

**Note:** To avoid issues caused by misconfiguration, Dell Force10 recommends configuring either DCBX or Egress QoS features, but not both simultaneously. If both DCBX and Egress QoS are enabled at the same time, the DCBX configuration will be applied and unexpected behavior will occur on the Egress QoS.

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 is supported only on platform E
Set a DSCP value for egress packets based on ingress QoS classification, as shown in Figure 39-2. 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 39-12. Marking DSCP Values for Egress Packets

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.
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 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 39-3 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.

**Note:** Dell Force10 recommends assigning bandwidth to all queues. If queues are left un-allocated, the remaining bandwidth is shared equally among the un-allocated queues. If the sum of the allocated bandwidth percentage exceeds 100% 1% from the allocated queues will be assigned to each un-allocated queues.

<table>
<thead>
<tr>
<th>Table 39-3. Default Bandwidth Weights for C-Series and S-Series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queue</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

The key difference between allocating bandwidth by weight on the C-Series and S-Series and allocating bandwidth by percentage on the E-Series 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.

A key similarity between allocating bandwidth by percentage and allocating by weight is that assigning a weight or percentage 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.

**Table 39-4** shows an example of choosing bandwidth weights for all four queues to achieve a target bandwidth allocation.

<table>
<thead>
<tr>
<th>Table 39-4. Assigning Bandwidth Weights for the C-Series and S-Series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queue</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>
Specify WRED drop precedence

Specify WRED drop precedence is supported only on platform E 54810.

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.

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.

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 39-5 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.
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 39-6 specifies the queue to which the classified traffic is sent based on the dot1p value.

<table>
<thead>
<tr>
<th>dot1p</th>
<th>E-Series Queue ID</th>
<th>C-Series Queue ID</th>
<th>S-Series Queue ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

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 39-5.

The behavior is similar for trust dot1p fallback in a Layer2 input policy map; the dot1p-to-queue mapping is according to Table 39-6.
To enable Fall Back to trust diffserve or dot1p:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classify packets according to their DSCP value as a secondary option in case no match occurs against the configured class maps.</td>
<td>trust {diffserve</td>
<td>dot1p} fallback</td>
</tr>
</tbody>
</table>

**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 39-6 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 828). The `service-class 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 S4810

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
<td>qos-rate-adjust overhead-bytes</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Default: Disabled</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-Series and S-Series Range: 1-31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E-Series Range: 1-144</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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 S4810**

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 39-13); 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.
You can create a custom WRED profile or use one of the five pre-defined profiles.

### Table 39-7. Pre-defined WRED Profiles (E-Series)

<table>
<thead>
<tr>
<th>Default Profile Name</th>
<th>Minimum Threshold</th>
<th>Maximum Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>wred_drop</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>wred_ge_y</td>
<td>1024</td>
<td>2048</td>
</tr>
<tr>
<td>wred_ge_g</td>
<td>2048</td>
<td>4096</td>
</tr>
<tr>
<td>wred_teng_y</td>
<td>4096</td>
<td>8192</td>
</tr>
<tr>
<td>wred_teng_g</td>
<td>8192</td>
<td>16384</td>
</tr>
</tbody>
</table>

### Table 39-8. Pre-defined WRED Profiles (S4810)

<table>
<thead>
<tr>
<th>Default Profile Name</th>
<th>Minimum Threshold</th>
<th>Maximum Threshold</th>
<th>Maximum Drop Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>wred_drop</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>wred_teng_y</td>
<td>467</td>
<td>4671</td>
<td>100</td>
</tr>
<tr>
<td>wred_teng_g</td>
<td>467</td>
<td>4671</td>
<td>50</td>
</tr>
<tr>
<td>wred_fortyg_y</td>
<td>467</td>
<td>4671</td>
<td>50</td>
</tr>
<tr>
<td>wred_fortyg_g</td>
<td>467</td>
<td>4671</td>
<td>25</td>
</tr>
</tbody>
</table>

### 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 its 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 829) 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`.

**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 39-14.

**Figure 39-14. Displaying WRED Profiles (E-Series)**

```
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
```

**Figure 39-15. Displaying WRED Profiles (S4810)**

```
FTOS#show qos wred-profile

Wred-profile-name       min-threshold   max-threshold   max-drop-rate
wred_drop               0               0              100
wred_teng_y             467             4671            100
wred_teng_g             467             4671            50
wred_fortyg_y           467             4671            50
wred_fortyg_g           467             4671            25

FTOS#
```
Display WRED Drop Statistics

Display the number of packets FTOS dropped by WRED Profile using the command show qos statistics from EXEC Privilege mode.

Figure 39-16. show qos statistics Command Example (E-Series)

```
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.

Figure 39-17. show qos statistics Command Example (S4810)

```
FTOS#show qos statistics wred-profile
Interface Te 0/0
Drop-statistic  WRED-name     Dropped Pkts
Green           WRED1          51623
Yellow          WRED2          51300
Out of Profile
FTOS#
```
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 39-18, 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 39-18. test cam-usage Command Example**

```
FTOS# test cam-usage service-policy input pmap_l2 linecard 0 port-set 0

<table>
<thead>
<tr>
<th>Linecard</th>
<th>Port-pipe</th>
<th>CAM Partition</th>
<th>Available CAM</th>
<th>Estimated CAM</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>L2ACL</td>
<td>500</td>
<td>200</td>
<td>Allowed(2)</td>
</tr>
</tbody>
</table>
```
Routing Information Protocol (RIP)

Routing Information Protocol (RIP) is supported only on platforms: E C S Z S4810

Routing Information Protocol (RIP) is based on a distance-vector algorithm, it tracks distances or hop counts to nearby routers when establishing network connections.

- Protocol Overview on page 841
- Implementation Information on page 842
- Configuration Information on page 842
- RIP Configuration Example on page 850

RIP protocol standards are listed in the Chapter 57, 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 40-1 displays the defaults for RIP in FTOS.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interfaces running RIP</td>
<td>Listen to RIPv1 and RIPv2</td>
</tr>
<tr>
<td></td>
<td>Transmit RIPv1</td>
</tr>
<tr>
<td>RIP timers</td>
<td>update timer = 30 seconds</td>
</tr>
<tr>
<td></td>
<td>invalid timer = 180 seconds</td>
</tr>
<tr>
<td></td>
<td>holddown timer = 180 seconds</td>
</tr>
<tr>
<td></td>
<td>flush timer = 240 seconds</td>
</tr>
<tr>
<td>Auto summarization</td>
<td>Enabled</td>
</tr>
<tr>
<td>ECMP paths supported</td>
<td>16</td>
</tr>
</tbody>
</table>

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 843 (mandatory)
- Configure RIP on interfaces on page 844 (optional)
- Control RIP routing updates on page 844 (optional)
- Set send and receive version on page 846 (optional)
- Generate a default route on page 848 (optional)
• Control route metrics on page 849 (optional)
• Summarize routes on page 848 (optional)
• Control route metrics on page 849
• Debug RIP on page 849

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>router rip</td>
<td>CONFIGURATION</td>
<td>Enter ROUTER RIP mode and enable the RIP process on FTOS.</td>
</tr>
<tr>
<td>2</td>
<td>network ip-address</td>
<td>ROUTER RIP</td>
<td>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.</td>
</tr>
</tbody>
</table>

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 shown in the following example in the ROUTER RIP mode.

Figure 40-1.  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 as shown in the following example.
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.

---

**Figure 40-2. 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 control the source of RIP route information, use the following commands, in the ROUTER RIP mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>neighbor ip-address</td>
<td>ROUTER RIP</td>
<td>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.</td>
</tr>
<tr>
<td>passive-interface interface</td>
<td>ROUTER RIP</td>
<td>Disable a specific interface from sending or receiving RIP routing information.</td>
</tr>
</tbody>
</table>

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.

To apply prefix lists to incoming or outgoing RIP routes, use the following commands in the ROUTER RIP mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>distribute-list prefix-list-name in</td>
<td>ROUTER RIP</td>
<td>Assign a configured prefix list to all incoming RIP routes.</td>
</tr>
<tr>
<td>distribute-list prefix-list-name out</td>
<td>ROUTER RIP</td>
<td>Assign a configured prefix list to all outgoing RIP routes.</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>redistribute {connected</td>
<td>ROUTER RIP</td>
<td>Include directly connected or user-configured (static) routes in RIP.</td>
</tr>
<tr>
<td></td>
<td>static} [metric [metric-value]] [route-map map-name]</td>
<td></td>
</tr>
<tr>
<td>redistribute isis [level-1</td>
<td>ROUTER RIP</td>
<td>Include IS-IS routes in RIP.</td>
</tr>
<tr>
<td></td>
<td>level-1-2</td>
<td>level-2] [metric [metric-value]] [route-map map-name]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>`version {1</td>
<td>2}`</td>
<td>ROUTER RIP</td>
</tr>
</tbody>
</table>

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 40-3 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.
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip rip receive version [1] [2]</td>
<td>INTERFACE</td>
<td>Set the RIP version(s) received on that interface.</td>
</tr>
<tr>
<td>ip rip send version [1] [2]</td>
<td>INTERFACE</td>
<td>Set the RIP version(s) sent out on that interface.</td>
</tr>
</tbody>
</table>

To configure an interface to receive or send both versions of RIP, include 1 and 2 in the command syntax. Figure 40-4 displays the command syntax for sending both RIPv1 and RIPv2 and receiving only RIPv2.

Figure 40-4. 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 40-5 confirms that both versions are sent out that interface. This interface no longer sends and receives the same RIP versions as FTOS does globally.
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>default-information originate</td>
<td>ROUTER RIP</td>
<td>Specify the generation of a default route in RIP.</td>
</tr>
<tr>
<td>[always] [metric value]</td>
<td></td>
<td>Configure the following parameters:</td>
</tr>
<tr>
<td>[route-map route-map-name]</td>
<td></td>
<td>• always: enter this keyword to always generate a default route.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• value range: 1 to 16.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• route-map-name: name of a configured route map.</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>distance weight [ip-address mask]</td>
<td>ROUTER RIP</td>
<td>Apply a weight to all routes or a specific route and ACL. Configure the following parameters:</td>
</tr>
<tr>
<td>[access-list-name]</td>
<td></td>
<td>• weight range: 1 to 255 (default is 120)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• ip-address mask: the IP address in dotted decimal format (A.B.C.D), and the mask in slash format (/x).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• access-list-name: name of a configured IP ACL.</td>
</tr>
<tr>
<td>offset-list access-list-name {in</td>
<td>ROUTER RIP</td>
<td>Apply an additional number to the incoming or outgoing route metrics. Configure the following parameters:</td>
</tr>
<tr>
<td>out} offset [interface]</td>
<td></td>
<td>• prefix-list-name: the name of an established Prefix list to determine which incoming routes will be modified</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• offset range: 0 to 16.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• interface: the type, slot, and number of an interface.</td>
</tr>
</tbody>
</table>

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 40-6** shows the confirmation when the debug function is enabled.

**Figure 40-6. 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 40-7** — “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 851
- Core 2 Output on page 851
- RIP Configuration on Core 3 on page 853
- Core 3 RIP Output on page 853
- RIP Configuration Summary on page 855

**Figure 40-7. RIP Topology Example**
Configuring RIPv2 on Core 2

Figure 40-8. Configuring RIPv2 on Core 2

Core 2 Output

The screenshots in this section are:

- Figure 40-9: Using `show ip rip database` command to display Core 2 RIP database
- Figure 40-10: Using `show ip route` command to display Core 2 RIP setup
- Figure 40-11: Using `show ip protocols` command to display Core 2 RIP activity

Figure 40-9. Example of RIP Configuration Response from Core 2

```bash
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          auto-summary
[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          auto-summary
[120/1] via 10.11.20.1, 00:00:03, GigabitEthernet 2/31
192.168.2.0/24          auto-summary
Core2#
```
Figure 40-10. 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, NL - 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

<table>
<thead>
<tr>
<th>Destination</th>
<th>Gateway</th>
<th>Dist/Metric Last Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>C 10.11.10.0/24</td>
<td>Direct, Gi 2/11</td>
<td>0/0 00:02:26</td>
</tr>
<tr>
<td>C 10.11.20.0/24</td>
<td>Direct, Gi 2/31</td>
<td>0/0 00:02:02</td>
</tr>
<tr>
<td>R 10.11.30.0/24</td>
<td>via 10.11.20.1, Gi 2/31</td>
<td>120/1 00:01:20</td>
</tr>
<tr>
<td>C 10.200.10.0/24</td>
<td>Direct, Gi 2/41</td>
<td>0/0 00:03:03</td>
</tr>
<tr>
<td>C 10.300.10.0/24</td>
<td>Direct, Gi 2/42</td>
<td>0/0 00:02:42</td>
</tr>
<tr>
<td>R 192.168.1.0/24</td>
<td>via 10.11.20.1, Gi 2/31</td>
<td>120/1 00:01:20</td>
</tr>
<tr>
<td>R 192.168.2.0/24</td>
<td>via 10.11.20.1, Gi 2/31</td>
<td>120/1 00:01:20</td>
</tr>
<tr>
<td>Core2#</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 192.168.1.0/24</td>
<td>via 10.11.20.1, Gi 2/31</td>
<td>120/1 00:05:22</td>
</tr>
<tr>
<td>R 192.168.2.0/24</td>
<td>via 10.11.20.1, Gi 2/31</td>
<td>120/1 00:05:22</td>
</tr>
</tbody>
</table>

Core2#

Figure 40-11. 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

<table>
<thead>
<tr>
<th>Interface</th>
<th>Recv</th>
<th>Send</th>
</tr>
</thead>
<tbody>
<tr>
<td>GigabitEthernet 2/42</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>GigabitEthernet 2/41</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>GigabitEthernet 2/31</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>GigabitEthernet 2/11</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Routing for Networks:
- 10.300.10.0
- 10.200.10.0
- 10.11.20.0
- 10.11.10.0

Routing Information Sources:
- Gateway: 10.11.20.1, Distance: 120, Last Update: 00:00:12

Distance: (default is 120)

Core2#
RIP Configuration on Core 3

Figure 40-12. 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 40-13**: Using `show ip rip database` command to display Core 3 RIP database
- **Figure 40-14**: Using `show ip route` command to display Core 3 RIP setup
- **Figure 40-15**: Using `show ip protocols` command to display Core 3 RIP activity

Figure 40-13. 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 40-14. 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

<table>
<thead>
<tr>
<th>Destination</th>
<th>Gateway</th>
<th>Dist/Metric Last Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>R 10.11.10.0/24</td>
<td>via 10.11.20.2, Gi 3/21</td>
<td>120/1 00:01:14</td>
</tr>
<tr>
<td>C 10.11.20.0/24</td>
<td>Direct, Gi 3/21</td>
<td>0/0 00:01:53</td>
</tr>
<tr>
<td>C 10.11.30.0/24</td>
<td>Direct, Gi 3/11</td>
<td>0/0 00:06:00</td>
</tr>
<tr>
<td>R 10.200.10.0/24</td>
<td>via 10.11.20.2, Gi 3/21</td>
<td>120/1 00:01:14</td>
</tr>
<tr>
<td>R 10.300.10.0/24</td>
<td>via 10.11.20.2, Gi 3/21</td>
<td>120/1 00:01:14</td>
</tr>
<tr>
<td>C 192.168.1.0/24</td>
<td>Direct, Gi 3/43</td>
<td>0/0 00:06:53</td>
</tr>
<tr>
<td>C 192.168.2.0/24</td>
<td>Direct, Gi 3/44</td>
<td>0/0 00:06:26</td>
</tr>
</tbody>
</table>

Core3#

Figure 40-15. 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

<table>
<thead>
<tr>
<th>Interface</th>
<th>Recv</th>
<th>Send</th>
</tr>
</thead>
<tbody>
<tr>
<td>GigabitEthernet 3/21</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>GigabitEthernet 3/11</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>GigabitEthernet 3/44</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>GigabitEthernet 3/43</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Routing for Networks:
10.11.20.0
10.11.30.0
192.168.2.0
192.168.1.0

Routing Information Sources:

<table>
<thead>
<tr>
<th>Gateway</th>
<th>Distance</th>
<th>Last Update</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.11.20.2</td>
<td>120</td>
<td>00:00:22</td>
</tr>
</tbody>
</table>

Distance: (default is 120)

Core3#
RIP Configuration Summary

Figure 40-16. 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 40-17. 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
```
Remote Monitoring (RMON)

Remote Monitoring (RMON) is supported on platform: E C S Z S4810

This chapter describes the Remote Monitoring (RMON):

- Implementation on page 857
- Fault Recovery on page 858

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).

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 Chapter 57, 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 0xFFFFFFFFFFFFFFFF (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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[no] rmon alarm number variable interval {delta</td>
<td>absolute} rising-threshold [value event-number] falling-threshold value event-number [owner string]</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>or</td>
<td></td>
<td>• <code>number</code>: Alarm number, should be an integer from 1 to 65,535, the value must be unique in the RMON Alarm Table</td>
</tr>
<tr>
<td>[no] rmon hc-alarm number variable interval {delta</td>
<td>absolute} rising-threshold value event-number falling-threshold value event-number [owner string]</td>
<td></td>
</tr>
</tbody>
</table>

| 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 64-bits 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.
The above example configures RMON alarm number 10. The alarm monitors the MIB variable 1.3.6.1.2.1.2.1.20.1 (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 1.3.6.1.2.1.2.1.20.1 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 1.3.6.1.2.1.2.1.20.1 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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>[no] rmon event number [log] [trap community] [description string] [owner string]</code></td>
<td>CONFIGURATION</td>
<td><code>number</code>: 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. <code>log</code>: (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. <code>trap community</code>: (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”. <code>description string</code>: (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. <code>owner string</code>: (Optional) Owner of this event, which is identical to the eventOwner in the eventTable of the RMON MIB. Default is a null-terminated string.</td>
</tr>
</tbody>
</table>

The following example shows the `rmon event` command.
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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[no] rmon collection statistics {controlEntry integer} [owner ownername]</td>
<td>CONFIGURATION INTERFACE (config-if)</td>
<td>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</td>
</tr>
</tbody>
</table>

The following command enables the RMON statistics collection on the interface, with an ID value of 20 and an owner of “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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>[no] rmon collection history</td>
<td>CONFIGURATION</td>
<td>controlEntry: Specifies the RMON group of statistics using a value.</td>
</tr>
<tr>
<td>{controlEntry integer} [owner</td>
<td>INTERFACE</td>
<td>integer: A value from 1 to 65,535 that identifies the RMON group of</td>
</tr>
<tr>
<td>ownername] [buckets bucket-number]</td>
<td>(config-if)</td>
<td>statistics. The value must be a unique index in the RMON History Table.</td>
</tr>
<tr>
<td>[interval seconds]</td>
<td></td>
<td>owner: (Optional) Specifies the name of the owner of the RMON group of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>statistics. Default is a null-terminated string.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ownername: (Optional) Records the name of the owner of the RMON group</td>
</tr>
<tr>
<td></td>
<td></td>
<td>of statistics.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>buckets: (Optional) Specifies the maximum number of buckets desired for</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the RMON collection history group of statistics.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bucket-number: (Optional) A value associated with the number of buckets</td>
</tr>
<tr>
<td></td>
<td></td>
<td>specified for the RMON collection history group of statistics. The value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>is limited to from 1 to 1000. Default is 50 (as defined in RFC-2819).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>interval: (Optional) Specifies the number of seconds in each polling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cycle.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>seconds: (Optional) The number of seconds in each polling cycle. The</td>
</tr>
<tr>
<td></td>
<td></td>
<td>value is ranged from 5 to 3,600 (Seconds). Default is 1,800 as defined</td>
</tr>
<tr>
<td></td>
<td></td>
<td>in RFC-2819.</td>
</tr>
</tbody>
</table>

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 41-4. rmon collection history Command Example

```
FTOS(conf-if-mgmt)#rmon collection history controlEntry 20 owner john
```
Rapid Spanning Tree Protocol (RSTP)

Rapid Spanning Tree Protocol (RSTP) is supported on platforms: E C S Z S4810

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 42-1.

Table 42-1. FTOS Supported Spanning Tree Protocols

<table>
<thead>
<tr>
<th>Dell Force10 Term</th>
<th>IEEE Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanning Tree Protocol (STP)</td>
<td>802.1d</td>
</tr>
<tr>
<td>Rapid Spanning Tree Protocol (RSTP)</td>
<td>802.1w</td>
</tr>
<tr>
<td>Multiple Spanning Tree Protocol (MSTP)</td>
<td>802.1s</td>
</tr>
<tr>
<td>Per-VLAN Spanning Tree Plus (PVST+)</td>
<td>Third Party</td>
</tr>
</tbody>
</table>

Configuring Rapid Spanning Tree

Configuring Rapid Spanning Tree is a two-step process:

2. Enable Rapid Spanning Tree Protocol. See page 49.

Related Configuration Tasks

- Add and Remove Interfaces on page 869
- Modify Global Parameters on page 869
- Modify Interface Parameters on page 870
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.

RSTP and VLT

VLT provides loop-free redundant topologies and does not require rapid spanning tree protocol (RSTP). RSTP can cause temporary port state blocking and may cause topology changes after link or node failures. Spanning tree topology changes are distributed to the entire layer 2 network, which can cause a network-wide flush of learned MAC and ARP addresses, requiring these addresses to be re-learned. However, enabling RSTP can detect potential loops caused by non-system issues such as cabling errors or incorrect configurations. RSTP is useful for potential loop detection but should be configured using the specifications below to minimize possible topology changes after link or node failure.

The following recommendations will help you avoid these issues and the associated traffic loss caused by using rapid spanning trees when VLT is enabled on both VLT peers:

- Any ports at the edge of the spanning tree's operating domain should be configured as edge ports, which are directly connected to end stations or server racks. Ports connected directly to layer 3-only routers not running STP should have RSTP disabled or be configured as edge ports.
- Ensure the primary VLT node is the root bridge and the secondary VLT peer node has the second-best bridge ID in the network. If the primary VLT peer node fails, the secondary VLT peer node becomes the root bridge, avoiding problems with spanning tree port state changes that occur when a VLT node fails or recovers.
- Even with this configuration, if the node has non-VLT ports using RSTP that are not configured as edge ports and are connected to other layer 2 switches, spanning tree topology changes can still be detected after VLT node recovery. To avoid this scenario, ensure that any non-VLT ports are configured as edge ports or have RSTP disabled.
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 42-1. Configuring Interfaces for Layer 2 Mode**

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>If the interface has been assigned an IP address, remove it.</td>
<td>no ip address</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>2</td>
<td>Place the interface in Layer 2 mode.</td>
<td>switchport</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>3</td>
<td>Enable the interface.</td>
<td>no shutdown</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

Verify that an interface is in Layer 2 mode and enabled using the `show config` command from INTERFACE mode.

**Figure 42-2. Verifying Layer 2 Configuration**

```
FTOS(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
 no ip address
 switchport
 no shutdown
!
```

Indicates that the interface is in Layer 2 mode
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter the PROTOCOL SPANNING TREE RSTP mode.</td>
<td>protocol spanning-tree rstp</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Enable Rapid Spanning Tree.</td>
<td>no disable</td>
<td>PROTOCOL SPANNING TREE RSTP</td>
</tr>
</tbody>
</table>

**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 42-3. 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 42-4. 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.
Confirm that a port is participating in Rapid Spanning Tree using the `show spanning-tree rstp brief` command from EXEC privilege mode.
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 1008 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 42-2 displays the default values for RSTP.

### Table 42-2. RSTP Default Values

<table>
<thead>
<tr>
<th>RSTP Parameter</th>
<th>Default Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward Delay</td>
<td>15 seconds</td>
</tr>
<tr>
<td>Hello Time</td>
<td>2 seconds</td>
</tr>
<tr>
<td>Max Age</td>
<td>20 seconds</td>
</tr>
<tr>
<td>Port Cost</td>
<td></td>
</tr>
<tr>
<td>100-Mb/s Ethernet interfaces</td>
<td>200000</td>
</tr>
<tr>
<td>1-Gigabit Ethernet interfaces</td>
<td>20000</td>
</tr>
<tr>
<td>10-Gigabit Ethernet interfaces</td>
<td>2000</td>
</tr>
<tr>
<td>Port Channel with 100 Mb/s Ethernet interfaces</td>
<td>180000</td>
</tr>
<tr>
<td>Port Channel with 1-Gigabit Ethernet interfaces</td>
<td>18000</td>
</tr>
<tr>
<td>Port Channel with 10-Gigabit Ethernet interfaces</td>
<td>1800</td>
</tr>
<tr>
<td>Port Priority</td>
<td>128</td>
</tr>
</tbody>
</table>

To change these parameters, use the following commands, on the root bridge:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change the forward-delay parameter.</td>
<td>forward-delay seconds</td>
<td>PROTOCOL SPANNING TREE RSTP</td>
</tr>
<tr>
<td>• Range: 4 to 30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Default: 15 seconds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change the hello-time parameter.</td>
<td>hello-time seconds</td>
<td>PROTOCOL SPANNING TREE RSTP</td>
</tr>
<tr>
<td><strong>Note:</strong> With large configurations (especially those with more ports) Dell Force10 recommends that you increase the hello-time.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range: 1 to 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Default: 2 seconds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change the max-age parameter.</td>
<td>max-age seconds</td>
<td>PROTOCOL SPANNING TREE RSTP</td>
</tr>
<tr>
<td>Range: 6 to 40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Default: 20 seconds</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

View the current values for global parameters using the `show spanning-tree rstp` command from EXEC privilege mode. See Figure 42-5.

### 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 42-2. 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change the port cost of an interface.  	Range: 0 to 65535  	Default: see Table 42-2.</td>
<td><code>spanning-tree rstp cost cost</code></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>Change the port priority of an interface.  	Range: 0 to 15  	Default: 128</td>
<td><code>spanning-tree rstp priority priority-value</code></td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

View the current values for interface parameters using the `show spanning-tree rstp` command from EXEC privilege mode. See Figure 42-5.

### 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. BPDU violations 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
</table>
| Enable EdgePort on an interface.  
[`bpduguard` | `spanning-tree rstp edge-port` | INTERFACE |
| `shutdown-on-violation`] |

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 42-7.
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 42-7. 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
<td>bridge-priority priority-value</td>
<td>PROTOCOL SPANNING TREE RSTP</td>
</tr>
</tbody>
</table>

A console message appears when a new root bridge has been assigned. Figure 42-8 shows the console message after the bridge-priority command is used to make R2 the root bridge.
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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure a hello time on the order of milliseconds.</td>
<td>hello-time milli-second <code>interval</code></td>
<td>PROTOCOL RSTP</td>
</tr>
<tr>
<td></td>
<td>Range: 50 - 950 milliseconds</td>
<td></td>
</tr>
</tbody>
</table>

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.
Software-Defined Networking (SDN)

Dell Networking operating software (FTOS) supports Software-Defined Networking (SDN). For more information, refer to the *SDN Deployment Guide*. 
Security features are supported on platforms: E C S Z 54810

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).

- AAA Accounting on page 877
- AAA Authentication on page 880
- AAA Authorization on page 883
- RADIUS on page 888
- TACACS+ on page 893
- Protection from TCP Tiny and Overlapping Fragment Attacks on page 897
- SCP and SSH on page 897
- Telnet on page 903
- VTY Line and Access-Class Configuration on page 910
- Trace Lists on page 904

For details on all commands discussed in this chapter, see the Security Commands chapter in the FTOS Command Line Reference Guide.

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 Line Reference Guide.

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 applying that list to various VTY lines.

Configuration Task List for AAA Accounting

The following sections present the AAA Accounting configuration tasks:
• Enable AAA Accounting on page 878 (mandatory)
• Suppress AAA Accounting for null username sessions on page 878 (optional)
• Configure Accounting of EXEC and privilege-level command usage on page 879 (optional)
• Configure AAA Accounting for terminal lines on page 879 (optional)
• Monitor AAA Accounting on page 879 (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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>aaa accounting suppress null-username</td>
<td>CONFIGURATION</td>
<td>Prevent accounting records from being generated for users whose username string is NULL</td>
</tr>
</tbody>
</table>
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:

```
Figure 44-1. 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
FTOS#
```

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show accounting</td>
<td>CONFIGURATION</td>
<td>Step through all active sessions and print all the accounting records for the actively accounted functions.</td>
</tr>
</tbody>
</table>
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.

Note: If a console user logs in with RADIUS authentication, the privilege level will be applied from the RADIUS server if the privilege level is configured for that user in RADIUS whether or not RADIUS authorization is configured.

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 881
- Enable AAA Authentication on page 882
- AAA Authentication—RADIUS on page 882

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 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:

```
aaa authentication enable {method-list-name | default} method1 [... method4]
```

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>aaa authentication enable default radius tacacs</td>
<td>CONFIGURATION</td>
<td>To enable RADIUS and to set up TACACS as backup.</td>
</tr>
<tr>
<td>radius-server host x.x.x.x key some-password</td>
<td>CONFIGURATION</td>
<td>To establish host address and password.</td>
</tr>
<tr>
<td>tacacs-server host x.x.x.x key some-password</td>
<td>CONFIGURATION</td>
<td>To establish host address and password.</td>
</tr>
</tbody>
</table>

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:

```
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 get enable authentication from the RADIUS server, and use TACACS as a backup, issue the following commands:

```
FTOS(config)# aaa authentication enable mymethodlist radius tacacs
```

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:

```plaintext
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 884 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 884 (mandatory)
- Configure the enable password command on page 885 (mandatory)
- Configure custom privilege levels on page 885 (mandatory)
- Specify LINE mode password and privilege on page 887 (optional)
- Enable and disabling privilege levels on page 888 (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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>enable password [level level]</td>
<td>CONFIGURATION</td>
<td>Configure a password for a privilege level. Configure the optional and required parameters:</td>
</tr>
<tr>
<td>[encryption-mode] password</td>
<td></td>
<td>• <code>level level</code>: Specify a level 0 to 15. Level 15 includes all levels.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>encryption-mode</code>: Enter 0 for plain text or 7 for encrypted text.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>password</code>: Enter a string.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>To change only the password for the enable command, configure only the <code>password</code> parameter.</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 1    | username name [access-class access-list-name] [privilege level] [nopassword | 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-mode: 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 44-2 is an example of a configuration to allow a user “john” to view only the EXEC mode commands and all snmp-server 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.
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 assigned to the user.
To specify a password for the terminal line, use the following commands, in any order, in the LINE mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>privilege level level</td>
<td>LINE</td>
<td>Configure a custom privilege level for the terminal lines.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- level level range: 0 to 15. Levels 0, 1 and 15 are pre-configured. Levels 2 to 14 are available for custom configuration.</td>
</tr>
<tr>
<td>password [encryption-type] password</td>
<td>LINE</td>
<td>Specify either a plain text or encrypted password.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Configure the following optional and required parameters:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- encryption-type: Enter 0 for plain text or 7 for encrypted text.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- password: Enter a text string up to 25 characters long.</td>
</tr>
</tbody>
</table>

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.

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 890 (mandatory)
- Apply the method list to terminal lines on page 891 (mandatory except when using default lists)
- Specify a RADIUS server host on page 891 (mandatory)
- Set global communication parameters for all RADIUS server hosts on page 892 (optional)
- Monitor RADIUS on page 893 (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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>aaa authentication login</td>
<td>CONFIGURATION</td>
<td>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.</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>aaa authorization exec</td>
<td>CONFIGURATION</td>
<td>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).</td>
</tr>
<tr>
<td>{method-list-name | default} radius tacacs+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| radius-server host {hostname \| ip-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 892.

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>radius-server deadtime seconds</code></td>
<td>CONFIGURATION</td>
<td>Set a time interval after which a RADIUS host server is declared dead.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>seconds</code> range: 0 to 2147483647. Default: 0 seconds</td>
</tr>
<tr>
<td><code>radius-server key [encryption-type] key</code></td>
<td>CONFIGURATION</td>
<td>Configure a key for all RADIUS communications between the system and RADIUS server hosts.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>encryption-type</code>: Enter 7 to encrypt the password. Enter 0 to keep the password as plain text.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>key</code>: Enter a string. The key can be up to 42 characters long. You cannot use spaces in the key.</td>
</tr>
<tr>
<td><code>radius-server retransmit retries</code></td>
<td>CONFIGURATION</td>
<td>Configure the number of times FTOS retransmits RADIUS requests.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>retries</code> range: 0 to 100. Default is 3 retries.</td>
</tr>
<tr>
<td><code>radius-server timeout seconds</code></td>
<td>CONFIGURATION</td>
<td>Configure the time interval the system waits for a RADIUS server host response.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>seconds</code> range: 0 to 1000. Default is 5 seconds.</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>debug radius</code></td>
<td>EXEC Privilege</td>
<td>View RADIUS transactions to troubleshoot problems.</td>
</tr>
</tbody>
</table>

**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 895
- TACACS+ Remote Authentication and Authorization on page 895
- Specify a TACACS+ server host on page 896
- Choose TACACS+ as the Authentication Method on page 893

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>tacacs-server host {ip-address</td>
<td>host}</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>aaa authentication login {method-list-name</td>
<td>default} tacacs+ [...method3]</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>3</td>
<td>line {aux 0</td>
<td>console 0</td>
<td>vty number [end-number]}</td>
</tr>
<tr>
<td>4</td>
<td>login authentication {method-list-name</td>
<td>default}</td>
<td>LINE</td>
</tr>
</tbody>
</table>

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 44-4, the TACACS+ is incorrect, but the user is still authenticated by the secondary method.
Figure 44-4. 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
authenticaiton 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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>debug tacacs+</td>
<td>EXEC Privilege</td>
<td>View TACACS+ transactions to troubleshoot problems.</td>
</tr>
</tbody>
</table>

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 44-5 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 44-5. Specify a TACACS+ server host

When configuring a TACACS+ server host, you can set different communication parameters, such as the key password.

To specify a TACACS+ server host and configure its communication parameters, use the following command in the CONFIGURATION mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| tacacs-server host {hostname | CONFIGURATION Enter the host name or IP address of the TACACS+ \-ip-address} [port port-number] server host. Configure the optional communication \[timeout seconds] [key key] 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::2202
Trying 2200:2200:2200:2200::2202...
Connected to 2200:2200:2200:2200::2202.
Escape character is '^]'.
Login: admin
Password:
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. 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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| `ssh {hostname} [-l username | -p port-number | -v {1 | 2}` | EXEC Privilege | Open an SSH connection specifying the hostname, username, port number, and version of the SSH client. `hostname` is the IP address or hostname of the remote device.  
  • Enter an IPv4 or IPv6 address in dotted decimal format (A.B.C.D). |

To enable the SSH server for version 1 and 2, use the following command in the CONFIGURATION mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>`ip ssh server {enable</td>
<td>port port-number}`</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

To enable the SSH server for version 1 or 2 only, use the following command:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>`ip ssh server version {1</td>
<td>2}`</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

To view the SSH configuration, use the following command in EXEC Privilege mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>show ip ssh</code></td>
<td>EXEC Privilege</td>
<td>Display SSH connection information.</td>
</tr>
</tbody>
</table>

Figure 44-6 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.
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>On Chassis One, set the SSH port number (port 22 by default).</td>
<td><code>ip ssh server port number</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>On Chassis One, enable SSH.</td>
<td><code>ip ssh server enable</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>3</td>
<td>On Chassis Two, invoke SCP.</td>
<td><code>copy scp: flash:</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>4</td>
<td>On Chassis Two, in response to prompts, enter the path to the desired file and enter the port number specified in Step 1.</td>
<td></td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

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:

```
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.
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 900**
- **RSA Authentication of SSH on page 901**
- **Host-based SSH Authentication on page 901**

**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.
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>On the SSH client (Unix machine), generate an RSA key, as shown in Figure 44-9.</td>
<td><code>ssh-keygen -t rsa</code></td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td></td>
<td>Figure 44-9. Generating RSA Keys</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>admin@Unix_client#ssh-keygen -t rsa</td>
<td>Generating public/private rsa key pair.</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td></td>
<td>Enter file in which to save the key (/home/admin/.ssh/id_rsa):</td>
<td>/home/admin/.ssh/id_rsa already exists.</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td></td>
<td>Overwrite (y/n)?</td>
<td>y</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td></td>
<td>Enter passphrase (empty for no passphrase):</td>
<td>Enter same passphrase again:</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td></td>
<td>Your identification has been saved in /home/admin/.ssh/id_rsa.</td>
<td>Your public key has been saved in /home/admin/.ssh/id_rsa.pub.</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>2</td>
<td>Copy the public key <code>id_rsa.pub</code> to the Dell Force10 system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Disable password authentication if enabled.</td>
<td><code>no ip ssh password-authentication enable</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>4</td>
<td>Enable RSA authentication.</td>
<td><code>ip ssh rsa-authentication enable</code></td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>5</td>
<td>Bind the public keys to RSA authentication.</td>
<td><code>ip ssh rsa-authentication my-authorized-keys flash://public_key</code></td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

Host-based SSH Authentication

Authenticate a particular host. This method uses SSH version 2.

To configure host-based authentication:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure RSA Authentication. See RSA Authentication of SSH, above.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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`. 

### Table: Command Syntax and Command Mode

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Create hosts by copying the public RSA key to the the file <code>hosts</code> in the directory <code>.ssh</code>, and write the IP address of the host to the file.</td>
<td><code>cp /etc/ssh/ssh_host_rsa_key.pub /.ssh/hosts</code></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Create a list of IP addresses and usernames that are permitted to SSH in a file called <code>rhosts</code>, as shown in Figure 44-11.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Copy the file hosts and rhosts to the Dell Force10 system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Disable password authentication and RSA authentication, if configured</td>
<td><code>no ip ssh password-authentication</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>no ip ssh rsa-authentication</code></td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>6</td>
<td>Enable host-based authentication.</td>
<td><code>ip ssh hostbased-authentication enable</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>7</td>
<td>Bind hosts and rhosts to host-based authentication.</td>
<td><code>ip ssh pub-key-file flash://filename</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>ip ssh rhostsfile flash://filename</code></td>
<td></td>
</tr>
</tbody>
</table>
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:

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.

Configuration Tasks for Trace Lists

The following configuration steps include mandatory and optional steps.

- Creating a trace list (mandatory)
- Apply trace lists (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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ip trace-list trace-list-name</td>
<td>CONFIGURATION</td>
<td>Enter the TRACE LIST mode by creating an trace list.</td>
</tr>
</tbody>
</table>
| 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ip trace-list trace-list-name</td>
<td>CONFIGURATION</td>
<td>Create a trace list and assign it a unique name.</td>
</tr>
</tbody>
</table>
To create a filter for UDP packets with a specified sequence number, use these commands in the following sequence, starting in the CONFIGURATION mode:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ip trace-list access-list-name</td>
<td>CONFIGURATION</td>
<td>Create a trace list and assign it a unique name.</td>
</tr>
<tr>
<td>2</td>
<td>seq sequence-number {deny</td>
<td>permit} udp</td>
<td>TRACE LIST</td>
</tr>
<tr>
<td></td>
<td>{source mask</td>
<td>any</td>
<td>host ip-address}</td>
</tr>
<tr>
<td></td>
<td>[operator port [port]] {destination mask</td>
<td></td>
<td>• mask: a network mask</td>
</tr>
<tr>
<td></td>
<td>any</td>
<td>host ip-address} [operator port [port]]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[established] [count [byte]</td>
<td>log]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[port]] [count [byte]</td>
<td>log]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• count: count packets processed by the filter.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• byte: count bytes processed by the filter.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• log: is supported.</td>
</tr>
</tbody>
</table>

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 44-13 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.
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| {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. |

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
FTOS(config-trace-acl)#
Figure 44-14 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.
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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ip trace-group trace-list-name</code></td>
<td>CONFIGURATION</td>
<td>Enable a configured Trace list to filter traffic.</td>
</tr>
</tbody>
</table>

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 44-15. 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 44-1. VTY Access

<table>
<thead>
<tr>
<th>Authentication Method</th>
<th>VTY access-class support?</th>
<th>Username access-class support?</th>
<th>Remote authorization support?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Local</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>TACACS+</td>
<td>YES</td>
<td>NO</td>
<td>YES (with FTOS 5.2.1.0 and later)</td>
</tr>
<tr>
<td>RADIUS</td>
<td>YES</td>
<td>NO</td>
<td>YES (with FTOS 6.1.1.0 and later)</td>
</tr>
</tbody>
</table>

FTOS provides several ways to configure access classes for VTY lines, including:

- VTY Line Local Authentication and Authorization on page 910
- VTY Line Remote Authentication and Authorization on page 911

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).

Note: If a VTY user logs in with RADIUS authentication, the privilege level will be applied from the RADIUS server only if RADIUS authentication is configured.
Figure 44-16 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 44-16. 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 deny all
FTOS(conf)#aaa authentication login localmethod local
FTOS(conf)#line vty 0 9
FTOS(config-line-vty)#login authentication localmethod
FTOS(config-line-vty)#end
```

*Note: See also the section Chapter 7, Access Control Lists (ACLs).*

**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. The following example 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 44-17. 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)#aaa authentication login tacacsmethod tacacs+
FTOS(conf)#tacacs-server host 256.1.1.2 key Force10
FTOS(conf)#line vty 0 9
FTOS(config-line-vty)#login authentication tacacsmethod
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 44-18). Figure 44-18 shows how to deny incoming connections from subnet 10.0.0.0 without displaying a login prompt.

**Figure 44-18. Example Access Class Configuration Using TACACS+ Without Prompt**

```plaintext
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(config)#line vty 0 9
FTOS(config-line-vty)#access-class sourcemac
FTOS(config-line-vty)#end
```
Service Provider Bridging

Service Provider Bridging is supported on platforms: $E\ C\ S\ Z\ S4810$

This chapter contains the following major sections:

- VLAN Stacking
- VLAN Stacking Packet Drop Precedence
- Dynamic Mode CoS for VLAN Stacking
- Layer 2 Protocol Tunneling
- Provider Backbone Bridging

VLAN Stacking

VLAN Stacking is supported on platforms: $C\ E\ S\ Z$

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 45-1).
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 915.
2. Assign access and trunk ports to a VLAN. See page 915.
3. Make the VLAN VLAN-stacking capable.

Related Configuration Tasks

- Configure the Protocol Type Value for the Outer VLAN Tag on page 916
- FTOS Options for Trunk Ports on page 916
- Debug VLAN Stacking on page 917
- VLAN Stacking in Multi-vendor Networks on page 918
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.

To create access and trunk ports:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Assign the role of access port to a Layer 2 port on a provider bridge that is connected to a customer.</td>
<td>vlan-stack access</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>2</td>
<td>Assign the role of trunk port to a Layer 2 port on a provider bridge that is connected to another provider bridge.</td>
<td>vlan-stack trunk</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>3</td>
<td>Assign all access ports and trunk ports to service provider VLANs.</td>
<td>member</td>
<td>INTERFACE VLAN</td>
</tr>
</tbody>
</table>

Display the VLAN-Stacking configuration for a switchport using the command show config from INTERFACE mode, as shown in Figure 45-2.

**Figure 45-2. 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable VLAN-Stacking for the VLAN.</td>
<td>INTERFACE VLAN</td>
<td>vlan-stack compatible</td>
</tr>
</tbody>
</table>
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 45-3. Display the Members of a VLAN-Stacking-enabled VLAN**

```
FTOS#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

<table>
<thead>
<tr>
<th>NUM</th>
<th>Status</th>
<th>Q Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Active</td>
<td>Gi 13/0-5,18</td>
</tr>
<tr>
<td>2</td>
<td>Inactive</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Inactive</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Inactive</td>
<td>Gi 13/14-15</td>
</tr>
<tr>
<td>5</td>
<td>Inactive</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Active</td>
<td>Gi 13/13</td>
</tr>
</tbody>
</table>

| FTOS# |
```

**Configure the Protocol Type Value for the Outer VLAN Tag**

The Tag Protocol Identifier (TPID) field of the S-Tag is user-configurable:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select a value for the S-Tag TPID.</td>
<td><code>vlan-stack protocol-type</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Default: 9100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure a trunk port to carry untagged, single-tagged, and double-tagged traffic by making it a hybrid port.</td>
<td><code>portmode hybrid</code></td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

**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.
In Figure 45-4 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 45-4. Hybrid Port as VLAN-Stack Trunk Port and as Member of other VLANs

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Add the port to a 802.1Q VLAN as tagged or untagged.</td>
<td>[tagged</td>
<td>untagged] INTERFACE VLAN</td>
</tr>
</tbody>
</table>

Debug VLAN Stacking

To debug the internal state and membership of a VLAN and its ports, use the debug member command, as shown in Figure 45-5. The port notations in Figure 45-5 are as follows:

- **MT** — stacked trunk
- **MU** — stacked access port
- **T** — 802.1Q trunk port
- **U** — 802.1Q access port
- **NU** — Native VLAN (untagged)
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 45-6) 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 45-6, 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 45-6, the frame originating from Building C is tagged 0x9191 on egress of R1. R2’s TPID is 0x9100, but it is an E-Series TeraScale system and makes no distinction between 0x9191 and 0x9100, so it forwards the frame.
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 45-7, 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.
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 45-6, 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 45-8 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.
Table 45-1 details the outcome of matched and mis-matched TPIDs in a VLAN-stacking network with the E-Series.

### Table 45-1. E-Series Behaviors for Mis-matched TPID

<table>
<thead>
<tr>
<th>Network Position</th>
<th>Incoming Packet TPID</th>
<th>System TPID</th>
<th>Match Type</th>
<th>TeraScale Behavior</th>
<th>ExaScale Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td>0xUVWX</td>
<td>0xUVYZ</td>
<td>1st-byte match</td>
<td>switch as 0xUVYZ</td>
<td>drop</td>
</tr>
<tr>
<td></td>
<td>0xUVWZ</td>
<td>0xQRST</td>
<td>mismatch</td>
<td>drop</td>
<td>drop</td>
</tr>
<tr>
<td>Egress Access Point</td>
<td>0xUVWX</td>
<td>0xUVYZ</td>
<td>1st-byte match</td>
<td>switch as 0xUVYZ</td>
<td>drop</td>
</tr>
<tr>
<td></td>
<td>0x81WX</td>
<td>0x81YZ</td>
<td>1st-byte match</td>
<td>switch as is (no decapsulation)</td>
<td>drop</td>
</tr>
<tr>
<td></td>
<td>0xUVWZ</td>
<td>0xQRST</td>
<td>mismatch</td>
<td>drop</td>
<td>drop</td>
</tr>
</tbody>
</table>

**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 45-9. Versions 8.2.1.0 and later differentiate between 0x9100 and 0x91XY, as shown in Figure 45-11.
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 45-11. 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 45-11.

Therefore, a mismatched TPID results in the port not differentiating between tagged and untagged traffic.

Figure 45-9. Single and Double-tag TPID Match on the C-Series and S-Series
Figure 45-10. Single and Double-tag First-byte TPID Match on C-Series and S-Series

Figure 45-11. Single and Double-tag TPID Mismatch on the C-Series and S-Series
Table 45-2 details the outcome of matched and mismatched TPIDs in a VLAN-stacking network with the C-Series and S-Series.

### Table 45-2. C-Series and S-Series Behaviors for Mis-matched TPID

<table>
<thead>
<tr>
<th>Network Position</th>
<th>Incoming Packet TPID</th>
<th>System TPID</th>
<th>Match Type</th>
<th>Pre-8.2.1.0</th>
<th>8.2.1.0+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ingress Access Point</td>
<td>untagged</td>
<td>0xUVWX</td>
<td>—</td>
<td>switch to default VLAN</td>
<td>switch to default VLAN</td>
</tr>
<tr>
<td>single-tag (0x8100)</td>
<td>0xUVWX</td>
<td>single-tag mismatch</td>
<td>switch to default VLAN</td>
<td>switch to default VLAN</td>
<td></td>
</tr>
<tr>
<td>Core</td>
<td>untagged</td>
<td>0xUVWX</td>
<td>—</td>
<td>switch to default VLAN</td>
<td>switch to default VLAN</td>
</tr>
<tr>
<td>double-tag 0xUVWX</td>
<td>0xUVWX</td>
<td>double-tag match</td>
<td>switch to VLAN</td>
<td>switch to VLAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0xUVYZ</td>
<td>double-tag first-byte match</td>
<td>switch to VLAN</td>
<td>switch to default VLAN</td>
<td></td>
</tr>
<tr>
<td>Egress Access Point</td>
<td>untagged</td>
<td>0xUVWX</td>
<td>—</td>
<td>switch to default VLAN</td>
<td>switch to default VLAN</td>
</tr>
<tr>
<td>double-tag 0xUVWX</td>
<td>0xUVWX</td>
<td>double-tag match</td>
<td>switch to VLAN</td>
<td>switch to VLAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0xUVYZ</td>
<td>double-tag first-byte match</td>
<td>switch to VLAN</td>
<td>switch to default VLAN</td>
<td></td>
</tr>
</tbody>
</table>

### 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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Make packets eligible for dropping based on their DEI value. By default, packets are colored green, and DEI is marked 0 on egress.</td>
<td>dei enable</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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 45-3.

Table 45-3. Drop Eligibility Behavior

<table>
<thead>
<tr>
<th>Ingress Egress</th>
<th>DEI Disabled</th>
<th>DEI Enabled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Port Normal Port</td>
<td>Retain CFI</td>
<td>Set CFI to 0</td>
</tr>
<tr>
<td>Trunk Port Trunk Port</td>
<td>Retain inner tag CFI</td>
<td>Retain inner tag CFI</td>
</tr>
<tr>
<td></td>
<td>Retain outer tag CFI</td>
<td>Set outer tag CFI to 0</td>
</tr>
<tr>
<td>Access Port Trunk Port</td>
<td>Retain inner tag CFI</td>
<td>Retain inner tag CFI</td>
</tr>
<tr>
<td></td>
<td>Set outer tag CFI to 0</td>
<td>Set outer tag CFI to 0</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Precedence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>High priority packets that are the least preferred to be dropped.</td>
</tr>
<tr>
<td>Yellow</td>
<td>Lower priority packets that are treated as best-effort.</td>
</tr>
<tr>
<td>Red</td>
<td>Lowest priority packets that are always dropped (regardless of congestion status).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
<td>dei honor {0</td>
<td>1} {green</td>
</tr>
<tr>
<td>Display the DEI-honoring configuration.</td>
<td>show interface dei-honor [interface slot/port</td>
<td>linecard number port-set number]</td>
</tr>
</tbody>
</table>
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).

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.
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Allocate CAM space to enable queuing frames according to the C-Tag or the S-Tag.</td>
<td>cam-acl l2acl number ipv4acl number ipv6acl number ipv4qos number l2pt number ipmacacl number ecfmacl number {vman-qos</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td></td>
<td>vman-qos: mark the S-Tag dot1p and queue the frame according to the original C-Tag dot1p.</td>
<td>vman-qos-dual-fp} number Default: 0 FP blocks for vman-qos and vman-qos-dual-fp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>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.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>The new CAM configuration is stored in NVRAM and takes effect only after a save and reload.</td>
<td>copy running-config startup-config reload</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>3</td>
<td>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.</td>
<td>vlan-stack dot1p-mapping c-tag-dot1p values sp-tag-dot1p value</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

**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

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 45-13).
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 45-14).

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 Dell 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.
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

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Verify that the system is running the default CAM profile; you must use this CAM profile for L2PT.</td>
<td>show cam-profile</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>2</td>
<td>Enable protocol tunneling globally on the system.</td>
<td>protocol-tunnel enable</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>3</td>
<td>Tunnel BPDUs the VLAN.</td>
<td>protocol-tunnel stp</td>
<td>INTERFACE VLAN</td>
</tr>
</tbody>
</table>
Specify a Destination MAC Address for BPDUs

By default, FTOS uses a Dell Force10-unique MAC address for tunneling BPDUs. You can configure another value.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overwrite the BPU with a user-specified destination MAC address when BPDUs are tunneled across the provider network. Default: 01:01:e8:00:00:00</td>
<td>protocol-tunnel destination-mac</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set a maximum rate at which the RPM will process BPDUs for L2PT. Default: 75 pps E-Series Range: 75 to 3000 pps</td>
<td>protocol-tunnel rate-limit</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create at least one FP group for L2PT. See CAM Allocation on page 278 for details on this command.</td>
<td>cam-acl l2acl</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Save the running-config to the startup-config.</td>
<td>copy running-config startup-config</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>3</td>
<td>Reload the system.</td>
<td>reload</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>4</td>
<td>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</td>
<td>protocol-tunnel rate-limit</td>
<td>VLAN STACKING</td>
</tr>
</tbody>
</table>
Debug Layer 2 Protocol Tunneling

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display debugging information for L2PT.</td>
<td>debug protocol-tunnel</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

**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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>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+.</td>
<td>bpdu-destination-mac-address [stp</td>
<td>gvrp] provider-bridge-group</td>
</tr>
</tbody>
</table>
sFlow

Configuring sFlow is supported on platforms: E C S Z S4810

- Enable and Disable sFlow
- sFlow Show Commands
- Specify Collectors
- Polling Intervals
- Sampling Rate
- Back-off Mechanism
- sFlow on LAG ports
- Extended sFlow

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.
Implementation Information

The 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 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 then 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 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>[no] sflow enable</td>
<td>CONFIGURATION</td>
<td>Enable sFlow globally.</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>[no] sflow enable</td>
<td>INTERFACE</td>
<td>Enable sFlow on an interface.</td>
</tr>
</tbody>
</table>
sFlow Show Commands

FTOS includes the following sFlow display commands:

- Show sFlow Globally
- Show sFlow on an Interface
- Show sFlow on a Line Card

Show sFlow Globally

Use the following command to view sFlow statistics:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show sflow</td>
<td>EXEC</td>
<td>Display sFlow configuration information and statistics.</td>
</tr>
</tbody>
</table>

Figure 46-2 is a sample output from the show sflow command:

```
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
  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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show sflow interface</td>
<td>EXEC</td>
<td>Display sFlow configuration information and statistics on a specific interface.</td>
</tr>
<tr>
<td>interface-name</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 46-3 is a sample output from the show sflow interface command.
The configuration, shown in Figure 46-2, is also displayed in the running configuration (Figure 46-4):

**Figure 46-4. 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 statistics on a specified line card:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>show sflow linecard slot-number</td>
<td>EXEC</td>
<td>Display sFlow configuration information and statistics on the specified interface.</td>
</tr>
</tbody>
</table>

**Figure 46-5** is a sample output from the show sflow linecard command:

**Figure 46-5. 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 :??
  UDP packets exported via RPM :??
  UDP packets dropped :?
```
Specify Collectors

The `sflow collector` command allows identification of sFlow Collectors to which sFlow datagrams are forwarded. The user can specify up to two sFlow collectors. If two Collectors are specified, the samples are sent to both.

Collection through Management interface is supported on platform: E.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>sflow collector ip-address</code></td>
<td>CONFIGURATION</td>
<td>Identify sFlow collectors to which sFlow datagrams are forwarded.</td>
</tr>
<tr>
<td><code>agent-addr ip-address [number</code></td>
<td></td>
<td>Default UDP port: 6343</td>
</tr>
<tr>
<td><code>max-datagram-size number ] ]</code></td>
<td></td>
<td>Default max-datagram-size: 1400</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>sflow polling-interval interval value</code></td>
<td>CONFIGURATION or INTERFACE</td>
<td>Change the global default counter polling interval.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>interval value</code>—in seconds.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range: 15 to 86400 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default: 20 seconds</td>
</tr>
</tbody>
</table>

Sampling Rate

**Sampling Rate** is supported on platform E.

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 numbers and re-enter the command. (For more information on values in power-of-2, see Sub-sampling on page 941.)

The sample rate can be configured globally or by interface using the sample rate command:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Usage</th>
</tr>
</thead>
</table>
| `[no] sflow sample-rate`  | CONFIGURATION or INTERFACE | Change the global or interface sampling rate. Rate must be entered in factors of 2 (e.g., 4096, 8192).
| sample-rate               |                     | `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:

\[
\text{Actual sampling rate} = \text{sub-sampling rate} \times \text{hardware sampling rate}
\]

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 46-6.
If none of the extended information is enabled, the show output is as shown in Figure 46-7.

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 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 46-1. Extended Gateway Summary

<table>
<thead>
<tr>
<th>IP SA</th>
<th>IP DA</th>
<th>srcAS and srcPeerAS</th>
<th>dstAS and dstPeerAS</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>static/connected/IGP</td>
<td>static/connected/IGP</td>
<td>—</td>
<td>—</td>
<td>Extended gateway data is not exported because there is no AS information.</td>
</tr>
<tr>
<td>static/connected/IGP</td>
<td>BGP</td>
<td>0</td>
<td>Exported</td>
<td>src_as &amp; src_peer_as are zero because there is no AS information for IGP.</td>
</tr>
<tr>
<td>BGP</td>
<td>static/connected/IGP</td>
<td>—</td>
<td>—</td>
<td>Prior to FTOS version 7.8.1.0, extended gateway data is not be exported because IP DA is not learned via BGP. 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.</td>
</tr>
<tr>
<td>BGP</td>
<td>BGP</td>
<td>Exported</td>
<td>Exported</td>
<td>Extended gateway data is packed.</td>
</tr>
</tbody>
</table>
Simple Network Management Protocol (SNMP)

Simple Network Management Protocol (SNMP) is supported on platforms: E C S Z S4810

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 16 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 version 1 or version 2 requires only a single step:
1. Create a community. See page 947.

Configuring SNMP version 3 requires you to configure SNMP users in one of three methods. See Setting Up User-based Security (SNMPv3).

**Related Configuration Tasks**

The following list contains configuration tasks for SNMP:

- Managing Overload on Startup
- Read Managed Object Values
- Write Managed Object Values
- Subscribe to Managed Object Value Updates using SNMP
- Copy Configuration Files Using SNMP
- Manage VLANs using SNMP
- Enable and Disable a Port using SNMP
- Fetch Dynamic MAC Entries using SNMP
- Deriving Interface Indices
- Monitor Port-channels

**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.
- User ACLs override group ACLs.

**Setting up SNMP**

As previously stated, FTOS supports SNMP version 1 and version 2 which are community-based security models. The primary difference between the two versions is that version 2 supports two additional protocol operations (informs operation and `snmpgetbulk` query) and one additional object (counter64 object).

SNMP version 3 (SNMPv3) is a user-based security model that provides password authentication for user security and encryption for data security and privacy. Three sets of configurations are available for SNMP read/write operations: no password or privacy, password privileges, password and privacy privileges.

A maximum of 16 users can be configured even if they are in different groups.
Create a Community

For SNMPv1 and SNMPv2, you must create a community to enable the community-based security in FTOS. The management station generates requests to either retrieve or alter the value of a management object and is called the SNMP manager. 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose a name for the community.</td>
<td>`snmp-server community name {ro</td>
<td>rw}`</td>
</tr>
</tbody>
</table>

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 47-1.

Figure 47-1. Creating an SNMP Community

```
FTOS(config)#snmp-server community my-snmp-community ro
22:31:23: %RPM1-P:CP %SNMP-6-SNMP_WARM_START: Agent Initialized - SNMP WARM_START.
FTOS#show running-config snmp
!
snmp-server community mycommunity ro
```

Setting Up User-based Security (SNMPv3)

When setting up SNMPv3, you can set users up with one of the following three types of configuration for SNMP read/write operations. Users are typically associated to an SNMP group with permissions provided, such as OID view.

- **noauth**: no password or privacy. Select this option to set a user up with no password or privacy privileges. This is the basic configuration. Users must have a group and profile that do not require password privileges.
- **auth**: password privileges. Select this option to set up an user with password authentication
- **priv**: password and privacy privileges. Select this option to set up a user with password and privacy privileges.
To set up a user with view privileges only (no password or privacy privileges):

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure the user.</td>
<td><code>snmp-server user name group-name 3 noauth</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Configure an SNMP group.</td>
<td><code>snmp-server group group-name 3 noauth auth read name write name</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Configure an SNMPv3 view.</td>
<td>`snmp-server view view-name oid-tree {included</td>
<td>excluded}`</td>
</tr>
</tbody>
</table>

**Note:** To give a user read and write view privileges, repeat this step for each privilege type.

To set up a user with password privileges only, use the following commands:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure the user with an authorization password</td>
<td><code>snmp-server user name group-name 3 noauth auth md5 auth-password</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Configure an SNMP group.</td>
<td><code>snmp-server group groupname (oid-tree) auth read name write name</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Configure an SNMPv3 view.</td>
<td>`snmp-server view view-name 3 noauth {included</td>
<td>excluded}`</td>
</tr>
</tbody>
</table>

**Note:** To give a user read and write privileges, repeat this step for each privilege type.

To set up a user with password or privacy privileges:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure an SNMP group.</td>
<td><code>snmp-server group group-name (oid-tree) priv read name write name</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Configure the user with a secure authorization password and privacy password.</td>
<td><code>snmp-server user name group-name (oid-tree) auth md5 auth-password priv des56 priv password</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Configure an SNMPv3 view.</td>
<td>`snmp-server view view-name oid-tree {included</td>
<td>excluded}`</td>
</tr>
</tbody>
</table>
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.

Dell Force10 supports RFC 4001, *Textual Conventions for Internet Work Addresses* that defines values representing a type of internet address. These values display for ipAddressTable objects using the `snmpwalk` command.

In the following figure, the value “4” displays in the OID before the IP address for IPv4. For an IPv6 IP address, a value of “16” displays.

```
> snmpwalk -v 2c -c public 10.11.195.63 1.3.6.1.2.1.4.34
IP-MIB::ip.34.1.3.1.4.1.1.1.1 = INTEGER: 1107787778
IP-MIB::ip.34.1.3.1.4.2.1.1.1 = INTEGER: 1107787779
IP-MIB::ip.34.1.3.2.16.0.1.0.0.0.0.0.0.0.0.0.0.0.0.0.0.1 = INTEGER: 1107787779
IP-MIB::ip.34.1.3.2.16.0.2.0.0.0.0.0.0.0.0.0.0.0.0.0.0.1 = INTEGER: 1107787779
IP-MIB::ip.34.1.3.2.16.254.0.0.0.0.0.0.0.2.1.232.255.254.139.5.8 = INTEGER: 1107787778
IP-MIB::ip.34.1.4.1.4.1.1.1.1 = INTEGER: 1
IP-MIB::ip.34.1.4.1.4.2.1.1.1 = INTEGER: 1
IP-MIB::ip.34.1.4.2.16.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.1 = INTEGER: 1
```

There are several UNIX SNMP commands that read data:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read the value of a single managed object, as shown in Figure 47-3.</td>
<td>`snmpget -v version -c community agent-ip {identifier.instance</td>
</tr>
<tr>
<td>Figure 47-3. Reading the Value of a Managed Object</td>
<td></td>
</tr>
<tr>
<td>&gt; snmpget -v 2c -c mycommunity 10.11.131.161 sysUpTime.0</td>
<td>DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (32852616) 3 days, 19:15:26.16</td>
</tr>
<tr>
<td>&gt; snmpget -v 2c -c mycommunity 10.11.131.161 .1.3.6.1.2.1.1.3.0</td>
<td></td>
</tr>
<tr>
<td>Read the value of the managed object directly below the specified object, as shown in Figure 47-4.</td>
<td>`snmpgetnext -v version -c community agent-ip {identifier.instance</td>
</tr>
<tr>
<td>Figure 47-4. Reading the Value of the Next Managed Object in the MIB</td>
<td></td>
</tr>
<tr>
<td>&gt; snmpgetnext -v 2c -c mycommunity 10.11.131.161 .1.3.6.1.2.1.1.3.0</td>
<td>SNMPv2-MIB::sysContact.0 = STRING:</td>
</tr>
<tr>
<td>&gt; snmpgetnext -v 2c -c mycommunity 10.11.131.161 sysContact.0</td>
<td></td>
</tr>
<tr>
<td>Read the value of many objects at once, as shown in Figure 47-5.</td>
<td>`snmpwalk -v version -c community agent-ip {identifier.instance</td>
</tr>
</tbody>
</table>
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 47-6.  
  snmpset -v version -c community agent-ip {identifier.instance | descriptor.instance} syntax value
```

Figure 47-6. 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
```
To configure the system from the management station using SNMP:

### Task
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

**Command:**
```
snmpset -v version -c community agent-ip sysContact.0 s "contact-info"
```

### Task
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

**Command:**
```
snmpset -v version -c community agent-ip sysLocation.0 s "location-info"
```

## 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, egpNeighborLoss
- **Force10 enterpriseSpecific environment traps:** fan, supply, temperature
- **Force10 enterpriseSpecific protocol traps:** bgp, ecfm, stp, xstp

To configure the system to send SNMP notifications:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure the Dell Force10 system to send notifications to an SNMP server.</td>
<td>`snmp-server host ip-address [traps</td>
<td>informs] [version 1</td>
</tr>
</tbody>
</table>
Table 47-1 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 47-1. RFC 1157 Defined SNMP Traps on FTOS

<table>
<thead>
<tr>
<th>Command Option</th>
<th>Trap</th>
</tr>
</thead>
<tbody>
<tr>
<td>snmp authentication</td>
<td>SNMP_AUTH_FAIL: SNMP Authentication failed. Request with invalid community string.</td>
</tr>
<tr>
<td>snmp coldstart</td>
<td>SNMP_COLD_START: Agent Initialized - SNMP COLD_START. SNMP_WARM_START: Agent Initialized - SNMP WARM_START.</td>
</tr>
<tr>
<td>snmp linkdown</td>
<td>PORT_LINKDN: changed interface state to down:%d</td>
</tr>
<tr>
<td>snmp lookup</td>
<td>PORT_LINKUP: changed interface state to up:%d</td>
</tr>
</tbody>
</table>

Enable a subset of Dell Force10 enterprise specific SNMP traps using one of the listed command options in Table 47-2 with the command `snmp-server enable traps`. Note that the `envmon` option enables all environment traps including those that are enabled with the `envmon supply`, `envmon temperature`, and `envmon fan` options.

Table 47-2. Dell Force10 Enterprise-specific SNMP Traps

<table>
<thead>
<tr>
<th>Command Option</th>
<th>Trap</th>
</tr>
</thead>
<tbody>
<tr>
<td>envmon</td>
<td>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</td>
</tr>
</tbody>
</table>
### Table 47-2. Dell Force10 Enterprise-specific SNMP Traps

<table>
<thead>
<tr>
<th>Command Option</th>
<th>Trap</th>
</tr>
</thead>
<tbody>
<tr>
<td>MINOR_SFM: MInor alarm: No working standby SFM</td>
<td></td>
</tr>
<tr>
<td>MINOR_SFM_CLR: Minor alarm cleared: Working standby SFM present</td>
<td></td>
</tr>
<tr>
<td>TASK SUSPENDED: SUSPENDED - svce:%d - inst:%d - task:%s</td>
<td></td>
</tr>
<tr>
<td>RPM0-P:CP %CHMGR-2-CARD_PARITY_ERR</td>
<td></td>
</tr>
<tr>
<td>ABNORMAL_TASK_TERMINATION: CRASH - task:%s %s</td>
<td></td>
</tr>
<tr>
<td>CPU_THRESHOLD: Cpu %s usage above threshold. Cpu5SecUsage (%d)</td>
<td></td>
</tr>
<tr>
<td>CPU_THRESHOLD_CLR: Cpu %s usage drops below threshold. Cpu5SecUsage (%d)</td>
<td></td>
</tr>
<tr>
<td>MEM_THRESHOLD: Memory %s usage above threshold. MemUsage (%d)</td>
<td></td>
</tr>
<tr>
<td>MEM_THRESHOLD_CLR: Memory %s usage drops below threshold. MemUsage (%d)</td>
<td></td>
</tr>
<tr>
<td>DETECT_STN_MOVE: Station Move threshold exceeded for Mac %s in vlan %d</td>
<td></td>
</tr>
<tr>
<td>CAM-UTILIZATION: Enable SNMP envmon CAM utilization traps.</td>
<td></td>
</tr>
</tbody>
</table>
envmon supply
- FEM_FRLBLM: Major alarm: problem with power entry module %s
- FEM_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 °C is within threshold of %d°C)
- MAJOR_TEMP: Major alarm: chassis temperature high (%s °C temperature reaches or exceeds threshold of %d°C)
- MAJOR_TEMP_CLR: Major alarm cleared: chassis temperature lower (%s °C °C is within threshold of %d°C)

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

vlt
Enable VLT traps.

vrrp
Enable VRRP state change traps

xstp
%SPANMG-5-STP_NEW_ROOT: New Spanning Tree Root, Bridge ID Priority 32768, Address 0001.e801.fc35
%SPANMG-5-STP_TOPOLOGY_CHANGE: Bridge port GigabitEthernet 11/38 transitioned from Forwarding to Blocking state.
%SPANMG-5-MSTP_NEW_ROOT_BRIDGE: Elected root bridge for instance 0.
%SPANMG-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.
%SPANMG-5-MSTP_TOPOLOGY_CHANGE: Topology change BridgeAddr: 0001.e801.fc35 Mstp Instance Id 0 port Gi 11/38 transitioned from forwarding to discarding state.

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_REMOTK_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

entity
Enable entity change traps

<table>
<thead>
<tr>
<th>Command Option</th>
<th>Trap</th>
</tr>
</thead>
<tbody>
<tr>
<td>envmon supply</td>
<td>FEM_FRLBLM: Major alarm: problem with power entry module %s</td>
</tr>
<tr>
<td></td>
<td>FEM_OK: Major alarm cleared: power entry module %s is good</td>
</tr>
<tr>
<td></td>
<td>MAJOR_PS: Major alarm: insufficient power %s</td>
</tr>
<tr>
<td></td>
<td>MAJOR_PS_CLR: major alarm cleared: sufficient power</td>
</tr>
<tr>
<td></td>
<td>MINOR_PS: Minor alarm: power supply non-redundant</td>
</tr>
<tr>
<td></td>
<td>MINOR_PS_CLR: Minor alarm cleared: power supply redundant</td>
</tr>
<tr>
<td>envmon temperature</td>
<td>MINOR_TEMP: Minor alarm: chassis temperature</td>
</tr>
<tr>
<td></td>
<td>MINOR_TEMP_CLR: Minor alarm cleared: chassis temperature normal (%s °C is within threshold of %d°C)</td>
</tr>
<tr>
<td></td>
<td>MAJOR_TEMP: Major alarm: chassis temperature high (%s °C temperature reaches or exceeds threshold of %d°C)</td>
</tr>
<tr>
<td></td>
<td>MAJOR_TEMP_CLR: Major alarm cleared: chassis temperature lower (%s °C °C is within threshold of %d°C)</td>
</tr>
<tr>
<td>envmon fan</td>
<td>FAN_TRAY_BAD: Major alarm: fantray %d is missing or down</td>
</tr>
<tr>
<td></td>
<td>FAN_TRAY_OK: Major alarm cleared: fan tray %d present</td>
</tr>
<tr>
<td></td>
<td>FAN_BAD: Minor alarm: some fans in fan tray %d are down</td>
</tr>
<tr>
<td></td>
<td>FAN_OK: Minor alarm cleared: all fans in fan tray %d are good</td>
</tr>
<tr>
<td>vlt</td>
<td>Enable VLT traps.</td>
</tr>
<tr>
<td>vrrp</td>
<td>Enable VRRP state change traps</td>
</tr>
<tr>
<td>xstp</td>
<td>%SPANMG-5-STP_NEW_ROOT: New Spanning Tree Root, Bridge ID Priority 32768, Address 0001.e801.fc35</td>
</tr>
<tr>
<td></td>
<td>%SPANMG-5-STP_TOPOLOGY_CHANGE: Bridge port GigabitEthernet 11/38 transitioned from Forwarding to Blocking state.</td>
</tr>
<tr>
<td></td>
<td>%SPANMG-5-MSTP_NEW_ROOT_BRIDGE: Elected root bridge for instance 0.</td>
</tr>
<tr>
<td></td>
<td>%SPANMG-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.</td>
</tr>
<tr>
<td></td>
<td>%SPANMG-5-MSTP_TOPOLOGY_CHANGE: Topology change BridgeAddr: 0001.e801.fc35 Mstp Instance Id 0 port Gi 11/38 transitioned from forwarding to discarding state.</td>
</tr>
<tr>
<td>ecfm</td>
<td>%ECFM-5-ECFM_XCON_ALARM: Cross connect fault detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000</td>
</tr>
<tr>
<td></td>
<td>%ECFM-5-ECFM_ERROR_ALARM: Error CCM Defect detected by MEP 1 in Domain customer1 at Level 7 VLAN 1000</td>
</tr>
<tr>
<td></td>
<td>%ECFM-5-ECFM_MAC_STATUS_ALARM: MAC Status Defect detected by MEP 1 in Domain provider at Level 4 VLAN 3000</td>
</tr>
<tr>
<td></td>
<td>%ECFM-5-ECFM_REMOTK_ALARM: Remote CCM Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000</td>
</tr>
<tr>
<td></td>
<td>%ECFM-5-ECFM_RDI_ALARM: RDI Defect detected by MEP 3 in Domain customer1 at Level 7 VLAN 1000</td>
</tr>
<tr>
<td>entity</td>
<td>Enable entity change traps</td>
</tr>
</tbody>
</table>

**Table 47-2. Dell Force10 Enterprise-specific SNMP Traps**
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 47-2. Dell Force10 Enterprise-specific SNMP Traps

<table>
<thead>
<tr>
<th>Command Option</th>
<th>Trap</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trap SNMPv2-MIB::sysUpTime.0 = Timeticks: (1489568) 4:08:15.68,SNMPv2-MIB::snmpTrapOID.0 = OID: SNMPv2-SMI::mib-2.47.2.0.1, SNMPv2-SMI::enterprises.6027.3.6.1.1.2.0 = INTEGER: 7</td>
</tr>
<tr>
<td></td>
<td>SNMP Copy Config Command Completed</td>
</tr>
<tr>
<td></td>
<td>%RPM0-P:CP %SNMP-4-RMON_RISING_THRESHOLD: RMON rising threshold alarm from SNMP OID &lt;oid&gt;</td>
</tr>
<tr>
<td></td>
<td>%RPM0-P:CP %SNMP-4-RMON_FALLING_THRESHOLD: RMON falling threshold alarm from SNMP OID &lt;oid&gt;</td>
</tr>
<tr>
<td></td>
<td>%RPM0-P:CP %SNMP-4-RMON_HC_RISING_THRESHOLD: RMON high-capacity rising threshold alarm from SNMP OID &lt;oid&gt;</td>
</tr>
</tbody>
</table>

### Table 47-3. MIB Objects for Copying Configuration Files via SNMP

<table>
<thead>
<tr>
<th>MIB Object</th>
<th>OID</th>
<th>Object Values</th>
<th>Description</th>
</tr>
</thead>
</table>
| copySrcFileType | .1.3.6.1.4.1.6027.3.5.1.1.1.2 | 1 = FTOS file 2 = running-config 3 = startup-config | Specifies the type of file to copy from. Range is:  
• 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 | .1.3.6.1.4.1.6027.3.5.1.1.1.3 | 1 = flash 2 = slot0 3 = tftp 4 = ftp 5 = scp 6 = usbflash | Specifies the location of source file.  
• If the copySrcFileLocation is FTP or SCP, copyServerAddress, copyUserName, and copyUserPassword must be specified. |
To copy a configuration file:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create an SNMP community string with read/write privileges.</td>
<td>snmp-server community community-name rw</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Copy the f10-copy-config.mib MIB from the Dell Force10 iSupport webpage to the server to which you are copying the configuration file.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 47-3. MIB Objects for Copying Configuration Files via SNMP

<table>
<thead>
<tr>
<th>MIB Object</th>
<th>OID</th>
<th>Object Values</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>copySrcFileName</td>
<td>1.3.6.1.4.1.6027.3.5.1.1.1.4</td>
<td>Path</td>
<td>Specifies name of the file.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(if file is not</td>
<td>• If copySourceFileType is set to running-config or startup-config,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>in current directory) and filename.</td>
<td>copySrcFileName is not required.</td>
</tr>
<tr>
<td>copyDestFileType</td>
<td>1.3.6.1.4.1.6027.3.5.1.1.1.5</td>
<td>1 = FTOS file</td>
<td>Specifies the type of file to copy to.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = running-config</td>
<td>• If the copySourceFileType is running-config or startup-config,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = startup-config</td>
<td>the default copyDestFileLocation is flash.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• If the copyDestFileType is a binary the copyDestFileLocation and</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>copyDestFileName must be specified.</td>
</tr>
<tr>
<td>copyDestFileLocation</td>
<td>1.3.6.1.4.1.6027.3.5.1.1.1.6</td>
<td>1 = flash</td>
<td>Specifies the location of destination file.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = slot0</td>
<td>• If the copyDestFileLocation is FTP or SCP, copyServerAddress,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = tftp</td>
<td>copyUserName, and copyUserPassword must be specified.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 = ftp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 = scp</td>
<td></td>
</tr>
<tr>
<td>copyDestFileName</td>
<td>1.3.6.1.4.1.6027.3.5.1.1.1.7</td>
<td>Path (if file is</td>
<td>Specifies the name of destination file.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>not in default directory) and filename.</td>
<td></td>
</tr>
<tr>
<td>copyServerAddress</td>
<td>1.3.6.1.4.1.6027.3.5.1.1.1.8</td>
<td>IP Address of</td>
<td>The IP address of the server.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the server</td>
<td>• If the copyServerAddress is specified so must copyUserName, and</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>copyUserPassword.</td>
</tr>
<tr>
<td>copyUserName</td>
<td>1.3.6.1.4.1.6027.3.5.1.1.1.9</td>
<td>Username for</td>
<td>Username for the FTP, TFTP, or SCP server.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the server.</td>
<td>• If the copyUserName is specified so must copyUserPassword.</td>
</tr>
<tr>
<td>copyUserPassword</td>
<td>1.3.6.1.4.1.6027.3.5.1.1.1.10</td>
<td>Password for the server.</td>
<td>Password for the FTP, TFTP, or SCP server.</td>
</tr>
</tbody>
</table>
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:** 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: FTOS-COPY-CONFIG-MIB::copySrcFileType.101

Table 47-4. Copying Configuration Files via SNMP

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
</table>
| Step 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 ranges.
- `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:** In UNIX, enter the command `snmpset` for help using this command. Place the file `f10-copy-config.mib` in the directory from which you are executing the `snmpset` command or in the `snmpset` tool path.

  **Note:** Use the following options in the `snmpset` command to view additional information:
  -c: View the community, either public or private
  -m: View the MIB files for the SNMP command
  -r: Number of retries using the option
  -t: View the timeout
  -v: View the SNMP version (either 1, 2, 2d or 3)

  **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` in the directory from which you are executing the `snmpset` command or in the `snmpset` tool path.

  **Note:** Use the following options in the `snmpset` command to view additional information:
  -c: View the community, either public or private
  -m: View the MIB files for the SNMP command
  -r: Number of retries using the option
  -t: View the timeout
  -v: View the SNMP version (either 1, 2, 2d or 3)

  **Table 47-4. Copying Configuration Files via SNMP**

<table>
<thead>
<tr>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copy the running-config to the startup-config using the following command from the UNIX machine:</td>
</tr>
</tbody>
</table>
Table 47-4. Copying Configuration Files via SNMP

<table>
<thead>
<tr>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>snmpset -v 2c -c public force10system-ip-address copySrcFileType.index i 2 copyDestFileType.index i 3</td>
</tr>
</tbody>
</table>

Figure 47-7 show the command syntax using MIB object names. Figure 47-8 shows the same command using the object OIDs. In both cases, the object is followed by a unique index number.

**Figure 47-7. Copying Configuration Files via SNMP using Object-Name Syntax**

```bash
> snmpset -v 2c -r 0 -t 60 -c private -m ./f10-copy-config.mib 10.10.10.10 copySrcFileType.101 i 2 copyDestFileType.101 i 3
FTOS-COPY-CONFIG-MIB::copySrcFileType.101 = INTEGER: runningConfig(2)
FTOS-COPY-CONFIG-MIB::copyDestFileType.101 = INTEGER: startupConfig(3)
```

**Figure 47-8. Copying Configuration Files via SNMP using OID Syntax**

```bash
> snmpset -v 2c -c private -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
FTOS-COPY-CONFIG-MIB::copySrcFileType.100 = INTEGER: runningConfig(2)
FTOS-COPY-CONFIG-MIB::copyDestFileType.100 = INTEGER: startupConfig(3)
```

Copy the startup-config to the running-config using the following command from a UNIX machine:

```bash
snmpset -c private -v 2c force10system-ip-address copySrcFileType.index i 2 copyDestFileType.index i 3
```

**Figure 47-9. Copying Configuration Files via SNMP using Object-Name Syntax**

```bash
> snmpset -c public -v 2c -m ./f10-copy-config.mib 10.11.131.162 copySrcFileType.7 i 3 copyDestFileType.7 i 2
FTOS-COPY-CONFIG-MIB::copySrcFileType.7 = INTEGER: runningConfig(3)
FTOS-COPY-CONFIG-MIB::copyDestFileType.7 = INTEGER: startupConfig(2)
```

**Figure 47-10. Copying Configuration Files via SNMP using OID Syntax**

```bash
> snmpset -c public -v 2c 10.11.131.162 .1.3.6.1.4.1.6027.3.5.1.1.1.2.8 i 3 .1.3.6.1.4.1.6027.3.5.1.1.1.5.8 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:

```bash
snmpset -v 2c -c private ./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 47-4. Copying Configuration Files via SNMP

<table>
<thead>
<tr>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>• server-ip-address must be preceded by the keyword a.</td>
</tr>
<tr>
<td>• values for copyUsername and copyUserPassword must be preceded by the keyword s.</td>
</tr>
</tbody>
</table>

Figure 47-11. Copying Configuration Files via SNMP and FTP to a Remote Server

Copy the startup-config to the server via TFTP using the following command from the UNIX machine:

```bash
snmpset -v 2c -c public -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
```

FTOS-COPY-CONFIG-MIB::copySrcFileType.110 = INTEGER: runningConfig(2)
FTOS-COPY-CONFIG-MIB::copyDestFileName.110 = STRING: /home/startup-config
FTOS-COPY-CONFIG-MIB::copyDestFileLocation.110 = INTEGER: ftp(4)
FTOS-COPY-CONFIG-MIB::copyServerAddress.110 = IpAddress: 11.11.11.11
FTOS-COPY-CONFIG-MIB::copyUserName.110 = STRING: mylogin
FTOS-COPY-CONFIG-MIB::copyUserPassword.110 = STRING: mypass

Figure 47-12. Copying Configuration Files via SNMP and TFTP to a Remote Server

Copy a binary file from the server to the startup-configuration on the Dell Force10 system via FTP using the following command:

```bash
snmpset -v 2c -c private -m ./f10-copy-config.mib force10system-ip-address copySrcFileType.index i 3
    copyDestFileType.index i 1
    copyDestFileName.index s filepath/filename
    copyDestFileLocation.index i 3
    copyServerAddress.index a server-ip-address
```

Figure 47-13. Copying Configuration Files via SNMP and FTP from a Remote Server

Copy a binary file from the server to the startup-configuration on the Dell Force10 system via FTP using the following command:

```bash
snmpset -v 2c -c private -m ./f10-copy-config.mib force10system-ip-address copySrcFileType.index i 1
    copySrcFileType.index i 3
    copySrcFileLocation.index i 4
    copySrcFileName.index s filepath/filename
    copyDestFileType.index i 3
    copyServerAddress.index a server-ip-address
    copyUserName.index s server-login-id
    copyUserPassword.index s server-login-password
```

```bash
> 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 47-5.

<table>
<thead>
<tr>
<th>MIB Object</th>
<th>OID</th>
<th>Values</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>copyState</td>
<td>.1.3.6.1.4.1.6027.3.5.1.1.1.11</td>
<td>1 = running</td>
<td>Specifies the state of the copy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = successful</td>
<td>operation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = failed</td>
<td></td>
</tr>
<tr>
<td>copyTimeStarted</td>
<td>.1.3.6.1.4.1.6027.3.5.1.1.1.12</td>
<td>Time value</td>
<td>Specifies the point in the up-time</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>clock that the copy operation started.</td>
</tr>
<tr>
<td>copyTimeCompleted</td>
<td>.1.3.6.1.4.1.6027.3.5.1.1.1.13</td>
<td>Time value</td>
<td>Specifies the point in the up-time</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>clock that the copy operation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>completed.</td>
</tr>
<tr>
<td>copyFailCause</td>
<td>.1.3.6.1.4.1.6027.3.5.1.1.1.14</td>
<td>1 = bad file name</td>
<td>Specifies the reason the copy request</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = copy in progress</td>
<td>failed.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = disk full</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 = file exists</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 = file not found</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 = timeout</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>7 = unknown</td>
<td></td>
</tr>
<tr>
<td>copyEntryRowStatus</td>
<td>.1.3.6.1.4.1.6027.3.5.1.1.1.15</td>
<td>Row status</td>
<td>Specifies the state of the copy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>operation.</td>
</tr>
</tbody>
</table>

To obtain a value for any of the MIB Objects in Table 47-5:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Get a copy-config MIB object value.</td>
</tr>
</tbody>
</table>

```
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 47-15.

Figure 47-14 and Figure 47-15 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 47-14 shows the command syntax using MIB object names, and Figure 47-15 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 47-14. Obtaining MIB Object Values for a Copy Operation using Object-name Syntax

```shell
> snmpget -v 2c -c private -m ./f10-copy-config.mib 10.11.131.140 copyTimeCompleted.110
FTOS-COPY-CONFIG-MIB::copyTimeCompleted.110 = Timeticks: (1179831) 3:16:38.31
```

Figure 47-15. Obtaining MIB Object Values for a Copy Operation using OID Syntax

```shell
> snmpget -v 2c -c private 10.11.131.140 .1.3.6.1.4.1.6027.3.5.1.1.1.13.110
SNMPv2-SMI::enterprises.6027.3.5.1.1.1.13.110 = 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 to manage VLANs.

Create a VLAN

Use the dot1qVlanStaticRowStatus object to create a VLAN. The `snmpset` operation in Figure 47-16 creates VLAN 10 by specifying a value of 4 for instance 10 of the dot1qVlanStaticRowStatus object.

Figure 47-16. Creating a VLAN using SNMP

```shell
> 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 47-17.
Figure 47-17. 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
```

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 47-18.

Figure 47-18. 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
```

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 47-19.

Figure 47-19. 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 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 on the S55 and S60; 64 ports on the S4810. On the S55 and S60, the last stack unit is assigned 8 pairs; the eighth pair is unused. On the S4810, the last stack unit begins on the 66th bit.

The first hex pair, 00 in Figure 47-19, 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 47-19 shows the output for an S-Series. All hex pairs are 00, indicating that no ports are assigned to VLAN 10. In Figure 47-20, Port 0/2 is added to VLAN 10 as untagged. And the first hex pair changes from 00 to 04.

**Figure 47-20. Displaying Ports in a VLAN using SNMP**

```plaintext
[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

<table>
<thead>
<tr>
<th>NUM</th>
<th>Status</th>
<th>Description</th>
<th>Q Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Inactive</td>
<td></td>
<td>U Gi 0/2</td>
</tr>
</tbody>
</table>

[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
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 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 47-21.
- To add an untagged port to a VLAN, write the port to the dot1qVlanStaticEgressPorts and dot1qVlanStaticUntaggedPorts objects, as shown in Figure 47-22.

**Note:** Whether adding a tagged or untagged port, you must specify values for both dot1qVlanStaticEgressPorts and dot1qVlanStaticUntaggedPorts.

In Figure 47-21, Port 0/2 is added as an untagged member of VLAN 10.

**Figure 47-21. 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"
```

```
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 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 00
```

In Figure 47-22, Port 0/2 is added as a tagged member of VLAN 10.

Figure 47-22. Adding Tagged Ports to a VLAN using SNMP

Managing Overload on Startup

If you are running IS-IS, you can set a specific amount of time to prevent ingress traffic from being received after a reload and allow the routing protocol upgrade process to complete.

Use the following command to prevent ingress traffic on a router while the IS reload is implemented:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set the amount of time after an IS-IS reload is performed before ingress traffic is allowed at startup.</td>
<td>set-overload-bit on-startup isis</td>
</tr>
</tbody>
</table>

The following OIDs are configurable through the snmpset command.

The node OID is 1.3.6.1.4.1.6027.3.18

F10-ISIS-MIB::f10IsisSysOloadSetOverload
F10-ISIS-MIB::f10IsisSysOloadSetOloadOnStartupUntil
F10-ISIS-MIB::f10IsisSysOloadWaitForBgp
F10-ISIS-MIB::f10IsisSysOloadV6SetOverload
F10-ISIS-MIB::f10IsisSysOloadV6SetOloadOnStartupUntil
F10-ISIS-MIB::f10IsisSysOloadV6WaitForBgp

To enable overload bit for IPv4 set 1.3.6.1.4.1.6027.3.18.1.1 and IPv6 set 1.3.6.1.4.1.6027.3.18.1.4
To set time to wait set 1.3.6.1.4.1.6027.3.18.1.2 and 1.3.6.1.4.1.6027.3.18.1.5 respectively
To set time to wait till bgp session are up set 1.3.6.1.4.1.6027.3.18.1.3 and 1.3.6.1.4.1.6027.3.18.1.6
Enable and Disable a Port using SNMP

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Create an SNMP community on the Dell Force10 system.</td>
<td>snmp-server community</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>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 snmpwalk command to identify the interface index.</td>
<td>show interface</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>3</td>
<td>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.</td>
<td>snmpset with descriptor: snmpset -v version -c community agent-ip ifAdminStatus.ifindex i {1</td>
<td>2} snmpset with OID: snmpset -v version -c community agent-ip .1.3.6.1.2.1.2.2.1.7.ifindex i {1</td>
</tr>
</tbody>
</table>

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 47-6. MIB Objects for Fetching Dynamic MAC Entries in the Forwarding Database

<table>
<thead>
<tr>
<th>MIB Object</th>
<th>OID</th>
<th>Description</th>
<th>MIB</th>
</tr>
</thead>
<tbody>
<tr>
<td>dot1dTpFdbTable</td>
<td>.1.3.6.1.2.1.17.4.3</td>
<td>List the learned unicast MAC addresses on the default VLAN.</td>
<td>Q-BRIDGE MIB</td>
</tr>
<tr>
<td>dot1qTpFdbTable</td>
<td>.1.3.6.1.2.1.17.7.1.2</td>
<td>List the learned unicast MAC addresses on non-default VLANs.</td>
<td>F10-LINK-AGGREGATION-MIB</td>
</tr>
<tr>
<td>dot3aCurAggFdbTable</td>
<td>.1.3.6.1.4.1.6027.3.2</td>
<td>List the learned MAC addresses of aggregated links (LAG).</td>
<td>F10-LINK-AGGREGATION-MIB</td>
</tr>
</tbody>
</table>

In Figure 47-23, 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 0.0.1.232.6.149.172.

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 47-23. Fetching Dynamic MAC Addresses on the Default VLAN

<table>
<thead>
<tr>
<th>VlanId</th>
<th>Mac Address</th>
<th>Type</th>
<th>Interface</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>00:01:e8:06:95:ac</td>
<td>Dynamic</td>
<td>Gi 1/21</td>
<td>Active</td>
</tr>
</tbody>
</table>

---

### Query from Management Station

```bash
>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
```

In **Figure 47-24**, 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 47-24. Fetching Dynamic MAC Addresses on Non-default VLANs

<table>
<thead>
<tr>
<th>VlanId</th>
<th>Mac Address</th>
<th>Type</th>
<th>Interface</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>00:01:e8:06:95:ac</td>
<td>Dynamic</td>
<td>Gi 1/21</td>
<td>Active</td>
</tr>
</tbody>
</table>

---

### Query from Management Station

```bash
>snmpwalk -v 2c -c techpubs 10.11.131.162 .1.3.6.1.2.1.17.7.1.2.2.1
```

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 47-25. Fetching Dynamic MAC Addresses on the Default VLAN

<table>
<thead>
<tr>
<th>VlanId</th>
<th>Mac Address</th>
<th>Type</th>
<th>Interface</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>00:01:e8:06:95:ac</td>
<td>Dynamic</td>
<td>Po 1</td>
<td>Active</td>
</tr>
</tbody>
</table>

---

### Query from Management Station

```bash
>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 47-26.

**Figure 47-26. 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]
```

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 47-27. Interface Index Binary Calculations**

```
<table>
<thead>
<tr>
<th>1 bit</th>
<th>1 bit</th>
<th>5 bits</th>
<th>7 bits</th>
<th>4 bits</th>
<th>14 bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unused</td>
<td>P/L Flag</td>
<td>Slot Number</td>
<td>Port Number</td>
<td>Interface Type</td>
<td>Card Type</td>
</tr>
</tbody>
</table>
```

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 10001011000110000000111010 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 47-28. 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.
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 47-28 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.

```bash
$senthilnathan@lithium ~$ 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.6027.3.2.1.1.1.4.1 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.4.2 = INTEGER: 1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.5.1 = Hex-STRING: 00 00
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.5.2 = Hex-STRING: 00 00
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.6.1 = STRING: "Gi 5/84 " << Channel member for Po1
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.6.2 = STRING: "Gi 5/85 " << Channel member for Po2
dot3aCommonAggFdbIndex
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.6.1.1107755009.1 = INTEGER: 1107755009
dot3aCommonAggFdbVlanId
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.6.1.2.1107755009.1 = INTEGER: 1
dot3aCommonAggFdbTagConfig
SNMPv2-SMI::enterprises.6027.3.2.1.1.1.6.1.3.1107755009.1 = INTEGER: 2 (Tagged 1 or Untagged 2)
dot3aCommonAggFdbStatus
```
If we learn MAC addresses for the LAG, status will be shown for those as well.

dot3aCurAggVlanId
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.1.0.0.0.0.1.1 = INTEGER: 1

dot3aCurAggMacAddr
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.2.1.0.0.0.0.1.1 = Hex-STRING: 00 00 00 00 01

dot3aCurAggIndex
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.3.1.0.0.0.0.1.1 = INTEGER: 1

dot3aCurAggStatus
SNMPv2-SMI::enterprises.6027.3.2.1.1.4.1.4.1.0.0.0.0.1.1 = INTEGER: 1 << Status active, 2 – status inactive

Layer 3 LAG does not include this support.

SNMP trap works fine for the Layer 2 / Layer 3 / 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.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]:

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 \texttt{snmpwalk} command, the output for echo replies may be incorrectly displayed. Use the \texttt{show ip traffic} command to correctly display this information under ICMP statistics.

- When you query an icmpStatsInErrors object in the icmpStats table by using the \texttt{snmpget} or \texttt{snmpwalk} command, the output for IPv4 addresses may be incorrectly displayed. Use the \texttt{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 \texttt{snmpwalk} command, the echo response output may not be displayed. Use the \texttt{show ip traffic} command to correctly display ICMP statistics, such as echo response.
Stacking

Stacking is supported on the following platforms: S (S50/S25), S55, S60, S4810

Stacking is supported on the S4810 platform with FTOS version 8.3.7.1, version 8.3.10.2 and newer.

Note: The S4810 commands accept Unit ID numbers 0-11, though the S4810 supports stacking up to 3 units only with FTOS version 8.3.7.1 and version 8.3.10.2. The S4810 supports stacking up to 6 units on FTOS version 8.3.12.0. The S55 supports stacking up to 8 units.

Using the FTOS stacking feature, multiple S-Series switch units can be interconnected with dedicated stacking ports or front end user ports. (The S50, S55, and S60 use dedicated stacking ports; the S4810 uses front end user ports for stacking.) The stack becomes manageable as a single switch through the stack management unit.

Note: Stacking S55 and S60 units with other devices is not supported.

This chapter contains the following sections:

- S-Series Stacking Overview
- Stack Management Roles
- Stack Master Election
- Stack Group/Port Numbers
- High Availability on S-Series Stacks
- Important Points to Remember - S4810 Stacking
- S-Series Stacking Configuration Tasks
- Troubleshoot an S-Series Stack
- Removing Units or Front End Ports from a Stack

S-Series Stacking Overview

An S-Series stack is analogous to an E-Series or C-Series system with redundant RPMs and multiple line cards. FTOS elects a management (master) unit, a standby unit, and all other units are member units.
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.

**Stack Management Roles**

The stack elects the management units for the stack management:

- **Stack master**: The primary management unit, also called the **master** unit.
- **Standby**: The secondary management unit.
- **Stack units**: Also called **stack members**, these are the remaining units in the stack. The system supports up to four S4810 stack units.
- **Stack group**: On the S4810, each set of 4 10G ports or each individual 40G port correspond to a stack-group. The CLI is used to configure the front ports on the S4810 to be stacking-ports.

The master holds the control plane and the other units maintain a local copy of the forwarding databases. From the stack master you can configure:

- System-level features that apply to all stack members.
- Interface-level features for each stack member.

The master synchronizes the following information with the standby unit:

- Stack unit topology
- Stack running configuration (which includes ACL, LACP, STP, SPAN, etc.)
- Logs

The master switch maintains stack operation with minimal impact in the event of:

- Switch failure
- Inter-switch stacking link failure
- Switch insertion
- Switch removal

If the master switch goes off line, the standby replaces it as the new master and the switch with the next highest priority or MAC address becomes standby.

**Stack Master Election**

The stack elects a master and standby unit at bootup time based on two criteria:

- **Unit priority**: User-configurable. Range is from 1 to 14. A higher value (14) means a higher priority. Default: 1. By removing the stack-unit priority using the `no stack-unit priority` command, you can set the priority back to the default value of zero. The unit with the highest priority is elected the master management unit; the unit with the second highest priority is elected the standby unit.
- MAC address (in case of priority tie): The unit with the higher MAC value becomes the master unit. The stack takes the MAC address of the master unit and retains it unless it is reloaded.

To view which switch is the stack master, enter the `show system` command. Figure 48-1 shows sample output from an established stack.

A change in the stack master occurs when:

- You power down the stack master or bring the master switch offline.
- A failover of the master switch occurs.
- You disconnect the master switch from the stack.

When a stack reloads and all the units come up at the same time, for example, when all units boot up from flash, all units participate in the election and the master and standby are chosen based on the priority or MAC address. When the units do not boot up at the same time, such as when some units are powered down just after reloading and powered up later to join the stack, they do not participate in the election process even though the units that boot up late may have a higher priority configured. This happens because the master and standby have already been elected, hence the unit that boots up late joins only as a member. When an up and running standalone unit or stack is merged with another stack, based on election, the losing stack reloads and the master unit of the winning stack becomes the master of the merged stack. For more details, see sections “Add Units to an Existing S-Series Stack on page 985” and “Remove a Unit from an S-Series Stack on page 995”. It is possible to reset individual units to force them to give up the management role or reload the whole stack from the CLI to ensure a fully synchronized bootup.

**Figure 48-1. Displaying the Stack Master**

```
FTOS#show system brief
Stack MAC : 00:01:e8:8c:53:32
Reload Type : normal-reload [Next boot : normal-reload]

-- Stack Info --
Unit UnitType Status ReqTyp CurTyp Version Ports
----------------------------------------------------------------------------
0  Member  not present
1  Management  online  S4810  S4810  4810-8-3-12-1447  64
2  Standby  online  S4810  S4810  4810-8-3-12-1447  64
3  Member  online  S4810  S4810  4810-8-3-12-1447  64
4  Member  online  S4810  S4810  4810-8-3-12-1447  64
5  Member  not present
6  Member  not present
7  Member  not present
8  Member  not present
9  Member  not present
10  Member  not present
```

**Virtual IP**

The stack can be managed using a single IP, known as a virtual IP, that is retained in the stack even after a failover. The virtual IP address is used to log in to the current master unit of the stack. Both IPv4 and IPv6 addresses are supported as virtual IPs.
Failover Roles

If the stack master fails (e.g., is powered off), it is removed from the stack topology. The standby unit detects the loss of peering communication and takes ownership of the stack management, switching from the standby role to the master role. The distributed forwarding tables are retained during the failover, as is the stack MAC address. The lack of a standby unit triggers an election within the remaining units for a standby role.

Once the former master switch recovers, despite having a higher priority or MAC address, it will not recover its master role but instead takes the next available role.

To view failover details, use the `show redundancy` command.

MAC Addressing on S-Series Stacks

The S-Series has three MAC addresses: the chassis MAC, interface MAC, and null interface MAC. All interfaces in the stack use the interface MAC address of the management unit, 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 48-2 and Figure 48-3, a standalone is added to a stack. The standalone and the master unit have the same priority, but the standalone has a lower MAC address, so the standalone reboots. In Figure 48-3, 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 48-2. Adding a Standalone with a Lower MAC Address to a Stack—Before (S50-type)

----------------------------------------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 48-3. Adding a Standalone with a Lower MAC Address and Equal Priority to a Stack—After

Stacking LAG

When multiple links are used between stack units, FTOS automatically bundles them in a stacking LAG to provide aggregated throughput and redundancy. The stacking LAG is established automatically and transparently by FTOS (without user configuration) once peering is detected and behaves as follows:

- The stacking LAG dynamically aggregates; it can lose link members or gain new links.
- Shortest path selection inside the stack: If multiple paths exist between two units in the stack, the shortest path is used.

Supported Stacking Topologies

The S4810 supports stacking in a ring or a daisy chain topology. Dell Force10 recommends the ring topology when stacking S4810 switches to provide redundant connectivity.
High Availability on S-Series Stacks

S-Series stacks have master and standby management units analogous to Dell Force10 Route Processor Modules (Figure 48-5). The master unit synchronizes the running configuration and protocol states so that the system fails over in the event of a hardware or software fault on the master unit. In such an event, or when the master unit is removed, the standby unit becomes the stack manager and FTOS elects a new standby unit. FTOS resets the failed master unit: once online, it becomes a member unit; the remaining members remain online.
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 master unit (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 48-6.

- **Remote access**: You may access the master unit and standby unit in a stack through the dedicated management ethernet interfaces with SNMP, SSH, or via Telnet.
Stacking Cable Redundancy

You can connect two units with two or more stacking cables in case of a stacking port or cable failure. Removal of only one of the cables does not trigger a reset.

Important Points to Remember - S4810 Stacking

- For S4810 stacking using 40GE ports, any remaining 40 GE ports should not be configured as quad-mode.
- You may stack up to six S4810 systems.
- The S4810 cannot be stacked with other system types.
- Stacking and Virtual Link Trunking (VLT) cannot be enabled simultaneously on the S4810.
- Data ports are configured as stacking ports in predefined groups of 4-10G ports called stack-groups. When using the 40G ports, a single port can be configured as a stack port; each 40G port is a stack-group.
- All the ports in a stack-group are placed in stacking mode. Unused ports in that group cannot be used as data ports.
- The stack-group must contain only ports of the same type (all 10G or all 40G).
  - Stacking with 1G interfaces is not supported.
- Stacking on the S4810 is accomplished through front end user ports on the chassis.
- All stack units must have the same version of FTOS.
S-Series Stacking Installation Tasks

- Create an S-Series Stack
- Add Units to an Existing S-Series Stack
- Remove a Unit from an S-Series Stack
- Split an S-Series Stack

Create an S-Series Stack

Stacking is enabled on the S4810 using the front end ports. No configuration is allowed on front end ports used for stacking. Stacking can be made between 10G ports of two units or 40G ports of two units. The stack links between the two units will be grouped into a single LAG.

Stack Group/Port Numbers

By default, each unit in standalone mode is numbered stack-unit 0. A maximum of eight 10G stack links or two 40G stack links can be made between two units in a stack. The front end ports are divided into 16 stack groups, each with 40G of bandwidth. Stack groups 0 through 11 correspond to 10G stack groups with four ports each. Stack groups 12 to 15 are one 40G port each.

The front end ports accommodate SFP, SFP+ and QSFP+.

- Ports are divided into 16 stack-groups (0 to 15) as shown in the following example. The stack groups must be of a single speed - either all 10G or all 40G.
  - stack-group 0 corresponds to ports 0-3, stack-group 1 corresponds to ports 4-7, so on through stack-group 11
  - stack-group 12 corresponds to the 40G port 48, stack-group 13 corresponds to port 52, so on through stack group 15.

Figure 48-7. S4810 Stack-group assignments

You can connect the units while they are powered down or up. Stacking ports are bi-directional.
With FTOS 8.3.12.0, when a unit is added to a stack, the management unit performs a system check on the new unit to ensure the hardware type is compatible. A similar check is performed on the FTOS version. If the stack is running 8.3.12.0 and the new unit is running an earlier software version, the new unit is put into a card problem state.

- If the unit is running version 8.3.10.x, it is upgraded to use the same FTOS version as the stack, rebooted and join the stack.
- If the new unit is running an FTOS version prior to 8.3.10.x, the unit is put into a card problem state, FTOS is not upgraded, and a syslog message is raised. The unit must be upgraded to FTOS 8.3.12.0 before you can proceed.

Syslog messages are generated by the management unit:

- before the management unit downloads its FTOS version (FTOS 8.3.12.0 or later) to the new unit. The syslog includes the unit number, previous version and version being downloaded.
- when the firmware synchronization is complete.
- if the system check fails, a message such as a hardware incompatibility message or incompatible uboot version is generated. If the unit is placed in a card problem state, the management unit also generates an SNMP trap.
- if the software version of the new unit predates FTOS 8.3.12.0, the management unit puts the new unit into a card problem state and generates a syslog that identifies the unit, its FTOS version, and its incompatibility for firmware synchronization.

**Note:** You must enter the `stack-unit stack-unit stack-group stack-group` command when adding units to a stack to ensure the units are assigned to the correct groups.

**Note:** Any scripts used to streamline the stacking configuration process must be updated to reflect the Command Mode change from EXEC to CONFIGURATION to allow the scripts to work correctly.

### Enable Front End Port Stacking

To enable the front ports on an S4810 unit for stacking, use the following procedure.

**Note:** Once a port has been allocated for stacking, it can only be used for stacking. If stack-group 0 is allocated for stacking, then ports 0, 1, 2, and 3 can be used for stacking but not for Ethernet anymore. If only port 0 is used for stacking, ports 1, 2, and 3 are just spare; they cannot be used for Ethernet.

After the front ports are enabled, use the `show system stack-unit` command to view the port assignments,

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assign a stack group for each unit. Begin with the first port on the management unit. Next, configure both ports on each subsequent unit. Finally, return to the management unit and configure the last port. (See the example below.)</td>
<td><code>stack-unit id stack-group id</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

Range: 0-15
Creating a New Stack

Prior to creating a stack, know which unit will be the management unit and which will be the standby unit. You must also enable the front ports of the units for stacking. For more information, see “Enable Front End Port Stacking” on page 981.

To create a new stack:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Power up all units in the stack</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Verify that each unit has the same FTOS version prior to stacking them together.</td>
<td>show version</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>3</td>
<td>Manually configure unit numbers for each unit, so that the stacking is deterministic upon boot up. Renumbering will cause the unit to reboot. The stack-unit default for all new units is stack-unit 0.</td>
<td>stack-unit 0 renumber</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>
In the following example, stack unit 1 will be the master management unit, stack unit 2 will be the standby unit. The cables are connected to each unit.

Configure the stack groups on the units in the following order:
Configure the first stack group on unit 1:

```
stack-unit 1 stack-group 13
```

Configure the stack groups on unit 2:

```
stack-unit 2 stack-group 14
stack-unit 2 stack-group 15
```

Configure the stack groups on unit 3:

```
stack-unit 3 stack-group 12
stack-unit 3 stack-group 13
```

Configure the stack groups on unit 4:

```
stack-unit 4 stack-group 13
stack-unit 4 stack-group 14
```

Configure the final stack-group on unit 1 to complete the stack.

```
stack-unit 1 stack-group 12
```

When the stack-group configuration is complete, the system will print a syslog for reload, as shown below.

```
FTOS#configure
FTOS(config)#stack-unit 4 stack-group 13
FTOS(config)#02:39:12: %STKUNIT4-M:CP %IFMGR-6-STACK_PORTS_ADDED: Ports Fo 4/52 have been configured as stacking ports. Please save and reload for config to take effect
FTOS(config)#stack-unit 4 stack-group 14
FTOS(config)#02:39:15: %STKUNIT4-M:CP %IFMGR-6-STACK_PORTS_ADDED: Ports Fo 4/56 have been configured as stacking ports. Please save and reload for config to take effect
FTOS(config)#
FTOS#02:39:18: %STKUNIT4-M:CP %SYS-5-CONFIG_I: Configured from console
```

Reload each unit in the stack. After the reload is complete, the four units will come up as a stack with unit 1 as the management unit, unit 2 as the standby unit, and the remaining units as stack-members as shown in the following example. All units in the stack can be accessed from the management unit.

```
FTOS#show system brief
Stack MAC : 00:01:e8:8c:53:32
Reload Type : normal-reload [Next boot : normal-reload]

--- Stack Info ---
Unit UnitType Status ReqTyp CurTyp Version Ports
--- "--- --- "-"-"-" "-" "-"-"--- "-"-"-"---
0 Member not present
1 Management online S4810 S4810 4810-8-3-12-1447 64
2 Standby online S4810 S4810 4810-8-3-12-1447 64
3 Member online S4810 S4810 4810-8-3-12-1447 64
4 Member online S4810 S4810 4810-8-3-12-1447 64
5 Member not present
6 Member not present
7 Member not present
8 Member not present
```
The following example shows how to configure two new S4810 switches for stacking using 10G ports.

S4810-1#stack-unit 0 stack-group 0
Setting ports Te 0/0 Te 0/1 Te 0/2 Te 0/3 as stack group will make their interface configs obsolete after a reload.
[confirm yes/no]:yes

S4810-2#stack-unit 0 stack-group 0
Setting ports Te 0/0 Te 0/1 Te 0/2 Te 0/3 as stack group will make their interface configs obsolete after a reload.
[confirm yes/no]:yes

S4810-1#show system stack-ports
Topology: Ring
Interface  Connection  Link Speed  Admin Link Trunk
                                      (Gb/s)      Status  Status Group
------------------------------------------------------------------
0/0        1/0          10          up  up
0/1        1/1          10          up  up
0/2        1/2          10          up  up
0/3        1/3          10          up  up

S4810-1#

Add Units to an Existing S-Series Stack

Units can be added to an existing stack in one of three ways:

• by manually assigning a new un-configured unit a position in an existing stack.
• by adding a configured unit to an existing stack.
• by merging two stacks.

If you are adding units to an existing stack, you can either:

• allow FTOS to automatically assign the new unit a position in the stack, or
• manually determine each unit's position in the stack by configuring each unit to correspond with the stack before connecting it.

• 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 the examples below.

• 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 shown in Figure 48-10 and Figure 48-11.

After the new unit loads, it synchronizes its running and startup configurations with the stack.

Manually Assign a New Unit to an Existing Stack

To manually assign a new unit a position in an existing stack, use the following steps.

Note: For an S50 system, install stacking modules in the new unit while the unit is unpowered.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>On the stack, determine the next available stack-unit number, and the management priority of the management unit.</td>
<td>show system brief show system stack-unit</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>2</td>
<td>On the new unit, number it the next available stack-unit number.</td>
<td>stack-unit renumber</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>3</td>
<td>(OPTIONAL) On the new unit, assign a management priority based on whether you want the new unit to be the stack manager.</td>
<td>stack-unit priority</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>4</td>
<td>Assign a stack group to each unit.</td>
<td>stack-unit id stack-group id</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>5</td>
<td>Connect the new unit to the stack using stacking cables.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 48-8. Adding an S4810 Stack Unit with a Conflicting Stack Number—Before

FTOS#show system brief

Stack MAC : 00:01:e8:8a:df:e6

Reload Type : normal-reload

-- Stack Info --

<table>
<thead>
<tr>
<th>Unit</th>
<th>UnitType</th>
<th>Status</th>
<th>ReqTyp</th>
<th>CurTyp</th>
<th>Version</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Management</td>
<td>online</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-7-13</td>
<td>64</td>
</tr>
<tr>
<td>1</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Standby</td>
<td>online</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-7-13</td>
<td>64</td>
</tr>
</tbody>
</table>
Figure 48-9. Adding an S4810 Stack Unit with a Conflicting Stack Number—After

```
FTOS#show system brief
Stack MAC : 00:01:e8:8a:df:e6

Reload Type : normal-reload
```

<table>
<thead>
<tr>
<th>-- Stack Info --</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>0 Management</td>
</tr>
<tr>
<td>1 Member</td>
</tr>
<tr>
<td>2 Member</td>
</tr>
<tr>
<td>3 Standby</td>
</tr>
<tr>
<td>4 Member</td>
</tr>
<tr>
<td>5 Member</td>
</tr>
<tr>
<td>6 Member</td>
</tr>
<tr>
<td>7 Member</td>
</tr>
<tr>
<td>8 Member</td>
</tr>
<tr>
<td>9 Member</td>
</tr>
<tr>
<td>10 Member</td>
</tr>
<tr>
<td>11 Member</td>
</tr>
</tbody>
</table>

Add a Configured Unit to an Existing Stack

To add a configured unit to an existing stack use the following steps.

If a stack unit goes down and is removed from the stack, the logical provisioning configured for that stack-unit number is saved on the master and standby units. When a new unit is added to the stack, if a stack group configuration conflict occurs between the new unit and the provisioned stack unit, the configuration of the new unit takes precedence.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Add the configured unit to the top or bottom of the stack.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Power on the switch.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Attach cables to connect ports on the added switch to one or more existing switches in the stack.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 4    | Logon to the CLI and enter global configuration mode. | Login: username  
Password: *****  
FTOS> enable  
FTOS# configure | |
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 user port cables from the front end user port.

- FTOS selects a master stack manager from the two existing managers based on the priority of the stack.
- FTOS resets all the units in the losing stack; 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 to synchronize the configuration on the new stack members.

If a standalone switch already has stack groups configured.

FTOS Behavior: When you add a new switch to a stack:
- If the new unit has been configured with a stack number that is already assigned to a stack member, the stack avoids a numbering conflict by assigning the new switch the first available stack number.
- If the stack has been provisioned for the stack number that is assigned to the new unit, the pre-configured provisioning must match the switch type. If there is a conflict between the provisioned switch type and the new unit, a mismatch error message is displayed.

### Merge Two S-Series Stacks

#### Step Task
5 Configure the ports on the added switch for stacking, where:
- **stack-unit 0** defines the default ID unit-number in the initial configuration of a switch.
- **stack-group group-number** configures a port for stacking.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Configure the ports on the added switch for stacking, where:</td>
<td>stack-unit 0 stack-group group-number</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td></td>
<td>stack-unit 0 defines the default ID unit-number in the initial configuration of a switch.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>stack-group group-number configures a port for stacking.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Save the stacking configuration on the ports.</td>
<td>write memory</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>7</td>
<td>Reload the switch. FTOS automatically assigns a number to the new unit and adds it as member switch in the stack. The new unit synchronizes its running and startup configurations with the stack.</td>
<td>reload</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>8</td>
<td>Attach cables to connect the ports already configured as stack groups on the switch to one or more switches in the stack.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>FTOS automatically assigns a number to the new unit and adds it as member switch in the stack. The new unit synchronizes its running and startup configurations with the stack.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 master and the standby management units, it is unaffected by the split.
- If one of the new stacks receives only the master unit, that unit remains the stack manager, and FTOS elects a new standby management unit.
- If one of the new stacks receives only the standby unit, it becomes the master unit of the new stack, and FTOS elects a new standby unit.
- If one of the new stacks receives neither the master nor the standby 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**
- **Create a Virtual Stack Unit on an S-Series Stack**
- **Display Information about an S-Series Stack**
- **Influence Management Unit Selection on an S-Series Stack**
- **Manage Redundancy on an S-Series Stack**
- **Reset a Unit on an S-Series Stack**
- **Recover from Stack Link Flaps**

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 11, however, only six (6) S4810 units can be stacked. Stack numbers are stored in NVRAM and are preserved upon reload.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assign a stack-number to a unit.</td>
<td>stack-unit renumber</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

**Note:** Renumbering the stack manager triggers the whole stack to reload as indicated in **Message 1**. When the stack comes back online, the master unit remains the management unit.
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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create a virtual stack unit.</td>
<td>stack-unit provision</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

Display Information about an S-Series Stack

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display for stack-identity, status, and hardware information on every unit in a stack. See the example below.</td>
<td>show system</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Display most of the information in show system, but in a more convenient tabular form. See the example below.</td>
<td>show system brief</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Display the same information in show system, but only for the specified unit. See the example below.</td>
<td>show system stack-unit</td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>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. See the example below.</td>
<td>show system stack-ports [status</td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

Display information about an S4810 stack using the show system command.

```
Force10#show system
Stack MAC : 00:01:e8:8a:df:e6
Reload Type : normal-reload
-- Unit 0 --
Unit Type : Management Unit
Status : online
Next Boot : online
Required Type : S4810 - 52-port GE/TE/FG (SE)
Current Type : S4810 - 52-port GE/TE/FG (SE)
Master priority : 0
Hardware Rev : 3.0
Num Ports : 64
Up Time : 57 min, 0 sec
FTOS Version : 8-3-7-13
Jumbo Capable : yes
POE Capable : no
```
Burned In MAC : 00:01:e8:8a:df:eb
No Of MACs : 3

-- Power Supplies --
<table>
<thead>
<tr>
<th>Unit</th>
<th>Bay</th>
<th>Status</th>
<th>Type</th>
<th>FanStatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>absent</td>
<td>absent</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>up</td>
<td>AC</td>
<td>up</td>
</tr>
</tbody>
</table>

-- Fan Status --
<table>
<thead>
<tr>
<th>Unit</th>
<th>Bay</th>
<th>TrayStatus</th>
<th>Fan0 Speed</th>
<th>Fan1 Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>up</td>
<td>6960 up</td>
<td>6960 up</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>up</td>
<td>6720 up</td>
<td>6720 up</td>
</tr>
</tbody>
</table>

Unit Type : Member Unit
Status : not present
Required Type : S4810 - 52-port GE/TE/FG (S)

-- Unit 2 --
Unit Type : Member Unit
Status : not present

-- Unit 3 --
Unit Type : Standby Unit
Status : online
Next Boot : online
Required Type : S4810 - 52-port GE/TE/FG (S)
Current Type : S4810 - 52-port GE/TE/FG (S)
Master priority : 0
Hardware Rev : 3.0
Num Ports : 64
Up Time : 57 min, 3 sec
FTOS Version : 8-3-7-13
Jumbo Capable : yes
POE Capable : no
Burned In MAC : 00:01:e8:8a:df:bf
No Of MACs : 3

-----output truncated-----

Display information about an S4810 stack using the `show system brief` command

FTOS#show system brief

Stack MAC : 00:01:e8:8a:de:48
Reload Type : normal-reload

-- Stack Info --
<table>
<thead>
<tr>
<th>Unit</th>
<th>UnitType</th>
<th>Status</th>
<th>ReqTyp</th>
<th>CurTyp</th>
<th>Version</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Management</td>
<td>online</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-12-13</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Display information about an S4810 stack using the `show system stack-ports` command.

```
FTOS#show system stack-ports
Topology: Ring
Interface   Connection    Link Speed (Gb/s)  Admin Status  Link Status  Trunk Group
---------------------------------------------------
0/56        3/56          40             up         up
0/60        3/60          40             up         up
3/48                      40             up         down
3/52                      40             up         down
3/56        0/56          40             up         up
3/60        0/60          40             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 master and standby management units. If multiple units tie for highest priority, the unit with the highest MAC address prevails.

If management was determined by priority only, a change in management occurs when:

- the management unit is powered down or a failover occurs
- you disconnect the management unit from the stack

When the management unit fails, the unit disappears from the stack topology. At that time, the standby unit detects the communication loss and switches from the standby unit role to the management unit role in the stack. From the remaining units in the stack, the system selects a new standby unit based on the unit priority using the same algorithm used when the stack was initially created. When the failed unit recovers, it takes the next available role, usually that of a stack member.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Influence the selection of the stack management units. The unit with the numerically highest priority is elected the master management unit, and the unit with the second highest priority is the standby unit. Default: 0 Range: 1-14</td>
<td><code>stack-unit priority</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
### Manage Redundancy on an S-Series Stack

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reset the current management unit, and make the standby unit the new master unit. A new standby is elected. When the former stack master comes back online, it becomes a member unit.</td>
<td><code>redundancy force-failover stack-unit</code></td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Prevent the stack master from rebooting after a failover. This command does not affect a forced failover, manual reset, or a stack-link disconnect.</td>
<td><code>redundancy disable-auto-reboot stack-unit</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>Display redundancy information.</td>
<td><code>show redundancy</code></td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

### Reset a Unit on an S-Series Stack

You may reset any stack unit except for the master management unit (Message 2).

**Message 2** Master Reset Disallowed

% Error: Reset of master unit is not allowed.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reload a stack-unit</td>
<td><code>reset stack-unit</code></td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Reload a member unit, from the unit itself</td>
<td><code>reset-self</code></td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>Reset a stack-unit when the unit is in a problem state.</td>
<td><code>reset stack-unit {hard}</code></td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>

### Verifying a Stack Configuration

#### LED Status Indicators on an S4810 Stack

The light of the LED status indicator on the front panel identifies the unit’s role in the stack.

- Off indicates the unit is a stack member.
- Blinking Green indicates the unit is the stack standby.
- Solid Green indicates the unit is the stack master (management unit)
Display Status of Stacking Ports

To display the status of the stacking ports, including the topology:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display the stacking</td>
<td><code>show system stack-ports</code></td>
<td>EXEC Privilege</td>
</tr>
<tr>
<td>ports.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The following example shows four switches stacked together with two 40G links in a ring topology.

```
FTOS#show system stack-ports
Topology: Ring
Interface  Connection    Link Speed (Gb/s)        Admin Status    Link Status    Trunk Group
------------------------------------------------------------------
1/48        7/56          40             up        up
1/52        5/60          40             up        up
2/56        6/52          40             up        up
2/60        4/52          40             up        up
3/48        7/52          40             up        up
3/52        5/56          40             up        up
4/52        6/48          40             up        up
4/56        4/48          40             up        up
FTOS#
```

The following example shows the parameters for the management unit in the stack.

```
FTOS#show system stack-unit 1
--  Unit 1 --
  Unit Type       : Management Unit
  Status          : online
  Next Boot       : online
  Required Type   : S4810 - 52-port GE/TE/FG (SE)
  Current Type    : S4810 - 52-port GE/TE/FG (SE)
  Master priority : 0
  Hardware Rev    : 3.0
  Num Ports       : 64
  Up Time         : 1 min, 14 sec
  FTOS Version    : 4810-8-3-12-1447
  Jumbo Capable   : yes
  POE Capable     : no
  Boot Flash      : 1.2.0.2
  Memory Size     : 2147483648 bytes
  Temperature     : 44C
  Voltage         : ok
  Serial Number   : H1DL104400018
  Part Number     : Rev
  Vendor Id       :
  Date Code       :
  Country Code    :
  Piece Part ID   : N/A
  PFXD Revision   : N/A
  Service Tag     : N/A
  Expr Svc Code   : N/A
  Auto Reboot     : disabled
  Burned In MAC   : 00:01:e8:8c:53:32
  No Of MACs      : 3
```
--- Power Supplies ---

<table>
<thead>
<tr>
<th>Unit</th>
<th>Bay</th>
<th>Status</th>
<th>Type</th>
<th>FanStatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>absent</td>
<td>absent</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>up</td>
<td>AC</td>
<td>up</td>
</tr>
</tbody>
</table>

--- Fan Status ---

<table>
<thead>
<tr>
<th>Unit</th>
<th>Bay</th>
<th>TrayStatus</th>
<th>Fan0</th>
<th>Speed</th>
<th>Fan1</th>
<th>Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>up</td>
<td>up</td>
<td>7200</td>
<td>up</td>
<td>7200</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>up</td>
<td>up</td>
<td>7200</td>
<td>up</td>
<td>7440</td>
</tr>
</tbody>
</table>

Speed in RPM

The following example shows three switches stacked together in a daisy chain topology.

```plaintext
stack-2#show system stack-ports
Topology: Daisy chain
Interface  Connection    Link Speed (Gb/s)  Admin Status  Link Status  Trunk Group
-----------------------------------------------
0/36        1/36          10             up           up           
0/37        1/37          10             up           up           
0/38        1/38          10             up           up           
0/39        1/39          10             up           up           
0/44        2/36          10             up           up           
0/45        2/37          10             up           up           
0/46        2/38          10             up           up           
0/47        2/39          10             up           up           
1/36        0/36          10             up           up           
1/37        0/37          10             up           up           
1/38        0/38          10             up           up           
1/39        0/39          10             up           up           
2/36        0/44          10             up           up           
2/37        0/45          10             up           up           
2/38        0/46          10             up           up           
2/39        0/47          10             up           up           
stack-2#`
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 daisy chain 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 48-10. Removing an S4810 Stack Member—Before

Force10#show system brief

Stack MAC : 00:01:e8:8a:df:e6
Reload Type : normal-reload

-- Stack Info --

<table>
<thead>
<tr>
<th>Unit</th>
<th>UnitType</th>
<th>Status</th>
<th>ReqTyp</th>
<th>CurTyp</th>
<th>Version</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Management</td>
<td>online</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-7-13</td>
<td>64</td>
</tr>
<tr>
<td>1</td>
<td>Member</td>
<td>online</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-7-13</td>
<td>64</td>
</tr>
<tr>
<td>2</td>
<td>Member</td>
<td>not present</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-7-13</td>
<td>64</td>
</tr>
<tr>
<td>3</td>
<td>Standby</td>
<td>online</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-7-13</td>
<td>64</td>
</tr>
</tbody>
</table>

Figure 48-11. Removing an S4810 Stack Member—After

Force10#show system brief

Stack MAC : 00:01:e8:8a:df:e6
Reload Type : normal-reload

-- Stack Info --

<table>
<thead>
<tr>
<th>Unit</th>
<th>UnitType</th>
<th>Status</th>
<th>ReqTyp</th>
<th>CurTyp</th>
<th>Version</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Management</td>
<td>online</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-7-13</td>
<td>64</td>
</tr>
<tr>
<td>1</td>
<td>Member</td>
<td>not present</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-7-13</td>
<td>64</td>
</tr>
<tr>
<td>2</td>
<td>Member</td>
<td>not present</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-7-13</td>
<td>64</td>
</tr>
<tr>
<td>3</td>
<td>Standby</td>
<td>online</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-7-13</td>
<td>64</td>
</tr>
</tbody>
</table>

Remove Front End Port Stacking

To remove the configuration on the front end ports used for stacking, use the following procedure.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remove the stack group configuration that are configured.</td>
<td><strong>no stack-unit id stack-group id</strong></td>
<td><strong>CONFIGURATION</strong></td>
</tr>
<tr>
<td>Save the stacking configuration on the ports.</td>
<td><strong>write memory</strong></td>
<td><strong>EXEC Privilege</strong></td>
</tr>
<tr>
<td>Reload the switch.</td>
<td><strong>reload</strong></td>
<td><strong>EXEC Privilege</strong></td>
</tr>
</tbody>
</table>
Troubleshoot an S-Series Stack

- Recover from Stack Link Flaps
- Recover from a Card Problem State on an S-Series Stack
- Recover from a Card Mismatch State on 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 (master) and standby management units as KERN-2-INT messages if the flapping port belongs to either of these units.

In the following example, a stack-port on the master flaps. The remote member, Member 2, displays a console message, and the master and standby display KERN-2-INT messages.

To re-enable the downed stack-port, power cycle the offending unit.

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 48-12. To recover, disconnect the new unit from the stack, change the FTOS version to match the stack, and then reconnect it to the stack, as shown in Figure 48-13.

Figure 48-12. Card Problem Error on an S-Series Stack - Different FTOS Versions

stack=1#show system brief
Stack MAC : 00:01:e8:8a:fd:6e
Reload Type : normal-reload [Next boot : normal-reload]

<table>
<thead>
<tr>
<th>Unit</th>
<th>UnitType</th>
<th>Status</th>
<th>ReqTyp</th>
<th>CurTyp</th>
<th>Version</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Standby</td>
<td>card problem</td>
<td>S4810</td>
<td>unknown</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>1</td>
<td>Management</td>
<td>online</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-10-223</td>
<td>64</td>
</tr>
<tr>
<td>2</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

-- Power Supplies --

<table>
<thead>
<tr>
<th>Unit</th>
<th>Bay</th>
<th>Status</th>
<th>Type</th>
<th>FanStatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>down</td>
<td>DC</td>
<td>down</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>up</td>
<td>DC</td>
<td>up</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>absent</td>
<td>absent</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>up</td>
<td>AC</td>
<td>up</td>
</tr>
</tbody>
</table>

-- Fan Status --

<table>
<thead>
<tr>
<th>Unit</th>
<th>Bay</th>
<th>TrayStatus</th>
<th>Fan0</th>
<th>Speed</th>
<th>Fan1</th>
<th>Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>up</td>
<td>up</td>
<td>9360</td>
<td>up</td>
<td>9360</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>up</td>
<td>up</td>
<td>9600</td>
<td>up</td>
<td>9360</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>up</td>
<td>up</td>
<td>6720</td>
<td>up</td>
<td>6720</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>up</td>
<td>up</td>
<td>6960</td>
<td>up</td>
<td>6720</td>
</tr>
</tbody>
</table>

Speed in RPM

**Figure 48-13. Card Problem Error on an S-Series Stack - Different FTOS Versions**

stack-1#show system brief

Stack MAC : 00:01:e8:8a:fd:6e

Reload Type : normal-reload [Next boot : normal-reload]

<table>
<thead>
<tr>
<th>Unit</th>
<th>UnitType</th>
<th>Status</th>
<th>ReqTyp</th>
<th>CurTyp</th>
<th>Version</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Standby</td>
<td>online</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-10-223</td>
<td>64</td>
</tr>
<tr>
<td>1</td>
<td>Management</td>
<td>online</td>
<td>S4810</td>
<td>S4810</td>
<td>8-3-10-223</td>
<td>64</td>
</tr>
<tr>
<td>2</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Member</td>
<td>not present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

-- Power Supplies --

<table>
<thead>
<tr>
<th>Unit</th>
<th>Bay</th>
<th>Status</th>
<th>Type</th>
<th>FanStatus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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. See the following example. 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 48-14. Recovering from a Card Mismatch State on an S-Series Stack (S50N and S25N)

Standalone#show system brief
Stack MAC : 00:01:e8:d5:e8:81
-- 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#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

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
Storm Control

Storm Control is supported on platforms: E C S Z 54810

Storm Control for Multicast is supported on platforms: C S Z

The storm control feature enables you to control unknown-unicast and broadcast traffic on Layer 2 and Layer 3 physical interfaces.

FTOS Behavior: On the E-Series, FTOS supports broadcast control for Layer 3 traffic only. To control Layer 2 broadcast traffic use the command storm-control unknown-unicast. On the C-Series and S-Series, FTOS supports broadcast control (command storm-control broadcast) for Layer 2 and Layer 3 traffic.

FTOS Behavior: 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. These ports can be in the same/different port-pipes or on the same/different line cards.

FTOS Behavior: The minimum number of packets per second (PPS) that storm control can limit on the S4810 is 2.

Configure Storm Control

Storm control is supported in INTERFACE mode and CONFIGURATION mode

Configure storm control from INTERFACE mode

Configure storm control from INTERFACE mode using the command storm control. From INTERFACE mode:

• You can only configure storm control for ingress traffic.
• If you configure storm control from both INTERFACE and CONFIGURATION mode, the INTERFACE mode configurations override the CONFIGURATION mode configurations.
• The percentage of storm control is calculated based on the advertised rate of the line card, not by the speed setting.
Configure storm control from CONFIGURATION mode

Configure storm control from CONFIGURATION mode using the command storm control. From CONFIGURATION mode you can configure storm control for ingress and egress traffic.

Do not apply per-VLAN QoS on an interface that has storm-control enabled (either on an interface or globally)

- On the E-Series, when broadcast storm-control is enabled on an interface or globally on the ingress and DSCP marking for a DSCP value 1 is configured for the data traffic, the traffic goes to queue 1 instead of queue 0. Similarly, if unicast storm-control is enabled on an interface or globally on the ingress, and DSCP marking for a DSCP value 2 is configured for the data traffic, the traffic goes to queue 2 instead of queue 0.
Spanning Tree Protocol (STP)

Spanning Tree Protocol (STP) is supported on platforms: E C S Z 54810

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>
<thead>
<tr>
<th>Dell Force10 Term</th>
<th>IEEE Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanning Tree Protocol (STP)</td>
<td>802.1d</td>
</tr>
<tr>
<td>Rapid Spanning Tree Protocol (RSTP)</td>
<td>802.1w</td>
</tr>
<tr>
<td>Multiple Spanning Tree Protocol (MSTP)</td>
<td>802.1s</td>
</tr>
<tr>
<td>Per-VLAN Spanning Tree Plus (PVST+)</td>
<td>Third Party</td>
</tr>
</tbody>
</table>

Configuring Spanning Tree

Configuring Spanning Tree is a two-step process:

1. Configuring Interfaces for Layer 2 Mode.

Related Configuration Tasks

- Adding an Interface to the Spanning Tree Group
• Removing an Interface from the Spanning Tree Group
• Modifying Global Parameters
• Modifying Interface STP Parameters
• Enabling PortFast
• Preventing Network Disruptions with BPDU Guard
• STP Root Selection
• SNMP Traps for Root Elections and Topology Changes
• Configuring Spanning Trees as Hitless

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.
To configure the interfaces for Layer 2 and then enable them:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>If the interface has been assigned an IP address, remove it.</td>
<td><code>no ip address</code></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>2</td>
<td>Place the interface in Layer 2 mode.</td>
<td><code>switchport</code></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>3</td>
<td>Enable the interface.</td>
<td><code>no shutdown</code></td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

Verify that an interface is in Layer 2 mode and enabled using the `show config` command from INTERFACE mode.

```bash
FTOS(conf-if-gi-1/1)#show config
!
interface GigabitEthernet 1/1
  no ip address
  switchport
  no shutdown
FTOS(conf-if-gi-1/1)#
```
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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter the PROTOCOL SPANNING TREE mode.</td>
<td>protocol spanning-tree 0</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Enable Spanning Tree.</td>
<td>no disable</td>
<td>PROTOCOL SPANNING TREE</td>
</tr>
</tbody>
</table>

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.

FTOS(conf)#protocol spanning-tree 0
FTOS(config-span)#show config
!
protocol spanning-tree 0
   no disable
FTOS#

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.
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.

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
  Designated root has priority 32768, address 0001.e80d.2462
  Designated bridge has priority 32768, address 0001.e80d.2462
  Designated port id is 8.497, 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
```

---More---

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
  Designated root has priority 32768, address 0001.e80d.2462
  Designated bridge has priority 32768, address 0001.e80d.2462
  Designated port id is 8.497, 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
```
Confirm that a port is participating in Spanning Tree using the `show spanning-tree 0 brief` command from EXEC privilege mode.

```text
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

<table>
<thead>
<tr>
<th>Interface</th>
<th>Name</th>
<th>PortID</th>
<th>Prio</th>
<th>Cost</th>
<th>Sts</th>
<th>Cost</th>
<th>Bridge ID</th>
<th>PortID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gi 1/1</td>
<td>8.496</td>
<td>8</td>
<td>4</td>
<td>DIS</td>
<td>0</td>
<td>32768</td>
<td>0001.e80d.2462</td>
<td>8.496</td>
</tr>
<tr>
<td>Gi 1/2</td>
<td>8.497</td>
<td>8</td>
<td>4</td>
<td>DIS</td>
<td>0</td>
<td>32768</td>
<td>0001.e80d.2462</td>
<td>8.497</td>
</tr>
<tr>
<td>Gi 1/3</td>
<td>8.513</td>
<td>8</td>
<td>4</td>
<td>FWD</td>
<td>0</td>
<td>32768</td>
<td>0001.e80d.2462</td>
<td>8.513</td>
</tr>
<tr>
<td>Gi 1/4</td>
<td>8.514</td>
<td>8</td>
<td>4</td>
<td>FWD</td>
<td>0</td>
<td>32768</td>
<td>0001.e80d.2462</td>
<td>8.514</td>
</tr>
</tbody>
</table>
```

**Adding an Interface to the Spanning Tree Group**

To add a Layer 2 interface to the Spanning Tree topology:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable Spanning Tree on a Layer 2 interface.</td>
<td>spanning-tree 0</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

**Removing an Interface from the Spanning Tree Group**

To remove a Layer 2 interface from the Spanning Tree topology:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disable Spanning Tree on a Layer 2 interface.</td>
<td>no spanning-tree 0</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

**FTOS Behavior:** 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 50-2 displays the default values for Spanning Tree.

### Table 50-2. STP Default Values

<table>
<thead>
<tr>
<th>STP Parameter</th>
<th>Default Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward Delay</td>
<td>15 seconds</td>
</tr>
<tr>
<td>Hello Time</td>
<td>2 seconds</td>
</tr>
<tr>
<td>Max Age</td>
<td>20 seconds</td>
</tr>
<tr>
<td>Port Cost</td>
<td></td>
</tr>
<tr>
<td>100-Mb/s Ethernet interfaces</td>
<td>19</td>
</tr>
<tr>
<td>1-Gigabit Ethernet interfaces</td>
<td>4</td>
</tr>
<tr>
<td>10-Gigabit Ethernet interfaces</td>
<td>2</td>
</tr>
<tr>
<td>Port Channel with 100 Mb/s Ethernet interfaces</td>
<td>18</td>
</tr>
<tr>
<td>Port Channel with 1-Gigabit Ethernet interfaces</td>
<td>3</td>
</tr>
<tr>
<td>Port Channel with 10-Gigabit Ethernet interfaces</td>
<td>1</td>
</tr>
<tr>
<td>Port Priority</td>
<td>8</td>
</tr>
</tbody>
</table>

To change STP global parameters:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
</table>
| 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. Refer to the second example in Enabling Spanning Tree Protocol Globally.

### 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 50-2.

To change the port cost or priority of an interface:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change the port cost of an interface.</td>
<td>spanning-tree 0 cost</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>
| Range: 0 to 65535  
Default: see Table 50-2. | | |
| Change the port priority of an interface. | spanning-tree 0 priority  
`priority-value` | INTERFACE |
| Range: 0 to 15  
Default: 8 | | |

View the current values for interface parameters using the `show spanning-tree 0` command from EXEC privilege mode. Refer to the second example in Enabling Spanning Tree Protocol Globally.

### Enabling PortFast

The PortFast feature enables interfaces to begin forwarding 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 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.

⚠️ **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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable PortFast on an interface</td>
<td>spanning-tree stp-id portfast [bpduguard</td>
<td>[shutdown-on-violation]]</td>
</tr>
</tbody>
</table>

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 50-3.

**Figure 50-3. PortFast Enabled on Interface**

```
FTOS#(conf-if-gi-1/1)#show conf
!
interface GigabitEthernet 1/1
 no ip address
 switchport
 spanning-tree 0 portfast
 no shutdown
 FTOS#(conf-if-gi-1/1)#
```

Indicates that the interface is in PortFast mode

### 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/Edgeport (edgeports) do not expect to receive BPDUs. If an edgeport 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 edgeport upon receiving a BPDU to prevent network disruptions, and FTOS displays **Message 1**. Enable BPDU Guard using the option `bpduguard` when enabling PortFast or EdgePort. The `bpduguard shutdown-on-violation` option causes the interface hardware to be shutdown when it receives a BPDU. Otherwise, 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 50-4** 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:** *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
```

```
<table>
<thead>
<tr>
<th>Interface</th>
<th>Name</th>
<th>PortID</th>
<th>Prio</th>
<th>Cost</th>
<th>Sts</th>
<th>Bridge ID</th>
<th>PortID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gi 0/6</td>
<td>128.263</td>
<td>128</td>
<td>20000</td>
<td>FWD</td>
<td>20000</td>
<td>32768</td>
<td>0001.e805.fb07 128.653</td>
</tr>
<tr>
<td>Gi 0/7</td>
<td>128.264</td>
<td>128</td>
<td>20000</td>
<td>EDS</td>
<td>20000</td>
<td>32768</td>
<td>0001.e85d.0e90 128.264</td>
</tr>
</tbody>
</table>
```

```
<table>
<thead>
<tr>
<th>Interface</th>
<th>Name</th>
<th>Role</th>
<th>PortID</th>
<th>Prio</th>
<th>Cost</th>
<th>Link-type</th>
<th>Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gi 0/6</td>
<td>Root</td>
<td></td>
<td>128.263</td>
<td>128</td>
<td>20000</td>
<td>P2P</td>
<td>No</td>
</tr>
<tr>
<td>Gi 0/7</td>
<td>ErrDis</td>
<td></td>
<td>128.264</td>
<td>128</td>
<td>20000</td>
<td>P2P</td>
<td>No</td>
</tr>
</tbody>
</table>
```

```
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).
The 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. You can also specify that a bridge is the root or the secondary root.

To change the bridge priority or specify that a bridge is the root or secondary root:

**FTOS Behavior:** BPDU Guard and BPDU filtering (refer to Removing an Interface from the Spanning Tree Group) 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 Guard

**STP Root Guard** is supported only on platforms: 

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
</table>
| 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 50-5) from EXEC privilege mode.

**Figure 50-5. 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
```

**Root Guard Scenario**

For example, in Figure 50-6 (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 50-6 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 BPU 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 50-6 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 50-6. STP Root Guard Prevents Bridging Loops**

---

Port State:
- **STP Block**
- **STP Root-Inconsistent**
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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
<td>`spanning-tree {0</td>
<td>mstp</td>
</tr>
</tbody>
</table>

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 RSTP, MSTP, and PVST+ collectively 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** **S** **4810**

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 LACP to be hitless using the command `redundancy protocol lacp`. Configure all Spanning Tree types to be hitless using the command `redundancy protocol xstp` from CONFIGURATION mode, as shown in Figure 50-7.

**Figure 50-7. 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** **4810**

**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 50-8 (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 50-8 (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 50-8 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 50-8. STP Loop Guard Prevents Forwarding Loops

Port State:
- **X**: STP Block
- **1 way Link Failure**
- **X**: STP Loop-Inconsistent
- **Green circle**: No traffic is transmitted
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).
- 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
<td>spanning-tree {0</td>
<td>mstp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INTERFACE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PORT-CHANNEL</td>
</tr>
</tbody>
</table>

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.

The example below shows 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.

```
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
```
System Time and Date

System Time and Date settings and Network Time Protocol are supported on platforms: E C S Z S4810

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
  • Protocol Overview
  • Implementation Information
  • Configuring Network Time Protocol
• FTOS Time and Date
  • Configuring time and date settings
  • Set daylight saving time

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.
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 1026.

Related Configuration Tasks

- Configure NTP broadcasts on page 1027
- Set the Hardware Clock with the Time Derived from NTP on page 1026
- Set the Hardware Clock with the Time Derived from NTP on page 1026
- Disable NTP on an interface on page 1027
- Configure a source IP address for NTP packets on page 1027 (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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify the NTP server to which the Dell Force10 system will</td>
<td>ntp server ip-address</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>synchronize.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Display the system clock state with respect to NTP using the command show ntp status from EXEC Privilege mode, as shown in Figure 51-2.

**Figure 51-2. Displaying the System Clock State with respect to NTP**

```
R6_E300(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 CD63BCC2.0CBBDD000 (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 51-3.

**Figure 51-3. 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

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periodically update the system hardware clock with the time value</td>
<td>ntp update-calendar</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>derived from NTP.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set the interface to receive NTP packets.</td>
<td>ntp broadcast client</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

Table 51-1.

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ntp disable</td>
<td>INTERFACE</td>
<td>Disable NTP on the interface.</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>ntp source interface</td>
<td>CONFIGURATION</td>
<td>Enter the following keywords and slot/port or number information:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 1-Gigabit Ethernet interface, enter the keyword GigabitEthernet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a loopback interface, enter the keyword loopback followed by a number between 0 and 16383.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a SONET interface, enter the keyword sonet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 10-Gigabit Ethernet interface, enter the keyword TenGigabitEthernet followed by the slot/port information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a VLAN interface, enter the keyword vlan followed by a number from 1 to 4094.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• For a 40-Gigabit Ethernet interface, enter the keyword fortyGigE followed by the slot/port information.</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ntp authenticate</td>
<td>CONFIGURATION</td>
<td>Enable NTP authentication.</td>
</tr>
</tbody>
</table>
To view the NTP configuration, use the `show running-config ntp` command (Figure 40) in the EXEC privilege mode. Figure 51-5 shows an encrypted authentication key. All keys are encrypted.

**Figure 51-5. show running-config ntp Command Example**

```
FTOS#show running ntp
1
 ntp authenticate
 ntp authentication-key 345 md5 5A60910F3D211F02  encrypted key
 ntp server 11.1.1.1 version 3
 ntp trusted-key 345
FTOS#
```
<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| 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. |

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: 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.

• **Origininate 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 saving time
  - Set Daylight Saving Time Once
  - Set Recurring Daylight Saving Time
Set the time and date for the switch hardware clock

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>calendar set time month day year</td>
<td>EXEC Privilege</td>
<td>Set the hardware clock to the current time and date.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>time</strong>: 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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>month</strong>: 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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>day</strong>: 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</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>year</strong>: Enter a four-digit number as the year. Range: 1993 to 2035.</td>
</tr>
</tbody>
</table>

```
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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>clock set time month day year</td>
<td>EXEC Privilege</td>
<td>Set the system software clock to the current time and date.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>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</td>
</tr>
<tr>
<td></td>
<td></td>
<td>year: Enter a four-digit number as the year. Range: 1993 to 2035.</td>
</tr>
</tbody>
</table>

FTOS#clock set 16:20:00 19 september 2009

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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>clock timezone timezone-name offset</td>
<td>CONFIGURATION</td>
<td>Set the clock to the appropriate timezone.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>timezone-name: Enter the name of the timezone. Do not use spaces.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>offset: Enter one of the following:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• a number from 1 to 23 as the number of hours in addition to UTC for the timezone.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• a minus sign (-) followed by a number from 1 to 23 as the number of hours</td>
</tr>
</tbody>
</table>

FTOS#
<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTOS#conf</td>
<td>FTOS(conf)</td>
<td>set daylight saving time time zone</td>
</tr>
<tr>
<td>FTOS(conf)#clock timezone Pacific -8</td>
<td></td>
<td>set daylight saving time time zone</td>
</tr>
<tr>
<td>FTOS(conf)#01:40:19: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Timezone configuration changed from &quot;UTC 0 hrs 0 mins&quot; to &quot;Pacific -8 hrs 0 mins&quot;</td>
<td>FTOS#</td>
<td></td>
</tr>
</tbody>
</table>

**Set daylight saving time**

FTOS supports setting the system to daylight saving 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 saving time on a one-time basis.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>clock summer-time</td>
<td>CONFIGURATION</td>
<td>Set the clock to the appropriate timezone and daylight saving time.</td>
</tr>
<tr>
<td>time-zone date</td>
<td></td>
<td>time-zone: Enter the three-letter name for the time zone. This name is</td>
</tr>
<tr>
<td>start-month start-day</td>
<td></td>
<td>displayed in the show clock output.</td>
</tr>
<tr>
<td>start-year start-time</td>
<td></td>
<td>start-month: Enter the name of one of the 12 months in English. You</td>
</tr>
<tr>
<td>end-month end-day</td>
<td></td>
<td>can enter the name of a day to change the order of the display to</td>
</tr>
<tr>
<td></td>
<td></td>
<td>time day month year.</td>
</tr>
<tr>
<td>end-time</td>
<td></td>
<td>start-day: Enter the number of the day. Range: 1 to 31.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>You can enter the name of a month to change the order of the display to time day month year.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>start-year: Enter a four-digit number as the year. Range: 1993 to 2035</td>
</tr>
<tr>
<td></td>
<td></td>
<td>start-time: Enter the time in hours:minutes. For the hour variable, use the 24-hour format, example, 17:15 is 5:15 pm.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>end-day: Enter the number of the day. Range: 1 to 31.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>You can enter the name of a month to change the order of the display to time day month year.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>end-year: Enter a four-digit number as the year. Range: 1993 to 2035.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>end-time: Enter the time in hours:minutes. For the hour variable, use the 24-hour format, example, 17:15 is 5:15 pm.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>offset: (OPTIONAL) Enter the number of minutes to add during the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>summer-time period. Range: 1 to1440. Default: 60 minutes</td>
</tr>
</tbody>
</table>

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 saving time on a specific day every year.

If you have already set daylight saving 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.

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>clock summer-time</td>
<td>CONFIGURATION</td>
<td>Set the clock to the appropriate timezone and adjust to daylight saving time every year.</td>
</tr>
<tr>
<td>time-zone recurring</td>
<td></td>
<td>time-zone: Enter the three-letter name for the time zone. This name is displayed in the show clock output.</td>
</tr>
<tr>
<td>start-week start-day</td>
<td></td>
<td>start-week: (OPTIONAL) Enter one of the following as the week that daylight saving begins and then enter values for start-day through end-time:</td>
</tr>
<tr>
<td>start-month start-time</td>
<td></td>
<td>• week-number: Enter a number from 1-4 as the number of the week in the month to start daylight saving time.</td>
</tr>
<tr>
<td>end-week end-day</td>
<td></td>
<td>• first: Enter this keyword to start daylight saving time in the first week of the month.</td>
</tr>
<tr>
<td>end-month end-time</td>
<td></td>
<td>• last: Enter this keyword to start daylight saving time in the last week of the month.</td>
</tr>
<tr>
<td>[offset]</td>
<td></td>
<td>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.</td>
</tr>
</tbody>
</table>
**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 saving ends:  
- **week-number:** enter a number from 1-4 as the number of the week to end daylight saving time.  
- **first:** enter the keyword first to end daylight saving time in the first week of the month.  
- **last:** enter the keyword last to end daylight saving 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 to 1440.  
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

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTOS(conf)#clock summer-time pacific recurring ?</td>
<td>&lt;1-4&gt;</td>
<td>Week number to start</td>
</tr>
<tr>
<td></td>
<td>first</td>
<td>Week number to start</td>
</tr>
<tr>
<td></td>
<td>last</td>
<td>Week number to start</td>
</tr>
<tr>
<td>&lt;cr&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTOS(conf)#clock summer-time pacific recurring</td>
<td>FTOS(conf)#02:10:57: %RPM0-P:CP %CLOCK-6-TIME CHANGE: Summertime configuration changed from &quot;Summer time starts 00:00:00 Pacific Sat Mar 14 2009 ; Summer time ends 00:00:00 pacific Sat Nov 7 2009&quot; to &quot;Summer time starts 02:00:00 Pacific Sun Mar 8 2009;Summer time ends 02:00:00 pacific Sun Nov 1 2009&quot;</td>
<td></td>
</tr>
</tbody>
</table>

Uplink Failure Detection (UFD) is supported on the following platforms: S (S50 only), MXL, S54810

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 52-2 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.
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 52-2.
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 626) to implement a rapid failover solution. For example, in Figure 52-2 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:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax and Mode</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>uplink-state-group group-id</td>
<td>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.</td>
</tr>
<tr>
<td>2</td>
<td>{upstream</td>
<td>downstream} interface</td>
</tr>
<tr>
<td>3</td>
<td>downstream disable links {number</td>
<td>all}</td>
</tr>
<tr>
<td>4</td>
<td>downstream auto-recover</td>
<td>(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.</td>
</tr>
</tbody>
</table>
### Clearing a UFD-Disabled Interface

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:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Description</th>
</tr>
</thead>
</table>
| clear ufd-disable {interface interface | 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.  
**clear ufd-disable {interface interface | uplink-state-group group-id}** re-enables all UFD-disabled downstream interfaces in the group.  
Range: 1 to 16.  

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax and Mode</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td><strong>description</strong> text</td>
<td>(Optional) Enters a text description of the uplink-state group. Maximum length: 80 alphanumeric characters.</td>
</tr>
<tr>
<td></td>
<td>Command Mode: UPLINK-STATE-GROUP</td>
<td></td>
</tr>
</tbody>
</table>
| 6    | **no enable**            | (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. |
|      | Command Mode: UPLINK-STATE-GROUP |             |
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 (S50)

```
02:36:43: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Te 0/46
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 0/46
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Fo 13/0
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Fo 13/1
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Fo 13/3
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Fo 13/5
02:36:43: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Downstream interface set to UFD error-disabled: Fo 13/6
02:37:29: %RPM0-P:CP %IFMGR-5-ASTATE_DN: Changed interface Admin state to down: Te 0/47
02:37:29: %RPM0-P:CP %IFMGR-5-OSTATE_DN: Changed interface state to down: Te 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: Fo 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: Fo 13/0
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD error-disabled: Fo 13/1
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD error-disabled: Fo 13/3
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD error-disabled: Fo 13/5
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Downstream interface cleared from UFD error-disabled: Fo 13/6
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Fo 13/0
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Fo 13/1
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Fo 13/3
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Fo 13/5
02:38:53: %RPM0-P:CP %IFMGR-5-OSTATE_UP: Changed interface state to up: Fo 13/6
```
# Displaying Uplink Failure Detection

To display information on the Uplink Failure Detection feature, enter any of the following `show` commands:

<table>
<thead>
<tr>
<th>Show Command Syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>show uplink-state-group [group-id] [detail]</strong></td>
<td>Displays status information on a specified uplink-state group or all groups. Valid <code>group-id</code> values are 1 to 16. <strong>detail</strong> displays additional status information on the upstream and downstream interfaces in each group (see Figure 52-3).</td>
</tr>
<tr>
<td><strong>show interfaces interface</strong></td>
<td>Displays the current status of a port or port-channel interface assigned to an uplink-state group. <strong>interface</strong> specifies one of the following interface types: Fast Ethernet: Enter <code>fastethernet slot/port</code>. 1-Gigabit Ethernet: Enter <code>gigabitethernet slot/port</code>. 10-Gigabit Ethernet: Enter <code>tengigabitethernet slot/port</code>. Port channel: Enter <code>port-channel {1-512}</code>. 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 52-4).</td>
</tr>
<tr>
<td><strong>show running-config uplink-state-group [group-id]</strong></td>
<td>Displays the current configuration of all uplink-state groups (Figure 52-5) or a specified group (Figure 52-6). Valid <code>group-id</code> values are 1 to 16.</td>
</tr>
<tr>
<td>Or <strong>show configuration</strong></td>
<td>Command Mode: UPLINK-STATE-GROUP</td>
</tr>
</tbody>
</table>

---
Figure 52-3. show uplink-state-group Command Output (S50)

```
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 52-4.  show interfaces Command: UFD Output (S50)

```
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 Mb/s, 0 packets/sec, 0.00% of line-rate
    Output 00.00 Mb/s, 0 packets/sec, 0.00% of line-rate
Time since last interface status change: 00:01:23
```

Figure 52-5.  show running-config uplink-state-group Command: UFD Output (S50)

```
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 52-6.  show configuration Command: UFD Output (S50)

```
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 52-7 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 52-7. Configuring Uplink Failure Detection (S50)

```
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#
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)
```
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 the Dell Force10 Technical Support Center. You can reach Technical Support:

- On the Web: www.force10networks.com/support/
- By email: support@force10networks.com
Virtual LANs (VLAN)

Virtual LANs (VLAN) are supported on platforms: E C S Z 54810

This section contains the following subsections:

- Default VLAN
- Port-Based VLANs
- VLANs and Port Tagging
- Configuration Task List for VLANs
- Enable Null VLAN as the Default VLAN

Virtual LANs, or VLANs, are a logical broadcast domain or logical grouping of interfaces in a LAN in which all data received is kept locally and broadcast to all members of the group. When in Layer 2 mode, VLANs move traffic at wire speed and can span multiple devices. FTOS supports up to 4093 port-based VLANs and 1 Default VLAN, as specified in IEEE 802.1Q.

VLANs provide the following benefits:

- Improved security because you can isolate groups of users into different VLANs
- Ability to create one VLAN across multiple devices

For more information on VLANs, refer to IEEE Standard 802.1Q Virtual Bridged Local Area Networks. In this guide, see also:

- Bulk Configuration in Chapter 23, Interfaces
- VLAN Stacking

For a complete listing of all commands related to FTOS VLANs, see these *FTOS Command Reference* chapters:

- *Interfaces* chapter
- Port Authentication (802.1x) section in the *Security* chapter
- Chapter 20, GARP VLAN Registration Protocol (GVRP).
- Chapter 45, Service Provider Bridging
- Chapter 38, Per-VLAN Spanning Tree Plus (PVST+).
- For E-Series, see also the *ACL VLAN Group and Force10 Resilient Ring Protocol* chapters.
Table 54-1 displays the defaults for VLANs in FTOS.

### Table 54-1. VLAN Defaults on FTOS

<table>
<thead>
<tr>
<th>Feature</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanning Tree group ID</td>
<td>All VLANs are part of Spanning Tree group 0</td>
</tr>
<tr>
<td>Mode</td>
<td>Layer 2 (no IP address is assigned)</td>
</tr>
<tr>
<td>Default VLAN ID</td>
<td>VLAN 1</td>
</tr>
</tbody>
</table>

### Default VLAN

When interfaces are configured for Layer 2 mode, they are automatically placed in the Default VLAN as untagged interfaces. Only untagged interfaces can belong to the Default VLAN.

Figure 54-1 displays the outcome of placing an interface in Layer 2 mode. To configure an interface for Layer 2 mode, use the `switchport` command. In Step 1, the `switchport` command places the interface in Layer 2 mode.

In Step 2, the `show vlan` command in EXEC privilege mode indicates that the interface is now part of the Default VLAN (VLAN 1).

#### Figure 54-1. Interfaces and the Default VLAN Example

```
FTOS(conf)#int gi 3/2
FTOS(conf-if)#no shut
FTOS(conf-if)#switchport
FTOS(conf-if)#show config
!
interface GigabitEthernet 3/2
   no ip address
   switchport
   no shutdown
FTOS(conf-if)#end
FTOS#show vlan
Codes: * - Default VLAN, G - GVRP VLANs

    NUM    Status  Q Ports
    * 1      Active  U Gi 3/2
    2      Active  T Po1(So 0/0-1) T Gi 3/0
FTOS#
```

By default, VLAN 1 is the Default VLAN. To change that designation, use the `default vlan-id` command in the CONFIGURATION mode. You cannot delete the Default VLAN.

**Note:** An IP address cannot be assigned to the Default VLAN. To assign an IP address to a VLAN that is currently the Default VLAN, create another VLAN and assign it to be the Default VLAN. For details on assigning IP addresses, see Assign an IP address to a VLAN.
Untagged interfaces must be part of a VLAN. To remove an untagged interface from the Default VLAN, you must create another VLAN and place the interface into that VLAN. Alternatively, enter the no switchport command, and FTOS removes the interface from the Default VLAN.

A tagged interface requires an additional step to remove it from Layer 2 mode. Since tagged interfaces can belong to multiple VLANs, you must remove the tagged interface from all VLANs, using the no tagged interface command. Only after the interface is untagged and a member of the Default VLAN can you use the no switchport command to remove the interface from Layer 2 mode. For more information, see VLANs and Port Tagging.

**Port-Based VLANs**

Port-based VLANs are a broadcast domain defined by different ports or interfaces. In FTOS, a port-based VLAN can contain interfaces from different line cards within the chassis. FTOS supports 4094 port-based VLANs.

**Note:** E-Series ExaScale platforms support 4094 VLANs with FTOS version 8.2.1.0 and later. Earlier ExaScale supports 2094 VLANS.

Port-based VLANs offer increased security for traffic, conserve bandwidth, and allow switch segmentation. Interfaces in different VLANs do not communicate with each other, adding some security to the traffic on those interfaces. Different VLANs can communicate between each other by means of IP routing. Because traffic is only broadcast or flooded to the interfaces within a VLAN, the VLAN conserves bandwidth. Finally, you can have multiple VLANs configured on one switch, thus segmenting the device.

Interfaces within a port-based VLAN must be in Layer 2 mode and can be tagged or untagged in the VLAN ID.

**VLANs and Port Tagging**

To add an interface to a VLAN, it must be in Layer 2 mode. After you place an interface in Layer 2 mode, it is automatically placed in the Default VLAN. FTOS supports IEEE 802.1Q tagging at the interface level to filter traffic. When tagging is enabled, a tag header is added to the frame after the destination and source MAC addresses. That information is preserved as the frame moves through the network. Figure 54-2 illustrates the structure of a frame with a tag header. The VLAN ID is inserted in the tag header.

**Figure 54-2. Tagged Frame Format**

<table>
<thead>
<tr>
<th>Ethernet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preamble</td>
</tr>
<tr>
<td>6 octets</td>
</tr>
</tbody>
</table>

The tag header contains some key information used by FTOS:
The VLAN protocol identifier identifies the frame as tagged according to the IEEE 802.1Q specifications (2 bytes).

Tag Control Information (TCI) includes the VLAN ID (2 bytes total). The VLAN ID can have 4,096 values, but 2 are reserved.

**Note:** The insertion of the tag header into the Ethernet frame increases the size of the frame to more than the 1518 bytes specified in the IEEE 802.3 standard. Some devices that are not compliant with IEEE 802.3 may not support the larger frame size.

Information contained in the tag header allows the system to prioritize traffic and to forward information to ports associated with a specific VLAN ID. Tagged interfaces can belong to multiple VLANs, while untagged interfaces can belong only to one VLAN.

### Configuration Task List for VLANs

This section contains the following VLAN configuration tasks:

- Create a port-based VLAN (mandatory)
- Assign interfaces to a VLAN (optional)
- Assign an IP address to a VLAN (optional)
- Enable Null VLAN as the Default VLAN

### Create a port-based VLAN

The Default VLAN as VLAN 1 is part of the system startup configuration and does not require configuration. To configure a port-based VLAN, you must create the VLAN and then add physical interfaces or port channel (LAG) interfaces to the VLAN.

To create a port-based VLAN, use the following command in the CONFIGURATION mode:

<table>
<thead>
<tr>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>interface vlan *vlan-id *</td>
<td>CONFIGURATION</td>
<td>Configure a port-based VLAN (if the vlan-id is different from the Default VLAN ID) and enter INTERFACE VLAN mode. After you create a VLAN, you must assign interfaces in Layer 2 mode to the VLAN to activate the VLAN.</td>
</tr>
</tbody>
</table>
Use the `show vlan` command (Figure 54-3) in the EXEC privilege mode to view the configured VLANs.

**Figure 54-3. show vlan Command Example**

![Show VLAN Command Example](image)

A VLAN is active only if the VLAN contains interfaces and those interfaces are operationally up. In Figure 54-3, VLAN 1 is inactive because it contains the interfaces that are not active. The other VLANs listed in the Figure 54-3 contain enabled interfaces and are active.

**Note:** In a VLAN, the `shutdown` command stops Layer 3 (routed) traffic only. Layer 2 traffic continues to pass through the VLAN. If the VLAN is not a routed VLAN (that is, configured with an IP address), the `shutdown` command has no affect on VLAN traffic.

When you delete a VLAN (using the `no interface vlan vlan-id` command), any interfaces assigned to that VLAN are assigned to the Default VLAN as untagged interfaces.

**Assign interfaces to a VLAN**

Only interfaces in Layer 2 mode can be assigned to a VLAN using the tagged and untagged commands. Use the `switchport` command to place an interface in Layer 2 mode.

These Layer 2 interfaces can further be designated as tagged or untagged. For more information, refer to the Interfaces chapter and Configure Layer 2 (Data Link) Mode. When an interface is placed in Layer 2 mode by the `switchport` command, the interface is automatically designated untagged and placed in the Default VLAN.

To view which interfaces are tagged or untagged and to which VLAN they belong, use the `show vlan` command. For example, Figure 54-3 shows that six VLANs are configured, and two interfaces are assigned to VLAN 2. The Q column in the `show vlan` command example notes whether the interface is tagged (T) or untagged (U). For more information on this command, see the command statement in the Layer 2 chapter of the *FTOS Command Reference*.

To view just the interfaces that are in Layer 2 mode, enter the `show interfaces switchport` command in the EXEC privilege mode or EXEC mode.
To tag frames leaving an interface in Layer 2 mode, you must assign that interface to a port-based VLAN to tag it with that VLAN ID. To tag interfaces, use these commands in the following sequence:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interface vlan vlan-id</td>
<td>CONFIGURATION</td>
<td>Access the INTERFACE VLAN mode of the VLAN to which you want to assign the interface.</td>
</tr>
<tr>
<td>2</td>
<td>tagged interface</td>
<td>INTERFACE</td>
<td>Enable an interface to include the IEEE 802.1Q tag header.</td>
</tr>
</tbody>
</table>

Figure 54-4 shows the steps to add a tagged interface (in this case, port channel 1) to VLAN 4.

**Figure 54-4. Example of Adding an Interface to Another VLAN**

<table>
<thead>
<tr>
<th>FTOS#show vlan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Codes: * - Default VLAN, G - GVRP VLANs</td>
</tr>
<tr>
<td>NUM</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>*</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

FTOS#config
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
FTOS(conf-if-vlan)#end
FTOS#show vlan

Codes: * - Default VLAN, G - GVRP VLANs

<table>
<thead>
<tr>
<th>NUM</th>
<th>Status</th>
<th>Q Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>Inactive</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Active</td>
<td>T Po1(So 0/0-1)</td>
</tr>
<tr>
<td>3</td>
<td>Active</td>
<td>T Po1(So 0/0-1)</td>
</tr>
<tr>
<td>4</td>
<td>Active</td>
<td>T Po1(So 0/0-1)</td>
</tr>
</tbody>
</table>

Except for hybrid ports, only a tagged interface can be a member of multiple VLANs. Hybrid ports can be assigned to two VLANs if the port is untagged in one VLAN and tagged in all others.

When you remove a tagged interface from a VLAN (using the no tagged interface command), it will remain tagged only if it is a tagged interface in another VLAN. If the tagged interface is removed from the only VLAN to which it belongs, the interface is placed in the Default VLAN as an untagged interface.
Use the untagged command to move untagged interfaces from the Default VLAN to another VLAN:

<table>
<thead>
<tr>
<th>Step</th>
<th>Command Syntax</th>
<th>Command Mode</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interface vlan vlan-id</td>
<td>CONFIGURATION</td>
<td>Access the INTERFACE VLAN mode of the VLAN to which you want to assign the interface.</td>
</tr>
<tr>
<td>2</td>
<td>untagged interface</td>
<td>INTERFACE</td>
<td>Configure an interface as untagged. This command is available only in VLAN interfaces.</td>
</tr>
</tbody>
</table>

The no untagged interface command removes the untagged interface from a port-based VLAN and places the interface in the Default VLAN. You cannot use the no untagged interface command in the Default VLAN. Figure 54-5 illustrates the steps and commands to move an untagged interface from the Default VLAN to another VLAN.

**Figure 54-5. Example of Moving an Untagged Interface to Another VLAN**

The only way to remove an interface from the Default VLAN is to place the interface in Default mode by entering the no switchport command in the INTERFACE mode.
Assign an IP address to a VLAN

VLANs are a Layer 2 feature. For two physical interfaces on different VLANs to communicate, you must assign an IP address to the VLANs to route traffic between the two interfaces.

The shutdown command in INTERFACE mode does not affect Layer 2 traffic on the interface; the shutdown command only prevents Layer 3 traffic from traversing over the interface.

**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.

To assign an IP address, use the following command in INTERFACE mode:

```
ip address ip-address mask [secondary]
```

**Purpose**
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** — This is the interface’s backup IP address. You can configure up to eight secondary IP addresses.

In FTOS, VLANs and other logical interfaces can be placed in Layer 3 mode to receive and send routed traffic. For details, see Bulk Configuration.

**VLAN Interface Counters**

VLAN counters can be enabled for either Ingress packets, egress packets, or both. VLAN counters are disabled by default, and are supported on E-Series ExaScale only.

```
enable vlan-counter [ingress | egress | all]
```

**Purpose**
Configure ingress, egress or both counters for VLAN interfaces.

To return to the default without any VLAN counters, enter no enable vlan-counter.

**Note:** VLAN output counters may show higher than expected values because source-suppression drops are counted.
Native VLANs

Traditionally, ports can be either untagged for membership to one VLAN or tagged for membership to multiple VLANs. An untagged port must be connected to a VLAN-unaware station (one that does not understand VLAN tags), and a tagged port must be connected to a VLAN-aware station (one that generates and understands VLAN tags).

Native VLAN support breaks this barrier so that a port can be connected to both VLAN-aware and VLAN-unaware stations. Such ports are referred to as hybrid ports. Physical and port-channel interfaces may be hybrid ports.

Native VLAN is useful in deployments where a Layer 2 port can receive both tagged and untagged traffic on the same physical port. The classic example is connecting a VOIP phone and a PC to the same port of the switch. The VOIP phone is configured to generate tagged packets (with VLAN = VOICE VLAN), and the attached PC generates untagged packets.

To configure a port so that it can be a member of an untagged and tagged VLANs:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Remove any Layer 2 or Layer 3 configurations from the interface.</td>
<td></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>2</td>
<td>Configure the interface for hybrid mode.</td>
<td>portmode hybrid</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>3</td>
<td>Configure the interface for switchport mode.</td>
<td>switchport</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>4</td>
<td>Add the interface to a tagged or untagged VLAN.</td>
<td>[tagged</td>
<td>untagged]</td>
</tr>
</tbody>
</table>

**Note:** An existing switchport or port channel interface cannot be configured for Native VLAN. Interfaces must have no other Layer 2 or Layer 3 configurations when entering the command portmode hybrid or a message like Message 1 is displayed.

**Message 1** Native VLAN Error

% Error: Port is in Layer-2 mode Gi 5/6.
Enable 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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disable the default VLAN, so that all ports belong to the Null VLAN until configured as a member of another VLAN.</td>
<td>default-vlan disable Default: the default VLAN is enabled (no default-vlan disable).</td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>
Virtual Link Trunking (VLT)

Virtual Link Trunking (VLT) is supported on platforms Z54810.

Overview

Virtual link trunking (VLT) allows physical links between two chassis to appear as a single virtual link to the network core. VLT reduces the role of Spanning Tree protocols by allowing LAG terminations on two separate distribution or core switches, and by supporting a loop free topology. (A Spanning Tree protocol is still needed to prevent the initial loop that may occur prior to VLT being established. After VLT is established, RSTP may be used to prevent loops from forming with new links that are incorrectly connected and outside the VLT domain.) VLT provides Layer 2 multipathing, creating redundancy through increased bandwidth, enabling multiple parallel paths between nodes and load-balancing traffic where alternative paths exist.

As shown in Figure 55-1, VLT presents a single logical Layer 2 domain from the perspective of attached devices that have a virtual link trunk terminating on separate chassis in the VLT domain. However, the two VLT chassis are independent L2/L3 switches for devices in the upstream network. L2/L3 control plane protocols and system management features function normally in VLT mode. Features such as VRRP and IGMP Snooping require state information to be coordinated between the two VLT chassis.
Figure 55-1. Virtual Link Trunking

VLT peer devices have independent management planes. A chassis interconnect trunk between the VLT chassis maintains synchronization of L2/L3 control planes across the two VLT peers. The chassis interconnect trunk uses 10GE or 40GE user ports on the chassis.

A separate backup link maintains heartbeat messages across an out-of-band management network. The backup link ensures that node failure conditions are correctly detected and are not confused with failures of the chassis interconnect trunk. VLT ensures that local traffic on a chassis does not traverse the chassis interconnect trunk and takes the shortest path to the destination via directly attached links.

Virtual link trunking offers the following benefits:

- Allows a single device to use a LAG across two upstream devices
- Provides a loop-free topology
- Uses all available uplink bandwidth
- Provides fast convergence if either the link or a device fails
- Optimized forwarding with VRRP
- Provides link-level resiliency
- Assures high availability
Enhanced VLT

An enhanced VLT (eVLT) configuration creates a port channel between two VLT domains by allowing two different VLT domains, using different VLT Domain ID numbers, connected by a standard LACP LAG to form a loop-free Layer 2 topology in the aggregation layer. This configuration supports a maximum of four (4) nodes per eVLT domain, increasing the number of available ports and allowing for dual redundancy of the VLT.

Additionally, a VLT domain that is a member of one eVLT can be used in another eVLT configuration with a different VLT domain. Routing protocols such as OSPF are not compatible with eVLT; however, VLT domains can be used for routing. A separate Layer 3 router is not required for inter-VLAN communication.

The following figure shows how the core/aggregation port density in the Layer 2 topology is increased using eVLT. For inter-VLAN routing and other Layer 3 routing, a separate Layer 3 router is required.

VLT Concepts Terminology

The following are some key VLT terms.

**Virtual link trunk (VLT)** - The combined port channel between an attached device and the VLT peer switches.

**VLT backup link** - The backup link monitors the vitality of a VLT peer switch. The backup link sends configurable, periodic keep alive messages between VLT peer switches.

**VLT interconnect (VLTI)** - The link used to synchronize states between the VLT peer switches. Both ends must be on 10Gor 40G interfaces.
VLT domain - This domain includes both VLT peer devices, the VLT interconnect, and all of the port channels in the VLT connected to the attached devices. It is also associated to the configuration mode that must be used to assign VLT global parameters.

VLT peer device - One of a pair of devices that are connected with the special port channel known as the VLT interconnect (VLTi).

### Configuring Virtual Link Trunking

VLT requires that you enable the feature and then configure the VLT domain, backup link, and VLT interconnect on both peer switches.

### Important Points to Remember

- VLT port channel interfaces must be switch ports.
- If RSTP is included on the system, it must be configured before VLT. See RSTP Configuration.
- Dell Force10 strongly recommends that the VLTi (VLT interconnect) must be a static LAG and that LACP should be disabled on the VLTi.
- The spanning tree root bridge should be at the Aggregation layer. If RSTP is enabled on the VLT device, refer to RSTP and VLT for guidelines to avoid traffic loss.
- If both VLT peers are rebooted in BMP mode and VLT LAGs are static, the DHCP server reply to the DHCP discover offer may not be forwarded by the ToR to the correct node. To avoid this scenario, configure the VLT LAGs to the ToR and the ToR port channel to the VLT peers with LACP. If supported by the ToR, enable the `lacp ungroup member-independent port-channel`.
- If the `lacp ungroup` feature is not supported on the ToR, VLT peers should be rebooted one at a time. After rebooting, verify that VLTi (ICL) is active before attempting DHCP connectivity.
- When IGMP snooping is enabled on the VLT peers, ensure the value of the `delay-restore` command is not less than the query interval.
- When Layer 3 routing protocols are enabled on VLT peers, make sure the `delay-restore` timer is set to a value that allows sufficient time for all routes to establish adjacency and exchange all the L3 routes between the VLT peers before the VLT ports are enabled.
- The `lacp ungroup member-independent` command should only be used if the system connects to nodes using BMP to upgrade or boot from the network.
- Ensure all port channels where LACP ungroup is applicable are configured as hybrid ports and as untagged members of a VLAN. BMP uses untagged DHCP packets to communicate with the DHCP server.
- If the DHCP server is located on the ToR and the VLTi (ICL) is down due to a failed link when a VLT node is rebooted in BMP mode, it will not be able to reach the DHCP server, resulting in BMP failure.
Configuration Notes

When you configure VLT, the following conditions apply:

- **VLT domain:**
  - A VLT domain supports two chassis members, which appear as a single logical device to network access devices connected to VLT ports through a port channel.
  - A VLT domain consists of the two core chassis, the interconnect trunk, backup link, and the LAG members connected to attached devices. The domain ID can be from 1 to 1000.
  - Each VLT domain has a unique MAC address that is created automatically by VLT or user-configured.
  - ARP tables are synchronized between the VLT peer nodes.
  - VLT peer switches operate as separate chassis with independent control and data planes for devices attached on non-VLT ports.
  - One chassis in the VLT domain is assigned a primary role; the other chassis takes the secondary role. The primary and secondary roles are required for scenarios when connectivity between the chassis is lost. VLT assigns the primary chassis role according to the lowest MAC address. The primary role is also user-configurable.
  - In a VLT domain, the peer switches must run the same FTOS software version.
  - You must separately configure each VLT peer switch with the same VLT domain ID and the VLT version. If the system detects mismatches between VLT peer switches in the VLT domain ID or VLT version, the VLT Interconnect (VLTi) will not activate. To find the reason for the VLTi being down, use the `show vlt statistics` command to verify there are mismatch errors, then use the `show vlt brief` command on each VLT peer to view the VLT version on the peer switch. If the VLT version is more than one release different from the current version in use, the VLTi will not activate.
  - The chassis members in a VLT domain support connection to orphan hosts and switches that are not connected to both switches in the VLT core.

- **VLT interconnect (VLTi):**
  - The interconnect must consist of either 10G or 40G ports. A maximum of eight 10G or four 40G ports is supported. A combination of 10G or 40G ports is not supported.
  - A VLT interconnect over 1G ports is not supported.
  - The port channel must be in default mode (not switchport) to be recognized by VLTi.
  - The system will automatically include required VLANs in VLTi. You do not need to manually select VLANs.
  - VLT peer switches operate as separate chassis with independent control and data planes for devices attached on non-VLT ports.
  - Port-channel link aggregation (LAG) across the ports in the VLT interconnect is required; individual ports are not supported. Dell Force10 strongly recommends configuring a static LAG for VLTi.
  - IGMP state information is synchronized between the VLT chassis over the VLT interconnect.
  - The traffic transmitted over VLT interconnect is prioritized, and allows you to configure the traffic class-to-queue assignment.
  - The VLT interconnect synchronizes L2 and L3 control-plane information across the two chassis.
The VLT interconnect is used for data traffic only when there is a link failure that requires the VLTi to be used in order for data packets to reach their final destination.

Unknown, multicast and broadcast traffic can be flooded across the VLT interconnect.

MAC addresses for VLANs configured across VLT peer chassis are synchronized over the VLT interconnect on an egress port such as a VLT LAG. MAC addresses are the same on both VLT peer nodes.

ARP entries configured across the VLTi are the same on both VLT peer nodes.

If you shut down the port channel used in the VLT interconnect on a peer switch in a VLT domain in which no backup link is configured, the switch’s role is displayed in `show vlt brief` command output as Primary instead of Standalone.

When you change the default VLAN ID on a VLT peer switch, the VLT interconnect may flap.

In a VLT domain, the following software features are supported on VLTi: LLDP, flow control, port monitoring, jumbo frames, DCB.

When the VLTi link is enabled, the link between the VLT peer switches is established if the following configured information is true on both peer switches:

- the VLT system MAC address matches.
- the VLT unit-id is not identical.

**Note:** If the VLT system MAC address or VLT unit-id is configured on only one of the VLT peer switches, the link between the VLT peer switches will be not be established. Each VLT peer switch must be correctly configured to establish the link between the peers.

- If the link between the VLT peer switches is established, changing the VLT system MAC address or the VLT unit-id will not cause the link between the VLT peer switches to become disabled. However, removing the VLT system MAC address or the VLT unit-id may cause the VLT ports to be disabled if the unit ID or system MAC address happens to be configured on only one VLT peer at any time.

- If the link between VLT peer switches is established, any change to the VLT system MAC address or unit-id will fail if the changes made create a mismatch by causing the VLT unit-ID to be the same on both peers and/or the VLT system MAC address does not match on both peers.

- If you replace a VLT peer node, pre-configure the switch with the VLT system MAC address, unit-id, and other VLT parameters before connecting it to the existing VLT peer switch using the VLTi connection.

- If the size of the MTU for VLTi members is less than 1496 bytes, MAC addresses may not be synced. Dell Force10 recommends retaining the default MTU allocation (1554 bytes) for VLTi members.

**VLT Backup link:**

- In the backup link between peer switches, heartbeat messages are exchanged between the two chassis for health checks. The default time interval between heartbeat messages over the backup link is 1 second. This interval is user-configurable; range: 1 to 5 seconds. DSCP marking on heartbeat messages is CS6.

- In order that the chassis backup link does not share the same physical path as the interconnect trunk, it is recommended that you use the management ports on the chassis and traverse an out-of-band management network. The backup link can use user ports, but not the same ports as used by the interconnect trunk.
• The chassis backup link does not carry control plane information or data traffic. Its use is restricted to health checks only.

• **Virtual link trunks (VLTs) between access devices and VLT peer switches:**
  - To connect servers and access switches with VLT peer switches, you use a VLT port channel (see Figure 55-1). Up to 48 port-channels are supported; up to 8 member links are supported in each port channel between the VLT domain and an access device.
  - The ID number of the port channel that connects an access device and a VLT switch is automatically generated by the discovery protocol running between VLT peers. The discovery protocol uses LACP properties to identify connectivity to a common client device and automatically generate a VLT number for port channels on VLT peers that connect to the device. The discovery protocol requires that an attached device should always run LACP over the port-channel interface.
  - VLT provides a loop-free topology for port channels with endpoints on different chassis in the VLT domain.
  - VLT uses shortest path routing so that traffic destined to hosts via directly attached links on a chassis does not traverse the chassis-interconnect link.
  - VLT allows multiple active parallel paths from access switches to VLT chassis.
  - VLT supports port-channel links with LACP between access switches and VLT peer switches. Dell Force10 recommends that you use static port channels on VLTi.
  - If VLTi connectivity with a peer is lost but the VLT backup connectivity indicates the peer is still alive, the VLT ports on the Secondary peer are orphaned and will be shut down.

In one possible topology, a switch uses the BMP feature to receive its IP address, configuration files, and boot image from a DHCP server that connects to the switch through the VLT domain. In the port-channel used by the switch to connect to the VLT domain, you must configure the port interfaces on each VLT peer as hybrid ports before adding them to the port channel (see Connect a VLT domain to an attached access device (switch or server)). To configure a port in hybrid mode so that it can carry untagged, single-tagged, and double-tagged traffic, enter the `portmode hybrid` command in interface configuration mode as described in Native VLANs.

• For example, if the DHCP server is located on the ToR and VLTi (ICL) is down (due to either an unavailable peer or a link failure), whether the VLT LAG is configured as static or LACP, when a single VLT peer is rebooted in BMP mode, it will not be able to reach the DHCP server, resulting in BMP failure.

• **Software features supported on VLT port-channels:**
  - In a VLT domain, the following software features are supported on VLT port-channels: 802.1p, designated router functionality of PIM Sparse-Mode, VRRP, Layer 3 VLANs, LLDP, flow control, port monitoring, jumbo frames, IGMP Snooping, sFlow, ingress and egress ACLs, and Layer 2 control protocols RSTP only).

**Note:** PVST+ passthrough is supported in a VLT domain. PVST+ BPDUs will not result in an interface shutdown. PVST+ BPDUs for a non-default VLAN will be flooded out as any other L2 multicast packet. On a default VLAN, RTSP will be part of the PVST+ topology in that specific VLAN (default VLAN).

• Refer to VLT and VRRP interoperability: for detailed information on how to use VRRP in a VLT domain.
• Refer to VLT and IGMP Snooping for information on configuring IGMP Snooping in a VLT domain.
• All system management protocols are supported on VLT ports, including SNMP, RMON, AAA, ACL, DNS, FTP, SSH, Syslog, NTP, RADIUS, SCP, TACACS+, Telnet, and LLDP.
• Layer 3 VLAN connectivity VLT peers is enabled by configuring a VLAN network interface for the same VLAN on both switches.
• IGMP snooping is supported over VLT ports. The multicast forwarding state is synchronized on both VLT peer switches. The IGMP snooping process on a VLT peer shares the learned group information with the other VLT peer over the chassis interconnect trunk.

- **Software features supported on VLT physical ports:**
  - In a VLT domain, the following software features are supported on VLT physical ports: 802.1p, LLDP, flow control, port monitoring, jumbo frames.

- **Software features not supported with VLT:**
  - In a VLT domain, the following software features are supported on non-VLT ports: 802.1x, ingress and egress ACLs, BGP, DHCP relay, DHCP Snooping, FRRP, IS-IS, IPv6 dynamic routing, OSPF, Active-Active PIM-SM, PIM-SSM, ingress and egress QOS, sFlow, and all protocols currently supported in FTOS.

- **VLT and VRRP interoperability:**
  - In a VLT domain, VRRP inter-operates with virtual link trunks that carry traffic to and from access devices (see Figure 55-1). The VLT peers belong to the same VRRP group and are assigned master and backup roles. Each peer actively forwards L3 traffic, reducing the traffic flow over the VLT interconnect.
  - VRRP elects the router with the highest priority as the master in the VRRP group. To ensure VRRP operation in a VLT domain, you must configure VRRP group priority on each VLT peer so that a peer is either the master or backup for all VRRP groups configured on its interfaces. See [Set VRRP Group (Virtual Router) Priority](#) for more information.
  - To verify that a VLT peer is consistently configured for either the master or backup role in all VRRP groups, enter the `show vrrp` command on each peer.
  - You must also configure the same L3 routing (static and dynamic) on each peer so that L3 reachability and routing tables are identical on both VLT peers. Both the VRRP master and backup peers should be able to locally forward L3 traffic in the same way.
  - In a VLT domain, although both VLT peers actively participate in L3 forwarding as the VRRP master or backup router, the `show vrrp` command output displays one peer as master and the other peer as backup.

- **Failure scenarios:**
  - On a link failover, when a VLT port channel fails, the traffic destined for that VLT port channel is re-directed to the VLTi to avoid flooding.
  - When a VLT switch determines that a VLT member port has failed (and that no other local member ports are available), the peer with the failed member port notifies the remote peer that it no longer has an active member port for a link. The remote peer then enables data forwarding across the VLT interconnect for packets that would otherwise have been forwarded over the failed member port. This mechanism ensures reachability and provides loop management.
  - If all ports in the chassis interconnect trunk fail, or if the messaging infrastructure fails to communicate across the interconnect trunk, the VLT management system uses the backup link interface to determine whether the failure is a link-level failure or whether in fact the remote peer has failed entirely. If the remote peer is still alive (heartbeat messages are still being received), the VLT secondary switch disables its VLT port channels. If keepalive messages from the peer are not being received, the peer continues to forward traffic, assuming that it is the last device available in
the network. In either case, upon recovery of the peer link or reestablishment of message forwarding across the interconnect trunk, the two VLT peers resynchronize any MAC addresses learned while communication was interrupted, and the VLT system continues normal data forwarding.

- If the primary chassis is rebooted, the secondary chassis takes on the operational role of the primary. When operation of the original, primary chassis is restored, it takes on the operational role of the secondary chassis.

- The SNMP MIB reports VLT statistics.

**RSTP and VLT**

VLT provides loop-free redundant topologies and does not require rapid spanning tree protocol (RSTP). RSTP can cause temporary port state blocking and may cause topology changes after link or node failures. Spanning tree topology changes are distributed to the entire layer 2 network, which can cause a network-wide flush of learned MAC and ARP addresses, requiring these addresses to be re-learned. However, enabling RSTP can detect potential loops caused by non-system issues such as cabling errors or incorrect configurations. RSTP is useful for potential loop detection but should be configured using the specifications below to minimize possible topology changes after link or node failure.

The following recommendations will help you avoid these issues and the associated traffic loss caused by using rapid spanning trees when VLT is enabled on both VLT peers:

- Any ports at the edge of the spanning tree's operating domain should be configured as edge ports, which are directly connected to end stations or server racks. Ports connected directly to layer 3-only routers not running STP should have RSTP disabled or be configured as edge ports.

- Ensure the primary VLT node is the root bridge and the secondary VLT peer node has the second-best bridge ID in the network. If the primary VLT peer node fails, the secondary VLT peer node becomes the root bridge, avoiding problems with spanning tree port state changes that occur when a VLT node fails or recovers.

- Even with this configuration, if the node has non-VLT ports using RSTP that are not configured as edge ports and are connected to other layer 2 switches, spanning tree topology changes can still be detected after VLT node recovery. To avoid this scenario, ensure that any non-VLT ports are configured as edge ports or have RSTP disabled.

**VLT Bandwidth Monitoring**

When bandwidth usage of the VLTi (ICL) exceeds 80%, a syslog error message (Message 1) and an SNMP trap are generated.

**Message 1 VLTi Bandwidth Usage Exceeding Threshold Value Error**

```%
%STKUNIT0-M:CP %VLTMGR-6-VLT-LAG-ICL: Overall Bandwidth utilization of VLT-ICL-LAG (port-channel 25) crosses threshold. Bandwidth usage (80% )
```
When the bandwidth usage drops below the 80% threshold, the system generates another syslog message (Message 2) and an SNMP trap.

**Message 2**  Excessive VLTi Bandwidth Usage Drops Below Threshold Value Error

| %STKUNIT0-M:CP %VLTMGGR-6-VLT-LAG-ICL: Overall Bandwidth utilization of VLT-ICL-LAG (port-channel 25) reaches below threshold. Bandwidth usage (74% )VLT show remote port channel status |

### VLT and IGMP Snooping

When configuring IGMP Snooping with VLT, you must ensure the configurations on both sides of the VLT trunk are identical to get the same behavior on both sides of the trunk. When IGMP Snooping is configured on a VLT node, the dynamically learned groups and multicast router ports are automatically learned on the VLT peer node.

### VLT Port Delayed Restoration

With FTOS Release 8.3.12.0, when a VLT node boots up, if the VLT ports have been previously saved in the start-up configuration, they will not be immediately enabled. To ensure MAC and ARP entries from the VLT per node are downloaded to the newly enabled VLT node, the system allows time for the VLT ports on the new node to be enabled and begin receiving traffic.

The delay-restore feature waits for all saved configurations to be applied, then starts a configurable timer. After the timer expires, the VLT ports are enabled one-by-one in a controlled manner. The delay between bringing up each VLT port-channel is proportional to the number of physical members in the port-channel.

Use the delay-restore command to change the duration of the configurable timer; the default is 90 seconds.

If IGMP Snooping is enabled, IGMP queries are also sent out on the VLT ports at this time allowing any receivers to respond to the queries and update the multicast table on the new node.

This delay in bringing up the VLT ports also applies when the VLTi link recovers from a failure that caused the VLT ports on the secondary VLT peer node to be disabled.
PIM-Sparse Mode Support on VLT

The Designated Router functionality of the PIM Sparse-Mode multicast protocol is supported on VLT peer switches for multicast sources and receivers that are connected to VLT ports. The VLT peer switches can act as a last-hop router for IGMP receivers and as a first-hop router for multicast sources.

On each VLAN where the VLT peer nodes act as the first hop or last hop routers, one of the VLT peer nodes will be elected as the PIM Designated Router. If IGMP Snooping is configured along with PIM on the VLT VLANs, then VLTi must be configured as the static multicast router port on both VLT peer switches. This allows multicast traffic originating from the source connected to the VLT ports to reach the PIM router that has downstream neighbors.

The VLT peer nodes can also act as normal PIM routers on Layer 3 ports and on VLANS that do not have any VLT port members. In addition to being first-hop or last-hop routers, the peer node can also act as an intermediate router.

To route traffic to and from the multicast source and receiver that are connected to VLT ports, enable PIM-Sparse mode on the VLANs to which the VLT ports belong using the `ip pim sparse-mode` command. If IGMP Snooping is configured on these VLANs, the VLTi must be configured as a static multicast router port on both VLT peers.

Use the `show ip pim neighbor`, `show ip igmp snooping mrouter`, and `show running config` commands to verify the PIM neighbors on the VLT VLAN and on the multicast port:

VLT peer nodes cannot be configured rendezvous points, PIM routers cannot be connected to VLT ports; you must use a different port.
If the VLT node elected as the designated router fails, traffic loss will occur until another VLT node is elected the designated router.

## RSTP Configuration

The RSTP Spanning Tree protocol is supported in a VLT domain. Before you configure VLT on peer switches, you must configure the Rapid Spanning Tree Protocol (RSTP) in the network if it will be included in your configuration. RSTP is required for initial loop prevention during the VLT startup phase. RSTP may also be used for loop prevention in the network outside of the VLT port channel. For information on how to configure RSTP, see Chapter 42, Rapid Spanning Tree Protocol (RSTP).

RSTP must be running on both VLT peer switches. The primary VLT peer controls RSTP states, such as forwarding and blocking, on both the primary and secondary peers. It is recommended that you configure the primary VLT peer as the RSTP primary root device and configure the secondary VLT peer as the RSTP secondary root device.

BPDUs use the MAC address of the primary VLT peer as the RSTP bridge ID in the designated bridge ID field. The primary VLT peer sends these BPDUs on VLT interfaces connected to access devices. The MAC address for a VLT domain is automatically selected on the peer switches when you create the domain.

You must configure both ends of the VLT interconnect trunk with identical RSTP configurations. When VLT is enabled, the `show spanning-tree rstp brief` command output displays VLT information.

### Preventing Forwarding Loops in a VLT Domain

During the bootup of VLT peer switches, a forwarding loop may occur until the VLT configurations are applied on each switch and the primary/secondary roles are determined. To prevent the interfaces in the VLT interconnect trunk and RSTP-enabled VLT ports from entering a forwarding state and creating a traffic loop in a VLT domain, you must take the following steps:

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure RSTP in the core network and on each peer switch as described in Chapter 42, Rapid Spanning Tree Protocol (RSTP). Disabling RSTP on one VLT peer may result in a VLT domain failure.</td>
<td>no disable</td>
<td>PROTOCOL SPANNING TREE RSTP</td>
</tr>
<tr>
<td>2</td>
<td>Enable RSTP on each peer switch.</td>
<td>bridge-priority</td>
<td>PROTOCOL SPANNING TREE RSTP</td>
</tr>
<tr>
<td>3</td>
<td>Configure each peer switch with a unique bridge priority.</td>
<td>bridge-priority</td>
<td>PROTOCOL SPANNING TREE RSTP</td>
</tr>
</tbody>
</table>
Sample RSTP Configuration

Using Figure 55-1 as a sample VLT topology, the primary VLT switch will send BPDUs to an access device (switch or server) with its own RSTP bridge ID. BPDUs generated by an RSTP-enabled access device are only processed by the primary VLT switch. The secondary VLT switch tunnels the BPDUs that it receives to the primary VLT switch over the VLT interconnect. Only the primary VLT switch determines the RSTP roles and states on VLT ports, and ensures that the VLT interconnect link is never blocked.

In case of a primary VLT switch failure, the secondary switch starts sending BPDUs with its own bridge ID and inherits all the port states from the last synchronization with the primary switch. An access device never detects the change in primary/secondary roles and does not see it as a topology change.

The following figures show an example of the RSTP configuration that you must perform on each peer switch to prevent forwarding loops.

Figure 55-2. Configuring RSTP on VLT Peers to Prevent Forwarding Loops (VLT Peer 1)

```
FTOS_VLTpeer1(conf)#protocol spanning-tree rstp
FTOS_VLTpeer1(conf-rstp)#no disable
FTOS_VLTpeer1(conf-rstp)#bridge-priority 4096
```

Figure 55-3. Configuring RSTP on VLT Peers to Prevent Forwarding Loops (VLT Peer 2)

```
FTOS_VLTpeer2(conf)#protocol spanning-tree rstp
FTOS_VLTpeer2(conf-rstp)#no disable
FTOS_VLTpeer2(conf-rstp)#bridge-priority 0
```

VLT Configuration Procedure

To configure virtual link trunking and create a VLT domain in which two S4810 or Z9000 switches are physically connected and treated as a single port channel by access devices, you must configure the following settings on each VLT peer device:

Prerequisite: Before you begin, make sure that both VLT peer switches are running the same FTOS version and are configured for RSTP as described in Chapter 42, Rapid Spanning Tree Protocol (RSTP). For VRRP operation, ensure that VRRP groups and L3 routing on each VLT peer are configured as described in Virtual Router Redundancy Protocol (VRRP).

1. Configure the VLT interconnect for the VLT domain. The primary and secondary switch roles in the VLT domain are automatically assigned after both sides of the VLTi are configured.

   Note: If a third-party ToR unit is used, Dell Force10 recommends using static LAGs on the VLTi between VLT peers to avoid potential problems if the VLT peers are rebooted.

2. Enable VLT and create a VLT domain ID. VLT automatically selects a system MAC address.

3. Configure a backup link for the VLT domain.
4. (Optional) Manually reconfigure default VLT settings, such as MAC address and VLT primary/secondary roles.

5. Connect the peer switches in a VLT domain to an attached access device (switch or server).

### Configure a VLT interconnect

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure the port channel to be used for the VLT interconnect on a VLT switch and enter interface configuration mode. Enter the same port-channel number configured with the peer-link port-channel command.</td>
<td>interface port-channel id-number</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Remove an IP address from the interface.</td>
<td>no ip address</td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
<tr>
<td>3</td>
<td>Add one or more port interfaces to the port channel. The interface specifies one of the following interface types: 1-Gigabit Ethernet: Enter gigabitethernet slot/port. 10-Gigabit Ethernet: Enter tengigabitethernet slot/port.</td>
<td>channel-member interface</td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
<tr>
<td>4</td>
<td>Ensure that the port channel is active.</td>
<td>no shutdown</td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
</tbody>
</table>

**Note:** To be included in the VLTI, the port channel must be in default mode (no switchport or VLAN assigned).

5. Repeat Steps 1 to 4 on the VLT peer switch to configure the VLT interconnect.

### Configure a VLT backup link

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Specify the management interface to be used for the backup link through an out-of-band management network. Enter the slot (0-1) and the port (0).</td>
<td>interface managementethernet slot/port</td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Configure an IPv4 address (A.B.C.D) or IPv6 address (X::X:1::X) and mask (/x) on the interface. This is the IP address to be configured on the VLT peer with the back-up destination command.</td>
<td>{ip address ipv4-address/ mask</td>
<td>MANAGEMENT INTERFACE</td>
</tr>
<tr>
<td>3</td>
<td>Ensure that the interface is active.</td>
<td>no shutdown</td>
<td>MANAGEMENT INTERFACE</td>
</tr>
<tr>
<td>4</td>
<td>Repeat Steps 1 to 3 on the VLT peer switch.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Use the `delay-restore` command at any time to set an amount of time, in seconds, to delay the system from restoring the VLT port. Refer to VLT Port Delayed Restoration for more information.

### Configure a VLT port delay period

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter VLT-domain configuration mode for a specified VLT domain. Range of domain IDs: 1 to 1000.</td>
<td><code>vlt domain domain-id</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Enter an amount of time, in seconds, to delay the restoration of the VLT ports after the system is rebooted. Range: 1-1200 Default: 90 seconds</td>
<td><code>delay-restore delay-restore-time</code></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

### (Optional) Reconfigure default VLT settings

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter VLT-domain configuration mode for a specified VLT domain. Range of domain IDs: 1 to 1000.</td>
<td><code>vlt domain domain-id</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>(Optional) After you configure the VLT domain on each peer switch on both sides of the interconnect trunk, by default, the FTOS software elects a primary and secondary VLT peer device. Use the <code>primary-priority</code> command to reconfigure the primary role of VLT peer switches. To configure the primary role on a VLT peer, enter a lower value than the priority value of the remote peer. Priority values are from 1 to 65535. Default: 32768.</td>
<td><code>primary-priority value</code></td>
<td>VLT DOMAIN</td>
</tr>
<tr>
<td>3</td>
<td>(Optional) When you create a VLT domain on a switch, the FTOS software automatically creates a VLT-system MAC address used for internal system operations. Use the <code>system-mac</code> command to explicitly configure the default MAC address for the domain by entering a new MAC address in the format: aaaa.bbbb.cccc. You must also reconfigure the same MAC address on the VLT peer switch. Use this command to minimize the time required for the VLT system to synchronize the default MAC address of the VLT domain on both peer switches when one peer switch reboots.</td>
<td><code>system-mac mac-address</code></td>
<td>VLT DOMAIN</td>
</tr>
</tbody>
</table>

Virtual Link Trunking (VLT) | 1077
(Optional) Reconfigure default VLT settings

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>(Optional) When you create a VLT domain on a switch, the FTOS software automatically assigns a unique unit ID (0 or 1) to each peer switch. The unit IDs are used for internal system operations. Use the <code>unit-id</code> command to explicitly configure the default values on each peer switch. You must configure a different unit ID (0 or 1) on each peer switch. Use this command to minimize the time required for the VLT system to determine the unit ID assigned to each peer switch when one peer switch reboots.</td>
<td>`unit-id {0</td>
<td>1}`</td>
</tr>
</tbody>
</table>

Connect a VLT domain to an attached access device (switch or server)

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure the same port channel ID number on each peer switch in the VLT domain to connect to an attached device as follows:</td>
<td><code>interface port-channel id-number</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td></td>
<td>On a VLT peer switch: Configure the same port channel to be used to connect to an attached device and enter interface configuration mode.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Remove an IP address from the interface.</td>
<td><code>no ip address</code></td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
<tr>
<td>3</td>
<td>Place the interface in Layer 2 mode.</td>
<td><code>switchport</code></td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
<tr>
<td></td>
<td>Add one or more port interfaces to the port channel.</td>
<td><code>channel-member interface</code></td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
<tr>
<td></td>
<td><code>interface</code> specifies one of the following interface types: 1-Gigabit Ethernet: Enter <code>gigabitethernet slot/port</code>. 10-Gigabit Ethernet: Enter <code>tengigabitethernet slot/port</code>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Ensure that the port channel is active.</td>
<td><code>no shutdown</code></td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
<tr>
<td>6</td>
<td>Associate the port channel to the corresponding port channel in the VLT peer for the VLT connection to an attached device. Valid port-channel ID numbers are from 1 to 128.</td>
<td><code>vlt-peer-lag port-channel id-number</code></td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
<tr>
<td>7</td>
<td>Repeat Steps 1 to 6 on the VLT peer switch to configure the same port channel as part of the VLT domain.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>On an attached switch or server: Configure a port channel to connect to the VLT domain and add port channels to it. Figure 55-12 shows how to verify the port-channel configuration.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Use the `peer-down-vlan` parameter to configure the VLAN where a VLT peer will forward received packets over the VLTi from an adjacent VLT peer that is down. When a VLT peer with BMP reboots, untagged DHCP discover packets are sent to the peer over the VLTi. Using this configuration ensures the DHCP discover packets are forwarded to the VLAN that has the DHCP server.

### (Optional) Configure a VLT VLAN peer-down

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter VLT-domain configuration mode for a specified VLT domain. Range of domain IDs: 1 to 1000.</td>
<td><code>vlt domain domain-id</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Enter the port-channel number that will act as the interconnect trunk. Range: 1 to 128.</td>
<td><code>peer-link port-channel id-number</code></td>
<td>VLT DOMAIN CONFIGURATION</td>
</tr>
<tr>
<td>3</td>
<td>Enter the VLAN ID number of the VLAN where the VLT forwards packets received on the VLTi from an adjacent peer that is down. Range: 1 to 4094.</td>
<td><code>peer-link port-channel id-number peer-down-vlan vlan interface number</code></td>
<td>VLT DOMAIN CONFIGURATION</td>
</tr>
</tbody>
</table>

Use the following procedure to configure enhanced VLT between two VLT domains on your network. Refer to eVLT Configuration Example for a sample configuration.

### (Optional) Configure Enhanced VLT (eVLT)

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure the port channel to be used for the VLT interconnect on a VLT switch and enter interface configuration mode. Enter the same port-channel number configured with the <code>peer-link port-channel</code> command.</td>
<td><code>interface port-channel id-number</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>2</td>
<td>Add one or more port interfaces to the port channel. <code>interface</code> specifies one of the following interface types: 1-Gigabit Ethernet: Enter <code>gigabitethernet slot/port</code>, 10-Gigabit Ethernet: Enter <code>tengigabitethernet slot/port</code>.</td>
<td><code>channel-member interface</code></td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
<tr>
<td>3</td>
<td>Enter VLT-domain configuration mode for a specified VLT domain. Range of domain IDs: 1 to 1000.</td>
<td><code>vlt domain domain-id</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>4</td>
<td>Enter the port-channel number that will act as the interconnect trunk. Range: 1 to 128.</td>
<td><code>peer-link port-channel id-number</code></td>
<td>VLT DOMAIN CONFIGURATION</td>
</tr>
</tbody>
</table>
### (Optional) Configure Enhanced VLT (eVLT)

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Configure the IP address of the management interface on the remote VLT peer to be used as the endpoint of the VLT backup link for sending out-of-band hello messages. You can optionally specify the time interval used to send hello messages. Range: 1 to 5 seconds.</td>
<td>back-up destination <code>ip-address [interval seconds]</code></td>
<td>VLT DOMAIN CONFIGURATION</td>
</tr>
<tr>
<td>6</td>
<td>When you create a VLT domain on a switch, the FTOS software automatically creates a VLT-system MAC address used for internal system operations. Use the <code>system-mac</code> command to explicitly configure the default MAC address for the domain by entering a new MAC address in the format: <code>aaaa.bbbb.cccc</code>. You must also reconfigure the same MAC address on the VLT peer switch. Use this command to minimize the time required for the VLT system to synchronize the default MAC address of the VLT domain on both peer switches when one peer switch reboots.</td>
<td><code>system-mac mac-address</code></td>
<td>VLT DOMAIN CONFIGURATION</td>
</tr>
<tr>
<td>7</td>
<td>When you create a VLT domain on a switch, the FTOS software automatically assigns a unique unit ID (0 or 1) to each peer switch. The unit IDs are used for internal system operations. Use the <code>unit-id</code> command to explicitly configure the default values on each peer switch. You must configure a different unit ID (0 or 1) on each peer switch. Use this command to minimize the time required for the VLT system to determine the unit ID assigned to each peer switch when one peer switch reboots.</td>
<td>`unit-id {0</td>
<td>1}`</td>
</tr>
</tbody>
</table>

### Configure enhanced VLT.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Configure the port channel to be used for the VLT interconnect on a VLT switch and enter interface configuration mode. Enter the same port-channel number configured with the <code>peer-link port-channel</code> command.</td>
<td><code>interface port-channel id-number</code></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>9</td>
<td>Place the interface in Layer 2 mode.</td>
<td><code>switchport</code></td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
<tr>
<td>10</td>
<td>Associate the port channel to the corresponding port channel in the VLT peer for the VLT connection to an attached device. Valid port-channel ID numbers are from 1 to 128.</td>
<td><code>vlt-peer-lag port-channel id-number</code></td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
</tbody>
</table>
**Optional Configure Enhanced VLT (eVLT)**

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Ensure that the port channel is active.</td>
<td><strong>no shutdown</strong></td>
<td>INTERFACE PORT-CHANNEL</td>
</tr>
</tbody>
</table>

Add links to the eVLT port.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Configure a range of interfaces to bulk configure.</td>
<td><strong>interface range {port-channel id}</strong></td>
<td>CONFIGURATION</td>
</tr>
<tr>
<td>13</td>
<td>Enable LACP on the LAN port.</td>
<td><strong>port-channel-protocol lacp</strong></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>14</td>
<td>Configure the LACP port channel mode.</td>
<td><strong>port-channel number mode [active]</strong></td>
<td>INTERFACE</td>
</tr>
<tr>
<td>15</td>
<td>Ensure that the interface is active.</td>
<td><strong>no shutdown</strong></td>
<td>MANAGEMENT INTERFACE</td>
</tr>
<tr>
<td>16</td>
<td>Repeat steps 1 through 15 for the VLT peer node in Domain 1.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Repeat steps 1 through 15 for the first VLT node in Domain 2.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Repeat steps 1 through 15 for the VLT peer node in Domain 2.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To verify the configuration of a VLT domain, enter any of the show commands described in [Verifying a VLT Configuration](#).

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Configure the VLT domain with the same ID in VLT peer 1 and VLT peer 2</td>
<td><strong>vlt domain domain id</strong></td>
<td>VLT DOMAIN</td>
</tr>
</tbody>
</table>

Configure the VLTi between VLT peer 1 and VLT peer 2.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. LACP/Static LAG can be configured between the peer units (not shown).</td>
<td><strong>interface port-channel port-channel id</strong></td>
<td>CONFIGURATION</td>
</tr>
</tbody>
</table>

**Note:** To benefit from the protocol negotiations, Dell Force10 recommends VLTs used as facing hosts/switches are configured with LACP. Both peers should use the same port channel ID.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Configure the peer-link port-channel in the VLT domains of each peer unit.</td>
<td><strong>channel-member</strong></td>
<td>INTERFACE PORTCHANNEL</td>
</tr>
</tbody>
</table>

Configure the backup link between the VLT peer units.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Configure the peer 2 management ip/ interface ip for which connectivity is present in VLT peer 1.</td>
<td><strong>show running-config vlt</strong></td>
<td>EXEC Privilege</td>
</tr>
</tbody>
</table>
In the following sample VLT configuration steps, VLT peer 1 is S4810-2, VLT peer 2 is S4810-4, and the ToR is S60-1:

Note: If a third-party ToR unit is used, Dell Force10 recommends using static LAGs with VLT peers to avoid potential problems if the VLT peers are rebooted.

Configure the VLT domain with the same ID in VLT peer 1 and VLT peer 2:

```
s4810-2(conf)#vlt domain 5  
s4810-2(conf-vlt-domain)#  
s4810-4(conf)#vlt domain 5  
s4810-4(conf-vlt-domain)#
```

Configure the VLTi between VLT peer 1 and VLT peer 2:

1. LACP/Static LAG can be configured between the peer units (not shown)
2. Configure the peer-link port-channel in the VLT domains of each peer unit.

```
s4810-2(conf)#interface port-channel 1  
s4810-2(conf-if-po-1)#channel-member TenGigabitEthernet 0/4-7  
s4810-2(conf)#no shutdown

s4810-4(conf)#interface port-channel 1  
s4810-4(conf-if-po-1)#channel-member TenGigabitEthernet 0/4-7  
s4810-4(conf)#no shutdown
```
Configure the backup link between the VLT peer units.

1. Configure the peer 2 management ip/ interface ip for which connectivity is present in VLT peer 1.
2. Configure the peer 1 management ip/ interface ip for which connectivity is present in VLT peer 2.

s4810-2# show running-config vlt
!
  vlt domain 5
  peer-link port-channel 1
  back-up destination 10.11.206.58
s4810-2#
s4810-2# show interfaces managementethernet 0/0
  Internet address is 10.11.206.43/16
s4810-4# show running-config vlt
!
  vlt domain 5
  peer-link port-channel 1
  back-up destination 10.11.206.43
s4810-4#
s4810-4# show running-config interface managementethernet 0/0
  ip address 10.11.206.58/16
  no shutdown

Configure the VLT links between VLT peer 1 and VLT peer 2 to the top of rack unit. In the following example, port Te 0/40 in VLT peer 1 is connected to Te 0/48 of TOR and port Te 0/18 in VLT peer 2 is connected to Te 0/50 of TOR.

1. Configure the static LAG/LACP between ports connected from VLT peer 1 and VLT peer 2 to the top of rack unit.
2. Configure the VLT peer link port channel id in VLT peer 1 and VLT peer 2.
3. In the top of rack unit, configure LACP in the physical ports (shown for VLT peer 1 only. Repeat steps for VLT peer 2. The highlighted `vlt-peer-lag port-channel 2` indicates that port-channel 2 is the port-channel id configured in VLT peer 2).

s4810-2# show running-config interface tengigabitethernet 0/40
!
  interface TenGigabitEthernet 0/40
  no ip address
  
  port-channel-protocol LACP
  port-channel 2 mode active
  no shutdown
s4810-2#
 configuring VLT peer lag in VLT
s4810-2# show running-config interface port-channel 2
!
  interface Port-channel 2
  no ip address
  switchport
  vlt-peer-lag port-channel 2
  no shutdown
s4810-2#
s4810-2# show interfaces port-channel 2 brief
  Codes: L - LACP Port-channel

<table>
<thead>
<tr>
<th>LAG</th>
<th>Mode</th>
<th>Status</th>
<th>Uptime</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>2</td>
<td>L2L3</td>
<td>03:33:14</td>
<td>Te 0/40 (Up)</td>
</tr>
</tbody>
</table>

s4810-2#
In the ToR unit, configure LACP on the physical ports.

```
s60-1#show running-config interface tengigabitethernet 0/48
  !
  interface TenGigabitEthernet 0/48
  no ip address
  !
  port-channel-protocol LACP
  port-channel 100 mode active
  no shutdown
s60-1#show running-config interface tengigabitethernet 0/50
  !
  interface TenGigabitEthernet 0/50
  no ip address
  !
  port-channel-protocol LACP
  port-channel 100 mode active
  no shutdown
s60-1#show running-config interface port-channel 100
  !
  interface Port-channel 100
  no ip address
  switchport
  no shutdown
s60-1#show interfaces port-channel 100 brief
Codes: L - LACP Port-channel
  LAG  Mode  Status       Uptime      Ports
  L   100 L2    up           03:33:48    Te 0/48    (Up)
          Te 0/50    (Up)
s60-1#
```

Verify VLT is up. Verify that the VLTi (ICL) link, backup link connectivity (heartbeat status) and VLT peer link (peer chassis) are all up:
FTOS(conf)#show vlt brief
VLT Domain Brief
------------------
Domain ID: 10
Role: Primary
Role Priority: 32768
ICL Link Status: Up
HeartBeat Status: Not Established
VLT Peer Status: Up
Version: 5(1)
Local System MAC address: 00:01:e8:8b:14:3c
Remote System MAC address: 00:01:e8:8b:15:20
Remote system version: 5 (1)
Delay-Restore timer: 90 seconds

FTOS#FTOS(conf-if-vl-100)#show vlt detail
Local LAG Id  Peer LAG Id  Local Status   Peer Status  Active VLANs
------------  -----------  ------------  ------------  -------------
10            10           UP            UP            100, 200, 300, 400,

Verify the VLT LAG is up in both VLT peer units.

s4810-2#show interfaces port-channel 2 brief
Codes: L - LACP Port-channel

<table>
<thead>
<tr>
<th>LAG</th>
<th>Mode</th>
<th>Status</th>
<th>Uptime</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>2</td>
<td>L2L3</td>
<td>up</td>
<td>Te 0/40     (Up)</td>
</tr>
</tbody>
</table>

s4810-2#

s4810-4#show interfaces port-channel 2 brief
Codes: L - LACP Port-channel

<table>
<thead>
<tr>
<th>LAG</th>
<th>Mode</th>
<th>Status</th>
<th>Uptime</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>2</td>
<td>L2L3</td>
<td>up</td>
<td>Te 0/18     (Up)</td>
</tr>
</tbody>
</table>

s4810-4#
eVLT Configuration Example

The following example demonstrates the steps to configure enhanced VLT (eVLT) in a network. In this example there are two domains being configured. Domain 1 consists of Peer 1 and Peer 2; Domain 2 consists of Peer 3 and Peer 4 as shown below.

In Domain 1, configure Peer 1 first, then configure Peer 2. When that is complete, perform the same steps for the peer nodes in Domain 2. The interface used in this example is TenGigabitEthernet.

In Domain 1, configure the VLT domain and VLTi on Peer 1:

```
Domain_1_Peer1#configure
Domain_1_Peer1(conf)#interface port-channel 1
Domain_1_Peer1(conf-if-po-1)#channel-member TenGigabitEthernet 0/8-9
Domain_1_Peer1#no shutdown
Domain_1_Peer1(conf)#vlt domain 1000
Domain_1_Peer1(conf-vlt-domain)#peer-link port-channel 1
Domain_1_Peer1(conf-vlt-domain)#back-up destination 10.16.130.11
Domain_1_Peer1(conf-vlt-domain)#system-mac mac-address 00:0a:00:0a:00:0a
Domain_1_Peer1(conf-vlt-domain)#unit-id 0
```

Configure eVLT on Peer 1:

```
Domain_1_Peer1(conf)#interface port-channel 100
Domain_1_Peer1(conf-if-po-100)#switchport
Domain_1_Peer1(conf-if-po-100)#vlt-peer-lag port-channel 100
Domain_1_Peer1(conf-if-po-100)#no shutdown
```

Add links to the eVLT port-channel on Peer 1:

```
Domain_1_Peer1(conf)#interface range tengigabitethernet 0/16 - 17
Domain_1_Peer1(conf-if-range-te-0/16-17)#port-channel-protocol LACP
Domain_1_Peer1(conf-if-range-te-0/16-17)#port-channel 100 mode active
```
Next, configure the VLT domain and VLTi on Peer 2:

```
Domain_1_Peer2#configure
Domain_1_Peer2(config)#interface port-channel 1
Domain_1_Peer2(config-if-po-1)#channel-member TenGigabitEthernet 0/8-9
Domain_1_Peer2#no shutdown

Domain_1_Peer2(config)#vlt domain 200
Domain_1_Peer2(config-vlt-domain)#peer-link port-channel 1
Domain_1_Peer2(config-vlt-domain)#back-up destination 10.16.130.12
Domain_1_Peer2(config-vlt-domain)#system-mac mac-address 00:0a:00:0a:00:0a
Domain_1_Peer2(config-vlt-domain)#unit-id 1

Configure eVLT on Peer 2:

```
Domain_1_Peer2(config)#interface port-channel 100
Domain_1_Peer2(config-if-po-100)#switchport
Domain_1_Peer2(config-if-po-100)#vlt-peer-lag port-channel 100
Domain_1_Peer2(config-if-po-100)#no shutdown

Add links to the eVLT port-channel on Peer 2:

```
Domain_1_Peer2(config)#interface range tengigabitethernet 0/28 - 29
Domain_1_Peer2(config-if-range-te-0/16-17)#port-channel-protocol LACP
Domain_1_Peer2(config-if-range-te-0/16-17)#port-channel 100 mode active
Domain_1_Peer2(config-if-range-te-0/16-17)#no shutdown

In Domain 2, configure the VLT domain and VLTi on Peer 3:

```
Domain_2_Peer3#configure
Domain_2_Peer3(config)#interface port-channel 1
Domain_2_Peer3(config-if-po-1)#channel-member TenGigabitEthernet 0/8-9
Domain_2_Peer3#no shutdown
Domain_2_Peer3(config)#vlt domain 200
Domain_2_Peer3(config-vlt-domain)#peer-link port-channel 1
Domain_2_Peer3(config-vlt-domain)#back-up destination 10.18.130.11
Domain_2_Peer3(config-vlt-domain)#system-mac mac-address 00:0b:00:0b:00:0b
Domain_2_Peer3(config-vlt-domain)#unit-id 0

Configure eVLT on Peer 3:

```
Domain_2_Peer3(config)#interface port-channel 100
Domain_2_Peer3(config-if-po-100)#switchport
Domain_2_Peer3(config-if-po-100)#vlt-peer-lag port-channel 100
Domain_2_Peer3(config-if-po-100)#no shutdown

Add links to the eVLT port-channel on Peer 3:

```
Domain_2_Peer3(config)#interface range tengigabitethernet 0/19 - 20
Domain_2_Peer3(config-if-range-te-0/16-17)#port-channel-protocol LACP
Domain_2_Peer3(config-if-range-te-0/16-17)#port-channel 100 mode active
Domain_2_Peer3(config-if-range-te-0/16-17)#no shutdown

Next, configure the VLT domain and VLTi on Peer 4:

```
Domain_2_Peer4#configure
Domain_2_Peer4(config)#interface port-channel 1
Domain_2_Peer4(config-if-po-1)#channel-member TenGigabitEthernet 0/8-9
Domain_1_Peer4#no shutdown
Configure eVLT on Peer 4:

```
Domain_2_Peer4(conf)#interface port-channel 100
Domain_2_Peer4(conf-if-po-100)#switchport
Domain_2_Peer4(conf-if-po-100)#vlt-peer-lag port-channel 100
Domain_2_Peer4(conf-if-po-100)#no shutdown
```

Add links to the eVLT port-channel on Peer 4:

```
Domain_2_Peer4(conf)#interface range tengigabitethernet 0/31 - 32
Domain_2_Peer4(conf-if-range-te-0/16-17)#port-channel-protocol LACP
Domain_2_Peer4(conf-if-range-te-0/16-17)#port-channel 100 mode active
Domain_2_Peer4(conf-if-range-te-0/16-17)#no shutdown
```

PIM-Sparse Mode Configuration Example

The sample configuration below shows how to configure the PIM Sparse mode designated router functionality on the VLT domain with two VLT port-channels that are members of VLAN 4001. Refer to the figure in PIM-Sparse Mode Support on VLT.

**Enable PIM Multicast Routing on the VLT node globally**

```
VLT_Peer1(conf)#ip multicast-routing
```

**Enable PIM on the VLT port VLANs**

```
VLT_Peer1(conf)#interface vlan 4001
VLT_Peer1(conf-if-vl-4001)#ip address 140.0.0.1/24
VLT_Peer1(conf-if-vl-4001)#ip pim sparse-mode
VLT_Peer1(conf-if-vl-4001)#tagged port-channel 101
VLT_Peer1(conf-if-vl-4001)#tagged port-channel 102
VLT_Peer1(conf-if-vl-4001)#no shutdown
VLT_Peer1(conf-if-vl-4001)#exit
```

**Configure the VLTi port as a Static Multicast Router port for the VLAN**

```
VLT_Peer1(conf)#interface vlan 4001
VLT_Peer1(conf-if-vl-4001)#ip igmp snooping mrouter interface port-channel 128
VLT_Peer1(conf-if-vl-4001)#exit
VLT_Peer1(conf)#end
```

Repeat these steps on VLT Peer Node 2

```
VLT_Peer2(conf)#ip multicast-routing
VLT_Peer2(conf)#interface vlan 4001
VLT_Peer2(conf-if-vl-4001)#ip address 140.0.0.2/24
VLT_Peer2(conf-if-vl-4001)#ip pim sparse-mode
VLT_Peer2(conf-if-vl-4001)#tagged port-channel 101
VLT_Peer2(conf-if-vl-4001)#tagged port-channel 102
VLT_Peer2(conf-if-vl-4001)#no shutdown
VLT_Peer2(conf-if-vl-4001)#ip igmp snooping mrouter interface port-channel 128
VLT_Peer2(conf-if-vl-4001)#exit
VLT_Peer2(conf)#end
```
Verifying a VLT Configuration

To monitor the operation or verify the configuration of a VLT domain, enter any of the following `show` commands on the primary and secondary VLT switches:

<table>
<thead>
<tr>
<th>Show Command Syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>show vlt backup-link</td>
<td>Displays information on backup link operation (see Figure 55-4).</td>
</tr>
<tr>
<td>Command Mode: EXEC</td>
<td></td>
</tr>
<tr>
<td>show vlt brief</td>
<td>Displays general status information about VLT domains currently configured on the switch (see Figure 55-5).</td>
</tr>
<tr>
<td>Command Mode: EXEC</td>
<td></td>
</tr>
<tr>
<td>show vlt detail</td>
<td>Displays detailed information about the VLT-domain configuration, including local and peer port-channel IDs, local and peer VLT switch status, and number of active VLANs on each port channel (see Figure 55-6).</td>
</tr>
<tr>
<td>Command Mode: EXEC</td>
<td></td>
</tr>
<tr>
<td>show vlt role</td>
<td>Displays the VLT peer status, role of the local VLT switch, VLT system MAC address and system priority, and the MAC address and priority of the locally-attached VLT device (see Figure 55-7).</td>
</tr>
<tr>
<td>Command Mode: EXEC</td>
<td></td>
</tr>
<tr>
<td>show running-config vlt [domain-id]</td>
<td>Display the current configuration of all VLT domains (Figure 55-8) or a specified group on the switch.</td>
</tr>
<tr>
<td>Command Mode: EXEC</td>
<td>Valid <code>domain-id</code> values are 1 to 1000.</td>
</tr>
<tr>
<td>show vlt statistics</td>
<td>Displays statistics on VLT operation (see Figure 55-9).</td>
</tr>
<tr>
<td>Command Mode: EXEC</td>
<td></td>
</tr>
<tr>
<td>show interfaces interface <code>interface</code></td>
<td>Displays the current status of a port or port-channel interface used in the VLT domain.</td>
</tr>
<tr>
<td>Command Mode: EXEC</td>
<td><code>interface</code> specifies one of the following interface types:</td>
</tr>
<tr>
<td></td>
<td>Fast Ethernet: Enter <code>fastethernet slot/port</code>.</td>
</tr>
<tr>
<td></td>
<td>1-Gigabit Ethernet: Enter <code>gigabitethernet slot/port</code>.</td>
</tr>
<tr>
<td></td>
<td>10-Gigabit Ethernet: Enter <code>tengigabitethernet slot/port</code>.</td>
</tr>
<tr>
<td></td>
<td>Port channel: Enter <code>port-channel {1-128}</code>.</td>
</tr>
</tbody>
</table>

**Figure 55-4.** show vlt backup-link Command Output on VLT peer switches

```
FTOS#VLTpeer1#show vlt backup-link

VLT Backup Link
-----------------
Destination: 10.11.200.18
Peer HeartBeat status: Up
HeartBeat Timer Interval: 1
HeartBeat Timeout: 3
UDP Port: 34998
HeartBeat Messages Sent: 1026
HeartBeat Messages Received: 1025

FTOS#VLTpeer2#show vlt backup-link

VLT Backup Link
-----------------
```
Destination:                    10.11.200.20  
Peer HeartBeat status:          Up  
HeartBeat Timer Interval:       1  
HeartBeat Timeout:              3  
UDP Port:                       34998  
HeartBeat Messages Sent:        1030  
HeartBeat Messages Received:    1014  

Figure 55-5.  show vlt brief Command Output on VLT peer switches

FTOS(conf)#show vlt brief  
VLT Domain Brief
------------------
Domain ID:                    10  
Role:                         Primary  
Role Priority:                32768  
ICL Link Status:              Up  
HeartBeat Status:             Not Established  
VLT Peer Status:              Up  
Version:                      5(1)  
Local System MAC address:     00:01:e8:8b:14:3c  
Remote System MAC address:    00:01:e8:8b:15:20  
Remote system version:        5 (1)  
Delay-Restore timer:          90 seconds

Figure 55-6.  show vlt detail Command Output

FTOS#FTOS(conf-if-vl-100)#show vlt detail  
Local LAG Id  Peer LAG Id  Local Status   Peer Status  Active VLANs
------------  -----------  ------------  ------------  -------------
10            10           UP            UP            100, 200, 300, 400,

Figure 55-7.  show vlt role Command Output on VLT peer switches

FTOS#VLTpeer1#show vlt role  
VLT Role
--------
VLT Role:                     Primary  
System MAC address:           00:01:e8:8a:df:bc  
System Role Priority:         32768  
Local System MAC address:     00:01:e8:8a:df:bc  
Local System Role Priority:   32768

FTOS#VLTpeer2#show vlt role  
VLT Role
--------
VLT Role:                     Secondary  
System MAC address:           00:01:e8:8a:df:bc  
System Role Priority:         32768  
Local System MAC address:     00:01:e8:8a:df:e6  
Local System Role Priority:   32768
Sample Configuration: Virtual Link Trunking

To configure virtual link trunking, you must create a VLT domain, configure a backup link and interconnect trunk, and connect the peer switches in a VLT domain to an attached access device (switch or server).

Figure 55-10 and Figure 55-11 show a sample configuration of virtual link trunking on each VLT peer switch.

Figure 55-12 shows how to verify the connection to a VLT domain from an attached switch.
Figure 55-10. Configuring Virtual Link Trunking (VLT Peer 1)

FTOS_VLTpeer1(conf)#vlt domain 999
FTOS_VLTpeer1(conf-vlt-domain)#peer-link port-channel 100
FTOS_VLTpeer1(conf-vlt-domain)#back-up destination 10.11.206.35
FTOS_VLTpeer1(conf-vlt-domain)#exit

FTOS_VLTpeer1(conf)#interface ManagementEthernet 0/0
FTOS_VLTpeer1(conf-if-ma-0/0)#ip address 10.11.206.23/16
FTOS_VLTpeer1(conf-if-ma-0/0)#no shutdown
FTOS_VLTpeer1(conf-if-ma-0/0)#exit

FTOS_VLTpeer1(conf)#interface port-channel 100
FTOS_VLTpeer1(conf-if-po-100)#no ip address
FTOS_VLTpeer1(conf-if-po-100)#channel-member fortyGigE 0/56,60
FTOS_VLTpeer1(conf-if-po-100)#no shutdown
FTOS_VLTpeer1(conf-if-po-100)#exit

FTOS_VLTpeer1(conf)#interface port-channel 110
FTOS_VLTpeer1(conf-if-po-110)#no ip address
FTOS_VLTpeer1(conf-if-po-110)#switchport
FTOS_VLTpeer1(conf-if-po-110)#channel-member fortyGigE 0/52
FTOS_VLTpeer1(conf-if-po-110)#no shutdown
FTOS_VLTpeer1(conf-if-po-110)#vlt-peer-lag port-channel 110
FTOS_VLTpeer1(conf-if-po-110)#end

FTOS_VLTpeer1#show vlan id 10
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
  i - Internal untagged, I - Internal tagged, v - VLT untagged, V - VLT tagged

<table>
<thead>
<tr>
<th>NUM</th>
<th>Status</th>
<th>Description</th>
<th>Q Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Active</td>
<td>T Po100(Fo 0/56,60) V Po110(Fo 0/52)</td>
<td></td>
</tr>
</tbody>
</table>

Enable VLT and create a VLT domain with a backup-link and interconnect (VLTi)
Configure the backup link
Configure the VLT interconnect (VLTi)
Configure the port channel to an attached device
Verify that the port channels used in the VLT domain are assigned to the same VLAN
Figure 55-11. Configuring Virtual Link Trunking (VLT Peer 2)

```
FTOS_VLTpeer2(conf)#vlt domain 999
FTOS_VLTpeer2(conf-vlt-domain)#peer-link port-channel 100
FTOS_VLTpeer2(conf-vlt-domain)#back-up destination 10.11.206.23
FTOS_VLTpeer2(conf-vlt-domain)#exit

FTOS_VLTpeer2(conf)#interface ManagementEthernet 0/0
FTOS_VLTpeer2(conf-if-ma-0/0)#ip address 10.11.206.35/16
FTOS_VLTpeer2(conf-if-ma-0/0)#no shutdown
FTOS_VLTpeer2(conf-if-ma-0/0)#exit

FTOS_VLTpeer2(conf)#interface port-channel 100
FTOS_VLTpeer2(conf-if-po-100)#no ip address
FTOS_VLTpeer2(conf-if-po-100)#channel-member fortyGigE 0/46,50
FTOS_VLTpeer2(conf-if-po-100)#no shutdown
FTOS_VLTpeer2(conf-if-po-100)#exit

FTOS_VLTpeer2(conf)#interface port-channel 110
FTOS_VLTpeer2(conf-if-po-110)#no ip address
FTOS_VLTpeer2(conf-if-po-110)#switchport
FTOS_VLTpeer2(conf-if-po-110)#channel-member fortyGigE 0/48
FTOS_VLTpeer2(conf-if-po-110)#no shutdown
FTOS_VLTpeer2(conf-if-po-110)#vlt-peer-lag port-channel 110
FTOS_VLTpeer2(conf-if-po-110)#end

FTOS_VLTpeer2#show vlan id 10
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
i - Internal untagged, I - Internal tagged, v - VLT untagged, V - VLT tagged

NUM Status Description       Q Ports
10 Active                   T Po100(Fo 0/46,50)
                              V Po110(Fo 0/48)
```
Troubleshooting VLT

Use the following information to help troubleshoot different VLT issues that may occur.

Note: For information on VLT failure mode timing and its impact, contact your Dell Force10 representative.

<table>
<thead>
<tr>
<th>Description</th>
<th>Behavior at Peer Up</th>
<th>Behavior During Run Time</th>
<th>Action to Take</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth monitoring</td>
<td>A syslog error message and an SNMP trap is generated when the VLTi bandwidth usage goes above the 80% threshold and when it drops below 80%.</td>
<td>A syslog error message and an SNMP trap is generated when the VLTi bandwidth usage goes above its threshold.</td>
<td>Depending on the traffic that is received, the traffic can be offloaded into VLTi.</td>
</tr>
<tr>
<td>Domain ID mismatch</td>
<td>The VLT peer does not boot up. The VLTi is forced to a down state. A syslog error message and an SNMP trap are generated.</td>
<td>The VLT peer does not boot up. The VLTi is forced to a down state. A syslog error message and an SNMP trap are generated.</td>
<td>Verify that the domain ID matches on both VLT peers.</td>
</tr>
<tr>
<td>FTOS Version mismatch</td>
<td>A syslog error message is generated.</td>
<td>A syslog error message is generated.</td>
<td>Follow correct upgrade procedure for unit with mismatched FTOS version.</td>
</tr>
<tr>
<td>Remote VLT port channel status</td>
<td>N/A</td>
<td>N/A</td>
<td>Use the show vlt detail and show vlt brief commands to view VLT port channel status information.</td>
</tr>
<tr>
<td>Spanning tree mismatch at global level</td>
<td>All VLT port channels go down on both VLT peers. A syslog error message is generated.</td>
<td>No traffic is passed on the port channels. A one-time informational syslog message is generated.</td>
<td>During run time, a loop may occur as long as the mismatch lasts. To resolve, enable RSTP on both VLT peers.</td>
</tr>
<tr>
<td>Spanning tree mismatch at port level</td>
<td>A syslog error message is generated.</td>
<td>A one-time informational syslog message is generated.</td>
<td>Correct the spanning tree configuration on the ports.</td>
</tr>
<tr>
<td>System MAC mismatch</td>
<td>A syslog error message and an SNMP trap are generated. The VLT domain will not be formed. The VLTi will be in a down state.</td>
<td>A syslog error message and an SNMP trap are generated.</td>
<td>Verify that the unit ID of VLT peers is not the same on both units and that the MAC address is the same on both units.</td>
</tr>
<tr>
<td>Description</td>
<td>Behavior at Peer Up</td>
<td>Behavior During Run Time</td>
<td>Action to Take</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Unit ID mismatch</td>
<td>The VLT peer does not boot up. The VLTi is forced to a down state.</td>
<td>The VLT peer does not boot up. The VLTi is forced to a down state.</td>
<td>Verify the unit ID is correct on both VLT peers. Unit ID numbers must be sequential on peer units; i.e., if Peer 1 is unit ID “0”, Peer 2 unit ID must be “1”.</td>
</tr>
<tr>
<td></td>
<td>The VLT domain will not be formed. The VLTi will be in a down state.</td>
<td>A syslog error message is generated.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A syslog error message is generated.</td>
<td>A syslog error message is generated.</td>
<td></td>
</tr>
<tr>
<td>Version ID mismatch</td>
<td>A syslog error message and an SNMP trap are generated.</td>
<td>A syslog error message and an SNMP trap are generated.</td>
<td>Verify the FTOS software versions on the VLT peers is compatible. For more information, see the Release Notes for this release.</td>
</tr>
<tr>
<td>VLT LAG ID is not configured on one VLT peer</td>
<td>A syslog error message is generated. The peer with the VLT configured remains active.</td>
<td>The peer with the VLT configured remains active.</td>
<td>Verify the VLT LAG ID is configured correctly on both VLT peers.</td>
</tr>
<tr>
<td>VLT LAG ID mismatch</td>
<td>The VLT port channel is brought down. A syslog error message is generated.</td>
<td>The VLT port channel is brought down. A syslog error message is generated.</td>
<td>Perform a mismatch check after the VLT peer is established.</td>
</tr>
<tr>
<td>VLT LAG VLAN mismatch</td>
<td>A syslog error message is generated.</td>
<td>A syslog error message is generated.</td>
<td>Verify that the VLAN configuration is same for the VLT lags on both peers.</td>
</tr>
</tbody>
</table>
Virtual Router Redundancy Protocol (VRRP)

Virtual Router Redundancy Protocol (VRRP) is supported on platforms:

This chapter covers the following information:

- VRRP Overview
- VRRP Benefits
- VRRP Implementation
- VRRP Configuration
- Sample Configurations

VRRP Overview

Virtual Router Redundancy Protocol (VRRP) is designed to eliminate a single point of failure in a statically routed network.

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 56-1 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 56-1 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.

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**

E-Series supports an unlimited total number of VRRP groups on the switch while supporting up to 255 VRRP groups on a single interface (Table 56-1).

C-Series supports a total of 128 VRRP groups on the switch with varying number of maximum VRRP groups per interface (Table 56-1).

S25/S50 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 56-1).

The S55 and S60 support a total of 2000 VRRP groups on a switch and up to 100 VRRP groups per interface (Table 56-2).

The S4810 supports a total of 2000 VRRP groups on a switch and 512 VRRP groups per interface (Table 56-2).

**Note:** If you have a stacked S4810 system, Dell Force10 recommends that you configure the `mac-address-table station-move refresh-arp` command on the stack to prevent VRRP from being affected in case of a stack system failover.

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, S4810, S55, and S60, or the FP on the S25/S50. 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 56-1. Recommended VRRP Advertise Intervals

<table>
<thead>
<tr>
<th>Total VRRP Groups</th>
<th>Recommended Advertise Interval</th>
<th>Groups/Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E-Series</td>
<td>C-Series</td>
</tr>
<tr>
<td>Less than 250</td>
<td>1 second</td>
<td>1 second</td>
</tr>
<tr>
<td>Between 250 and 450</td>
<td>2 seconds</td>
<td>2 - 3 seconds</td>
</tr>
<tr>
<td>Between 450 and 600</td>
<td>3 seconds</td>
<td>4 seconds</td>
</tr>
<tr>
<td>Between 600 and 800</td>
<td>4 seconds</td>
<td>5 seconds</td>
</tr>
<tr>
<td>Between 800 and 1000</td>
<td>5 seconds</td>
<td>5 seconds</td>
</tr>
<tr>
<td>Between 1000 and 1200</td>
<td>7 seconds</td>
<td>7 seconds</td>
</tr>
<tr>
<td>Between 1200 and 1500</td>
<td>8 seconds</td>
<td>8 seconds</td>
</tr>
</tbody>
</table>

### Table 56-2. Recommended VRRP Advertise Intervals on the S4810, S55, S60

<table>
<thead>
<tr>
<th>Total VRRP Groups</th>
<th>Recommended Advertise Interval</th>
<th>Groups/Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S4810</td>
<td>S55</td>
</tr>
<tr>
<td>Less than 250</td>
<td>1 second</td>
<td>1 second</td>
</tr>
<tr>
<td>Between 250 and 450</td>
<td>2 - 3 seconds</td>
<td>2 - 3 seconds</td>
</tr>
<tr>
<td>Between 450 and 600</td>
<td>3 - 4 seconds</td>
<td>3 - 4 seconds</td>
</tr>
<tr>
<td>Between 600 and 800</td>
<td>4 seconds</td>
<td>4 seconds</td>
</tr>
<tr>
<td>Between 800 and 1000</td>
<td>5 seconds</td>
<td>5 seconds</td>
</tr>
<tr>
<td>Between 1000 and 1200</td>
<td>7 seconds</td>
<td>7 seconds</td>
</tr>
<tr>
<td>Between 1200 and 1500</td>
<td>8 seconds</td>
<td>8 seconds</td>
</tr>
</tbody>
</table>

The recommendations in Table 56-1 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 Configuration

By default, VRRP is not configured.

Configuration Task List for VRRP

The following list specifies the configuration tasks for VRRP:

- Create a Virtual Router (mandatory)
- Assign Virtual IP addresses (mandatory)
- Set VRRP Group (Virtual Router) Priority (optional)
- Configure VRRP Authentication (optional)
- Disable Preempt (optional)
- Change the Advertisement interval (optional)
- Track an Interface or Object (optional)
- VRRP initialization delay

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. In FTOS, a VRRP Group is identified by the Virtual Router Identifier (VRID).

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.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create a virtual router for that interface with a VRID.</td>
<td><strong>vrrp-group vrid</strong>&lt;br&gt;VRID Range: 1-255</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>

**Note:** The interface must already have a Primary IP Address defined, and be enabled.

Figure 56-2. Command Example: vrrp-group

FTOS(conf)#int gi 1/1
FTOS(conf-if-gi-1/1)#vrrp-group 111
FTOS(conf-if-gi-1/1-vrid-111)#
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.

E-Series supports an unlimited total number of VRRP Groups on the router while supporting up to 255 VRRP groups on a single interface (Table 56-1).

C-Series supports a total of 128 VRRP groups on the switch with varying number of maximum VRRP groups per interface (Table 56-1).

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 56-1).

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 on a single VRRP Group (VRID).

The following rules apply to virtual IP addresses:

- The virtual IP addresses must be in the same subnet as the primary or secondary IP addresses configured on the interface. Though a single VRRP group can contain virtual IP addresses belonging to multiple IP subnets configured on the interface, Dell Force10 recommends 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 MUST be set to 255. The interface then becomes the 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.

- If multiple VRRP groups are configured on an interface, only one of the VRRP Groups can contain the interface primary or secondary IP address.
Configure a Virtual IP address with these commands in the following sequence in the INTERFACE mode.

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Configure a VRRP group.</td>
<td><code>vrrp-group vrrp-id</code>&lt;br&gt;VRID Range: 1-255</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>2</td>
<td>Configure virtual IP addresses for this VRID.</td>
<td><code>virtual-address</code>&lt;br&gt;<code>ip-address1</code>&lt;br&gt;...&lt;br&gt;<code>ip-address12</code>&lt;br&gt;Range: up to 12 addresses</td>
<td>INTERFACE -VRID</td>
</tr>
</tbody>
</table>

**Figure 56-4. 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
FTOS(conf-if-gi-1/1-vrid-111)#
```

**Figure 56-5. Command Example Display: show config for the Interface**

Note that the Primary IP address and the Virtual IP addresses are on the same subnet in the following example.

```
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
 FTOS(conf-if-gi-1/1)#
```

**Figure 56-6** shows the same VRRP group (VRID 111) configured on multiple interfaces on different subnets.
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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure the priority for the VRRP group.</td>
<td>INTERFACE -VRID</td>
<td>priority priority</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range: 1 to 255</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Default: 100</td>
</tr>
</tbody>
</table>

**Figure 56-7. Command Example: priority in Interface VRRP mode**

```plaintext
FTOS(conf-if-gi-1/2)#vrrp-group 111
FTOS(conf-if-gi-1/2-vrid-111)#priority 125
```
Configure VRRP Authentication

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 the VRRP mode:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configure a simple text password.</td>
<td>authentication-type simple [encryption-type] password</td>
<td>INTERFACE-VRID</td>
</tr>
</tbody>
</table>

Parameters:
- `encryption-type`: 0 indicates unencrypted; 7 indicates encrypted
- `password`: plain text
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

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevent any BACKUP router with a higher priority from becoming the MASTER router.</td>
<td><code>no preempt</code></td>
<td>INTERFACE-VRID</td>
</tr>
</tbody>
</table>

**Figure 56-11. Command Example: no preempt**

```
FTOS(conf-if-gi-1/1/1-vrid-111)#vrrp-group 111
FTOS(conf-if-gi-1/1-vrid-111)#no preempt
FTOS(conf-if-gi-1/1-vrid-111)#show conf
```
Figure 56-12. 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:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change the advertisement interval setting.</td>
<td><code>advertise-interval seconds</code></td>
<td>INTERFACE-VRID</td>
</tr>
<tr>
<td></td>
<td>Range: 1-255 seconds</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Default: 1 second</td>
<td></td>
</tr>
</tbody>
</table>

Figure 56-13. 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)#
```
Track an Interface or Object

Set FTOS to monitor the state of any interface according to the Virtual group. Each VRRP group can track up to 12 interfaces, which may affect the priority of the VRRP group. If the tracked interface goes down, the VRRP group’s priority is decreased by a default value of 10 (also known as cost). If the tracked interface’s state goes up, the VRRP group’s priority is increased by 10.

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 the VRRP group may trigger an election. As the Master/Backup VRRP routers are selected based on the VRRP group’s priority, tracking features ensure that the best VRRP router is the Master for that group. The sum of all the costs of all the tracked interfaces must be less than the configured priority on the VRRP group. If the VRRP group is configured as Owner router (priority 255), tracking for that group is disabled, irrespective of the state of the tracked interfaces. The priority of the owner group always remains at 255.

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.
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, use the following command in the VRRP mode:

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitor an interface and, optionally, set a value to be subtracted from the interface’s VRRP group priority.</td>
<td><code>track interface [priority-cost cost]</code></td>
<td>INTERFACE-VRID</td>
</tr>
<tr>
<td>(Optional) Display the configuration and UP or DOWN state of tracked objects, including the client (VRRP group) that is tracking an object’s state.</td>
<td><code>show track</code></td>
<td>EXEC EXEC Privilege</td>
</tr>
<tr>
<td>(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.</td>
<td><code>show vrrp</code></td>
<td>EXEC EXEC Privilege</td>
</tr>
<tr>
<td>(Optional) Display the configuration of tracked objects in VRRP groups on a specified interface.</td>
<td><code>show running-config interface interface</code></td>
<td>EXEC EXEC Privilege</td>
</tr>
</tbody>
</table>

The sum of all the costs for all tracked interfaces must be less than the configured priority of the VRRP group.

**Figure 56-15. Command Example: track**

```
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 56-16. Command Example Display: track in VRID mode

```plaintext
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
   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 56-17. Command Example: show track

```plaintext
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
      Tracking 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
      Tracking by:
         VRRP GigabitEthernet 7/30 IPv6 VRID 1
```

Figure 56-18. Command Example: show vrrp

```plaintext
FTOS#show vrrp
ẟ-----------------
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
```
VRRP initialization delay

VRRP initialization delay is supported on the S4810 only.

When configured, VRRP is enabled immediately upon system reload or boot. VRRP initialization can be delayed to allow IGP and EGP protocols to be enabled prior to selecting the VRRP Master. This delay ensures that VRRP initializes with no errors or conflicts. The delay can be configured for up to 15 minutes, after which VRRP enables normally.

The delay timer is set on individual interfaces and is supported on all physical interfaces, VLANS and LAGs.

When both CLIs are configured, the later timer rules the VRRP enabling. For example, if `vrrp delay reload 600` and `vrrp delay minimum 300`, the following behavior occurs:

- When the system reloads, VRRP waits 600 seconds (10 minutes) to bring up VRRP on all interfaces that are up and configured for vrrp.
- When an interface comes up and becomes operational, the system waits 300 seconds (5 minutes) to bring up VRRP on that interface.

<table>
<thead>
<tr>
<th>Task</th>
<th>Command Syntax</th>
<th>Command Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set the delay time for VRRP initialization on an individual interface. This is the gap between an interface coming up and being operational, and VRRP enabling.</td>
<td><code>vrrp delay minimum seconds</code>&lt;br&gt;Seconds range: 0-900&lt;br&gt;Default: 0</td>
<td>INTERFACE</td>
</tr>
<tr>
<td>Set the delay time for VRRP initialization on all the interfaces in the system configured for VRRP. This is the gap between system boot up completion and VRRP enabling.</td>
<td><code>vrrp delay reload seconds</code>&lt;br&gt;Seconds range: 0-900&lt;br&gt;Default: 0</td>
<td>INTERFACE</td>
</tr>
</tbody>
</table>
Sample Configurations

VRRP for IPv4 Configuration

The configuration in Figure 56-20 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 56-20 shows the VRRP topology created with the CLI configuration in Figure 56-22.

Figure 56-20. VRRP for IPv4 Topology

State Master: R2 was the first interface configured with VRRP
Virtual MAC is automatically assigned and is the same on both Routers

State Backup: R3 was the second interface configured with VRRP

Virtual MAC address: 00:00:5e:00:01:63
Virtual IP address: 10.1.1.3
Authentication: (none)

Internet

VRID 99 10.1.1.3

10.1.1.1
GigE 3/21
R3

10.1.1.2
GigE 2/31
R2

R2#show vrrp
---------
GigabitEthernet 2/31, VRID: 99, Net: 10.1.1.1
State: Master, Priority: 100, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 661, Bad pkts rcvd: 0, Adv sent: 661, Gratuitous ARP sent: 1
Virtual MAC address: 00:00:5e:00:01:63
Virtual IP address: 10.1.1.3
Authentication: (none)
R2#

R3#show vrrp
----------
GigabitEthernet 3/21, VRID: 99, Net: 10.1.1.1
State: Backup, Priority: 100, Master: 10.1.1.1 (local)
Hold Down: 0 sec, Preempt: TRUE, AdvInt: 1 sec
Adv rcvd: 331, 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)
R3#
Figure 56-21. 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 10.1.1.1/24
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 56-22 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 56-22 shows the VRRP for IPv6 topology with the CLI configuration.

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.

Backup State: R3 is the backup in the VRRP group because the R3 interface has a lower IPv6 address.

You must configure both a virtual IPv6 address and a virtual link local (fe80) address for an IPv6 VRRP group.

---

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.
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.

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 1121.

Non-VLAN Scenario

Figure 56-24. VRRP in VRF: Non-VLAN Example

Figure 56-24 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 56-25. 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
```
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).
Switch-1

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 56-28. 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 56-29. Command Example: show running-config track interface**

```
FTOS#show running-config track 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 56-30. Command Example: show vrrp vrf**

```
FTOS#show vrrp vrf red
------------------------
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)
```
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
- Force10 — FRRP (Force10 Redundant Ring Protocol)
- 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

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>768</td>
<td>User Datagram Protocol</td>
<td>7.6.1 ✓ 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>793</td>
<td>Transmission Control Protocol</td>
<td>7.6.1 ✓ 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>854</td>
<td>Telnet Protocol Specification</td>
<td>7.6.1 ✓ 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>959</td>
<td>File Transfer Protocol (FTP)</td>
<td>7.6.1 ✓ 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>1321</td>
<td>The MD5 Message-Digest Algorithm</td>
<td>7.6.1 ✓ 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>1350</td>
<td>The TFTP Protocol (Revision 2)</td>
<td>7.6.1 ✓ 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>1661</td>
<td>The Point-to-Point Protocol (PPP)</td>
<td>✓</td>
</tr>
<tr>
<td>1989</td>
<td>PPP Link Quality Monitoring</td>
<td>✓</td>
</tr>
<tr>
<td>1990</td>
<td>The PPP Multilink Protocol (MP)</td>
<td>✓</td>
</tr>
<tr>
<td>1994</td>
<td>PPP Challenge Handshake Authentication Protocol (CHAP)</td>
<td>✓</td>
</tr>
<tr>
<td>2474</td>
<td>Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers</td>
<td>7.7.1 ✓ 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>2615</td>
<td>PPP over SONET/SDH</td>
<td>✓</td>
</tr>
<tr>
<td>2615</td>
<td>PPP over SONET/SDH</td>
<td>✓</td>
</tr>
<tr>
<td>2698</td>
<td>A Two Rate Three Color Marker</td>
<td>✓</td>
</tr>
<tr>
<td>3164</td>
<td>The BSD syslog Protocol</td>
<td>7.6.1 ✓ 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>draft-ietf-bfd-base-03</td>
<td>Bidirectional Forwarding Detection</td>
<td>7.6.1 ✓ 8.1.1</td>
</tr>
</tbody>
</table>
### General IPv4 Protocols

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>791</td>
<td>Internet Protocol</td>
<td>S</td>
</tr>
<tr>
<td>792</td>
<td>Internet Control Message Protocol</td>
<td>S</td>
</tr>
<tr>
<td>826</td>
<td>An Ethernet Address Resolution Protocol</td>
<td>S</td>
</tr>
<tr>
<td>1027</td>
<td>Using ARP to Implement Transparent Subnet Gateways</td>
<td>S</td>
</tr>
<tr>
<td>1035</td>
<td>Domain Names- Implementation and Specification (client)</td>
<td>S</td>
</tr>
<tr>
<td>1042</td>
<td>A Standard for the Transmission of IP Datagrams over IEEE 802 Networks</td>
<td>S</td>
</tr>
<tr>
<td>1191</td>
<td>Path MTU Discovery</td>
<td>S</td>
</tr>
<tr>
<td>1305</td>
<td>Network Time Protocol (Version 3) Specification, Implementation and Analysis</td>
<td>S</td>
</tr>
<tr>
<td>1519</td>
<td>Classless Inter-Domain Routing (CIDR): an Address Assignment and Analysis</td>
<td>S</td>
</tr>
<tr>
<td>1542</td>
<td>Clarifications and Extensions for the Bootstrap Protocol</td>
<td>S</td>
</tr>
<tr>
<td>1812</td>
<td>Requirements for IP Version 4 Routers</td>
<td>S</td>
</tr>
<tr>
<td>2131</td>
<td>Dynamic Host Configuration Protocol</td>
<td>S</td>
</tr>
<tr>
<td>2338</td>
<td>Virtual Router Redundancy Protocol (VRRP)</td>
<td>S</td>
</tr>
<tr>
<td>3021</td>
<td>Using 31-Bit Prefixes on IPv4 Point-to-Point Links</td>
<td>S</td>
</tr>
<tr>
<td>3046</td>
<td>DHCP Relay Agent Information Option</td>
<td>S</td>
</tr>
<tr>
<td>3069</td>
<td>VLAN Aggregation for Efficient IP Address Allocation</td>
<td>S</td>
</tr>
<tr>
<td>3128</td>
<td>Protection Against a Variant of the Tiny Fragment Attack</td>
<td>S</td>
</tr>
</tbody>
</table>

### General IPv6 Protocols

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>1886</td>
<td>DNS Extensions to support IP version 6</td>
<td>S</td>
</tr>
<tr>
<td>1981</td>
<td>Path MTU Discovery for IP version 6</td>
<td>S</td>
</tr>
</tbody>
</table>
### General IPv6 Protocols

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>2460</td>
<td>Internet Protocol, Version 6 (IPv6) Specification</td>
<td>7.8.1 7.8.1 √ 8.2.1</td>
</tr>
<tr>
<td>2461</td>
<td>Neighbor Discovery for IP Version 6 (IPv6) (Partial)</td>
<td>7.8.1 7.8.1 √ 8.2.1</td>
</tr>
<tr>
<td>2462</td>
<td>IPv6 Stateless Address Autoconfiguration (Partial)</td>
<td>7.8.1 7.8.1 √ 8.2.1</td>
</tr>
<tr>
<td>2463</td>
<td>Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification</td>
<td>7.8.1 7.8.1 √ 8.2.1</td>
</tr>
<tr>
<td>2464</td>
<td>Transmission of IPv6 Packets over Ethernet Networks</td>
<td>7.8.1 7.8.1 √ 8.2.1</td>
</tr>
<tr>
<td>2675</td>
<td>IPv6 Jumbograms</td>
<td>7.8.1 7.8.1 √ 8.2.1</td>
</tr>
<tr>
<td>2740</td>
<td>OSPF for IPv6</td>
<td>9.1 9.1</td>
</tr>
<tr>
<td>3587</td>
<td>IPv6 Global Unicast Address Format</td>
<td>7.8.1 7.8.1 √ 8.2.1</td>
</tr>
<tr>
<td>4291</td>
<td>Internet Protocol Version 6 (IPv6) Addressing Architecture</td>
<td>7.8.1 7.8.1 √ 8.2.1</td>
</tr>
</tbody>
</table>

### Border Gateway Protocol (BGP)

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>BGP Communities Attribute</td>
<td>7.8.1 7.7.1 √ 8.1.1</td>
</tr>
<tr>
<td>2385</td>
<td>Protection of BGP Sessions via the TCP MD5 Signature Option</td>
<td>7.8.1 7.7.1 √ 8.1.1</td>
</tr>
<tr>
<td>2439</td>
<td>BGP Route Flap Damping</td>
<td>7.8.1 7.7.1 √ 8.1.1</td>
</tr>
<tr>
<td>2545</td>
<td>Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing</td>
<td>7.8.1 7.8.1 √ 8.2.1</td>
</tr>
<tr>
<td>2796</td>
<td>BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)</td>
<td>7.8.1 7.7.1 √ 8.1.1</td>
</tr>
<tr>
<td>2842</td>
<td>Capabilities Advertisement with BGP-4</td>
<td>7.8.1 7.7.1 √ 8.1.1</td>
</tr>
<tr>
<td>2858</td>
<td>Multiprotocol Extensions for BGP-4</td>
<td>7.8.1 7.7.1 √ 8.1.1</td>
</tr>
<tr>
<td>2918</td>
<td>Route Refresh Capability for BGP-4</td>
<td>7.8.1 7.7.1 √ 8.1.1</td>
</tr>
<tr>
<td>3065</td>
<td>Autonomous System Confederations for BGP</td>
<td>7.8.1 7.7.1 √ 8.1.1</td>
</tr>
<tr>
<td>4360</td>
<td>BGP Extended Communities Attribute</td>
<td>7.8.1 7.7.1 7.6.1 8.1.1</td>
</tr>
<tr>
<td>4893</td>
<td>BGP Support for Four-octet AS Number Space</td>
<td>7.8.1 7.7.1 7.7.1 8.1.1</td>
</tr>
<tr>
<td>5396</td>
<td>Textual Representation of Autonomous System (AS) Numbers</td>
<td>8.1.2 8.1.2 8.1.2 8.2.1</td>
</tr>
</tbody>
</table>
### Border Gateway Protocol (BGP)

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>draft-ietf-idr-bgp4-20</td>
<td>A Border Gateway Protocol 4 (BGP-4)</td>
<td>7.8.1</td>
</tr>
<tr>
<td>draft-ietf-idr-restart-06</td>
<td>Graceful Restart Mechanism for BGP</td>
<td>7.8.1</td>
</tr>
</tbody>
</table>

### Open Shortest Path First (OSPF)

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>1587</td>
<td>The OSPF Not-So-Stubby Area (NSSA) Option</td>
<td>7.6.1</td>
</tr>
<tr>
<td>2154</td>
<td>OSPF with Digital Signatures</td>
<td>7.6.1</td>
</tr>
<tr>
<td>2328</td>
<td>OSPF Version 2</td>
<td>7.6.1</td>
</tr>
<tr>
<td>2370</td>
<td>The OSPF Opaque LSA Option</td>
<td>7.6.1</td>
</tr>
<tr>
<td>2740</td>
<td>OSPF for IPv6</td>
<td>7.8.1</td>
</tr>
<tr>
<td>3623</td>
<td>Graceful OSPF Restart</td>
<td>7.8.1</td>
</tr>
<tr>
<td>4222</td>
<td>Prioritized Treatment of Specific OSPF Version 2 Packets and Congestion Avoidance</td>
<td>7.6.1</td>
</tr>
</tbody>
</table>

### Intermediate System to Intermediate System (IS-IS)

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>1142</td>
<td>OSI IS-IS Intra-Domain Routing Protocol (ISO DP 10589)</td>
<td>✓</td>
</tr>
<tr>
<td>1195</td>
<td>Use of OSI IS-IS for Routing in TCP/IP and Dual Environments</td>
<td>✓</td>
</tr>
<tr>
<td>2763</td>
<td>Dynamic Hostname Exchange Mechanism for IS-IS</td>
<td>✓</td>
</tr>
<tr>
<td>2966</td>
<td>Domain-wide Prefix Distribution with Two-Level IS-IS</td>
<td>✓</td>
</tr>
<tr>
<td>3373</td>
<td>Three-Way Handshake for Intermediate System to Intermediate System (IS-IS) Point-to-Point Adjacencies</td>
<td>✓</td>
</tr>
<tr>
<td>3567</td>
<td>IS-IS Cryptographic Authentication</td>
<td>✓</td>
</tr>
<tr>
<td>5120</td>
<td>M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)</td>
<td>7.8.1</td>
</tr>
</tbody>
</table>
### Intermediate System to Intermediate System (IS-IS)

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>5306</td>
<td>Restart Signaling for IS-IS</td>
<td>8.3.1  8.3.1</td>
</tr>
<tr>
<td>draft-ietf-isis-igp-p2p-over-lan-06</td>
<td>Point-to-point operation over LAN in link-state routing protocols</td>
<td>✓ 8.1.1</td>
</tr>
<tr>
<td>draft-ietf-isis-ipv6-06</td>
<td>Routing IPv6 with IS-IS</td>
<td>7.5.1  8.2.1</td>
</tr>
<tr>
<td>draft-kaplan-isis-ext-eth-02</td>
<td>Extended Ethernet Frame Size Support</td>
<td>✓ 8.1.1</td>
</tr>
</tbody>
</table>

### Routing Information Protocol (RIP)

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>1058</td>
<td>Routing Information Protocol</td>
<td>7.8.1  7.6.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>2453</td>
<td>RIP Version 2</td>
<td>7.8.1  7.6.1 ✓ 8.1.1</td>
</tr>
</tbody>
</table>

### Multiprotocol Label Switching (MPLS)

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>2702</td>
<td>Requirements for Traffic Engineering Over MPLS</td>
<td>8.3.1</td>
</tr>
<tr>
<td>3031</td>
<td>Multiprotocol Label Switching Architecture</td>
<td>8.3.1</td>
</tr>
<tr>
<td>3032</td>
<td>MPLS Label Stack Encoding</td>
<td>8.3.1</td>
</tr>
<tr>
<td>3209</td>
<td>RSVP-TE: Extensions to RSVP for LSP Tunnels</td>
<td>8.3.1</td>
</tr>
<tr>
<td>3630</td>
<td>Traffic Engineering (TE) Extensions to OSPF Version 2</td>
<td>8.3.1</td>
</tr>
<tr>
<td>3784</td>
<td>Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)</td>
<td>8.3.1</td>
</tr>
<tr>
<td>3812</td>
<td>Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Management Information Base (MIB)</td>
<td>8.3.1</td>
</tr>
<tr>
<td>3813</td>
<td>Multiprotocol Label Switching (MPLS) Label Switching Router (LSR) Management Information Base (MIB)</td>
<td>8.3.1</td>
</tr>
<tr>
<td>4090</td>
<td>Fast Reroute Extensions to RSVP-TE for LSP Tunnels</td>
<td>8.3.1</td>
</tr>
<tr>
<td>4379</td>
<td>Detecting Multi-Protocol Label Switched Data Plane Failures (MPLS TE/LDP Ping &amp; Traceroute)</td>
<td>8.3.1</td>
</tr>
</tbody>
</table>
## Multicast

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>1112</td>
<td>Host Extensions for IP Multicasting</td>
<td>7.8.1, 7.7.1, ✓ 8.1.1</td>
</tr>
<tr>
<td>2236</td>
<td>Internet Group Management Protocol, Version 2</td>
<td>7.8.1, ✓ 8.1.1</td>
</tr>
<tr>
<td>2710</td>
<td>Multicast Listener Discovery (MLD) for IPv6</td>
<td>✓ 8.2.1</td>
</tr>
<tr>
<td>3376</td>
<td>Internet Group Management Protocol, Version 3</td>
<td>7.8.1, ✓ 8.1.1</td>
</tr>
<tr>
<td>3569</td>
<td>An Overview of Source-Specific Multicast (SSM)</td>
<td>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</td>
</tr>
<tr>
<td>3618</td>
<td>Multicast Source Discovery Protocol (MSDP)</td>
<td>✓ 8.1.1</td>
</tr>
<tr>
<td>3810</td>
<td>Multicast Listener Discovery Version 2 (MLDv2) for IPv6</td>
<td>✓ 8.2.1</td>
</tr>
<tr>
<td>3973</td>
<td>Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised)</td>
<td>✓</td>
</tr>
<tr>
<td>4541</td>
<td>Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches</td>
<td>7.6.1 IGMPv1/v2, 7.6.1 IGMPv1/v2/v3, MLPv1 Snooping, ✓ 8.2.1 IGMPv1/v2/v3, MLPv1 Snooping, 8.2.1 PIM-SM for IPv4/IPv6</td>
</tr>
</tbody>
</table>
## Network Management

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>1155</td>
<td>Structure and Identification of Management Information for TCP/IP-based Internets</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>1156</td>
<td>Management Information Base for Network Management of TCP/IP-based Internets</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>1157</td>
<td>A Simple Network Management Protocol (SNMP)</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>1212</td>
<td>Concise MIB Definitions</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>1215</td>
<td>A Convention for Defining Traps for use with the SNMP</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>1493</td>
<td>Definitions of Managed Objects for Bridges [except for the dot1dTpLearnedEntryDiscards object]</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>1724</td>
<td>RIP Version 2 MIB Extension</td>
<td>7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>1850</td>
<td>OSPF Version 2 Management Information Base</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>1901</td>
<td>Introduction to Community-based SNMPv2</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>2011</td>
<td>SNMPv2 Management Information Base for the Internet Protocol using SMIv2</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>2012</td>
<td>SNMPv2 Management Information Base for the Transmission Control Protocol using SMIv2</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>2013</td>
<td>SNMPv2 Management Information Base for the User Datagram Protocol using SMIv2</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>2024</td>
<td>Definitions of Managed Objects for Data Link Switching using SMIv2</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>2096</td>
<td>IP Forwarding Table MIB</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>2558</td>
<td>Definitions of Managed Objects for the Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH) Interface Type</td>
<td>✔️</td>
</tr>
<tr>
<td>2570</td>
<td>Introduction and Applicability Statements for Internet Standard Management Framework</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>2571</td>
<td>An Architecture for Describing Simple Network Management Protocol (SNMP) Management Framework</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>2572</td>
<td>Message Processing and Dispatching for the Simple Network Management Protocol (SNMP)</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>2574</td>
<td>User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>2575</td>
<td>View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)</td>
<td>7.6.1 7.5.1 ✔️ 8.1.1</td>
</tr>
<tr>
<td>RFC#</td>
<td>Full Name</td>
<td>FTOS support, per platform</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------------------------------------------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>2576</td>
<td>Coexistence Between Version 1, Version 2, and Version 3 of the Internet-standard Network Management Framework</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>2578</td>
<td>Structure of Management Information Version 2 (SMIv2)</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>2579</td>
<td>Textual Conventions for SMIv2</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>2580</td>
<td>Conformance Statements for SMIv2</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>2618</td>
<td>RADIUS Authentication Client MIB, except the following four counters:</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td></td>
<td>radiusAuthClientInvalidServerAddresses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>radiusAuthClientMalformedAccessResponses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>radiusAuthClientUnknownTypes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>radiusAuthClientPacketsDropped</td>
<td></td>
</tr>
<tr>
<td>2665</td>
<td>Definitions of Managed Objects for the Ethernet-like Interface Types</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>2674</td>
<td>Definitions of Managed Objects for Bridges with Traffic Classes, Multicast Filtering and Virtual LAN Extensions</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>2787</td>
<td>Definitions of Managed Objects for the Virtual Router Redundancy Protocol</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>2819</td>
<td>Remote Network Monitoring Management Information Base: Ethernet Statistics Table, Ethernet History Control Table, Ethernet History Table, Alarm Table, Event Table, Log Table</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>2863</td>
<td>The Interfaces Group MIB</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>2865</td>
<td>Remote Authentication Dial In User Service (RADIUS)</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>3273</td>
<td>Remote Network Monitoring Management Information Base for High Capacity Networks (64 bits): Ethernet Statistics High-Capacity Table, Ethernet History High-Capacity Table</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>3416</td>
<td>Version 2 of the Protocol Operations for the Simple Network Management Protocol (SNMP)</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>3418</td>
<td>Management Information Base (MIB) for the Simple Network Management Protocol (SNMP)</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>3434</td>
<td>Remote Monitoring MIB Extensions for High Capacity Alarms, High-Capacity Alarm Table (64 bits)</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>3580</td>
<td>IEEE 802.1X Remote Authentication Dial In User Service (RADIUS) Usage Guidelines</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>3815</td>
<td>Definitions of Managed Objects for the Multiprotocol Label Switching (MPLS), Label Distribution Protocol (LDP)</td>
<td>8.3.1</td>
</tr>
</tbody>
</table>
## Network Management (continued)

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>5060</td>
<td>Protocol Independent Multicast MIB</td>
<td>7.8.1</td>
</tr>
<tr>
<td></td>
<td>ANSI/TIA-1057 The LLDP Management Information Base extension module for TIA-TR41.4 Media Endpoint Discovery information</td>
<td>7.7.1</td>
</tr>
<tr>
<td></td>
<td>draft-ietf-tacacs-02 The TACACS+ Protocol</td>
<td>7.6.1</td>
</tr>
<tr>
<td></td>
<td>draft-ietf-idr-bgp4-mib-06 Definitions of Managed Objects for the Fourth Version of the Border Gateway Protocol (BGP-4) using SMIv2</td>
<td>7.8.1</td>
</tr>
<tr>
<td></td>
<td>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</td>
<td>7.8.1</td>
</tr>
<tr>
<td></td>
<td>IEEE 802.1AB Management Information Base module for LLDP configuration, statistics, local system data and remote systems data components.</td>
<td>7.7.1</td>
</tr>
<tr>
<td></td>
<td>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)</td>
<td>7.7.1</td>
</tr>
<tr>
<td></td>
<td>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)</td>
<td>7.7.1</td>
</tr>
<tr>
<td></td>
<td>ruzin-mstp-mib-02 (Traps) Definitions of Managed Objects for Bridges with Multiple Spanning Tree Protocol</td>
<td>7.6.1</td>
</tr>
<tr>
<td></td>
<td>sFlow.org sFlow Version 5</td>
<td>7.7.1</td>
</tr>
<tr>
<td></td>
<td>sFlow.org sFlow Version 5 MIB</td>
<td>7.7.1</td>
</tr>
<tr>
<td></td>
<td>FORCE10-BGP4-V2-MIB Dell Force10 BGP MIB (draft-ietf-idr-bgp4-mibv2-05)</td>
<td>7.8.1</td>
</tr>
<tr>
<td></td>
<td>FORCE10-CIDR-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)</td>
<td>7.6.1</td>
</tr>
</tbody>
</table>
## Network Management (continued)

<table>
<thead>
<tr>
<th>RFC#</th>
<th>Full Name</th>
<th>FTOS support, per platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>FORCE10-C</td>
<td>Dell Force10 C-Series Enterprise Chassis MIB</td>
<td>7.5.1</td>
</tr>
<tr>
<td>FORCE10-I</td>
<td>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 &quot;show interfaces&quot; output)</td>
<td>7.6.1 7.6.1 7.6.1 8.1.1</td>
</tr>
<tr>
<td>FORCE10-L</td>
<td>Dell Force10 Enterprise Link Aggregation MIB</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>FORCE10-C</td>
<td>Dell Force10 E-Series Enterprise Chassis MIB</td>
<td>✓ 8.1.1</td>
</tr>
<tr>
<td>FORCE10-C</td>
<td>Dell Force10 File Copy MIB (supporting SNMP SET operation)</td>
<td>7.7.1 7.7.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>FORCE10-M</td>
<td>Dell Force10 Monitoring MIB</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>FORCE10-P</td>
<td>Dell Force10 Product Object Identifier MIB</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>FORCE10-S</td>
<td>Dell Force10 S-Series Enterprise Chassis MIB</td>
<td>7.6.1</td>
</tr>
<tr>
<td>FORCE10-S</td>
<td>Dell Force10 Structure of Management Information</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>FORCE10-S</td>
<td>Dell Force10 System Component MIB (enables the user to view CAM usage information)</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>FORCE10-T</td>
<td>Dell Force10 Textual Convention</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
<tr>
<td>FORCE10-T</td>
<td>Dell Force10 Trap Alarm MIB</td>
<td>7.6.1 7.5.1 ✓ 8.1.1</td>
</tr>
</tbody>
</table>

Standards Compliance | 1133
MIB Location

Dell Force10 MIBs are under the **Force10 MIBs** subhead on the **Documentation** page of iSupport:


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:


If you have forgotten or lost your account information, contact Dell Force10 TAC for assistance.
Index

Numerics
10/100/1000 Base-T Ethernet line card, auto negotiation 508
100/1000 Ethernet interfaces
  port channels 482
4-Byte AS Numbers 186
802.1AB 1123
802.1D 1123
802.1p 1123
802.1p/Q 1123
802.1Q 1123
802.1s 1123
802.1w 1123
802.1X 1123
802.3ab 1123
802.3ac 1123
802.3ad 1123
802.3ae 1123
802.3af 1123
802.3ak 1123
802.3i 1123
802.3u 1123
802.3x 1123
802.3z 1123

A
AAA (Accounting, Authentication, and Authorization)
  security model 877
AAA Accounting 877
  command 878
  suppress null-username command 878
AAA Authentication
  authentication and authorization, local by default 883
  command 882
  configuring 881
  enable 882
  enable command 882
  enable method 881
  line method 881
  local method 881
  none method 881
  radius method 881
  tacacs+ 881
AAA Authorization
  AAA new-model enabled by default 883
ABR
  definition 739
  access-class 896
ACL
  definition 101
  IP ACL definition 101
  RADIUS 889
  ANSI/TIA-1057 643
Applying an ACL to Loopback 119
Area Border Router. See ABR.
AS 172
  support 194
AS-PATH ACL
  “permit all routes” statement 227
  configuring 213
AS_PATH attribute
  using 212
Authentication
  implementation 880
  TACACS+ 895
  VTY 910
Authorization
  TACACS+ 895
  VTY 910
auto negotiation 508
auto negotiation, line card 508
Auto-command 890

B
Balancing, load 489
Bare Metal Provisioning 251
  prerequisites 252
Base VLAN 794
BGP 172
  Attributes 177
  Autonomous Systems 172
  best path criteria 177
  changing COMMUNITY numbers in a path 220
  changing how MED attribute is used 222
  changing LOCAL_PREF default value 222
  changing the LOCAL_PREF default values for routes on a router 222
  clearing route dampening information 232
  Communities 177
  configuring a route reflector 227
  configuring an AS-PATH ACL 213
  configuring an IP community list 217
  configuring BGP confederations 229
  configuring BGP timers 233
  configuring route flap dampening 230
  configuring the router as next hop 223
  creating a peer group 201, 202
  default 193, 219
  Distance defaults 193
enabling a peer group 203
establishing BGP process 195
External BGP requirements 195
Fast External Fallover 193
filtering routes based on AS-PATH 226
filtering routes using a route map 226
filtering routes using IP Community list 219
filtering routes using prefix lists 225
graceful restart tasks 210
graceful restart, default role 210
graceful restart, default setting 193
graceful restart, enabling 211, 746
graceful restart, hot failover actions 210
graceful restart, implementing by neighbor or BGP peer-group 211
Implementing with FTOS 184
inbound and outbound policies 224
Internal BGP requirements 195
KEEPALIVE messages 194
LOCAL_PREF default value 193
MULTI_EXIT_DISC default value 193
Neighbor Adjacency changes 193
neighbors 194
resetting a BGP neighbor 232
route dampening information 229
Route Flap Damping Parameters 193
Route Reflectors 176
route reflectors 227
sending the COMMUNITY attribute 219
Sessions and Peers 174
specifying next hop address 223, 224
Timers defaults 193
timers negotiation 233
viewing all BGP path attributes 212
viewing the BGP configuration 196
viewing the status of BGP neighbors 197
viewing the status of peer groups 203
BPDU 869
Bridge MIB
  STP implementation 1004
Bridge Protocol Data Units. See BPDU.

C
CAM Profiling, When to Use 276
cam-acl 550
cam-profile 548
CLI
  case sensitivity 35
  editing commands 35
  partial keywords 35
CLI Modes
  LINE 30
COMMUNITY attribute
  changing in a path 220
  default 219
  NO_ADVERTISE 217, 220
  NO_EXPORT 217, 220
  NO_EXPORT_SUBCONFED 217, 220
Community list
  configuring 217
  community port 794
  community VLAN 793
  Console terminal line 63
  Control Plane Policing (CoPP) 291
coredumps 371
crypto key generate 899
C-Series and S-Series load-balancing 491
CSNP 574

D
debug ip ssh 899
Default VLAN
  changing the VLAN id 1054
  implementation 1054
  Layer 2 mode 1054
  remove interface 1059
  remove untagged interface 1055
  untagged interfaces 1054, 1057
default-information originate (OSPF IPv6) 758
directed broadcast 520
disable-on-sfm failure 500
display parameter 38
distribution algorithms 488
DNS 521
Document conventions 27
dynamic hostname exchange 573

E
ECMP
  link bundle monitoring 397
extended IP ACL 102

F
Fast Convergence after MSTP-Triggered Topology Changes 468
fast-convergence
  OSPF 742
File Transfer Protocol. See FTP
  flowcontrol 505
Force 10 Resilient Ring Protocol 419
forward delay 870, 1009
FRRP 419
FRRP Master Node 419
FRRP Transit Node 419
FTOS 727
FTP 61
  configuring client parameters 62
  configuring server parameters 62
  enabling server 61
  using VLANs 61

G
GARP VLAN Registration Protocol (GVRP) 431
grep option 37
grep pipe option 800
GVRP (GARP VLAN Registration Protocol) 431

H
Hash algorithm 491
hash algorithm, LAG 484, 486, 488
hashing algorithms for flows and fragments 488
hello time 870, 1009
host port 794
Hot Lock ACL 102
Hybrid ports 1058
hybrid ports 1061

I
I-D (Internet Draft) Compliance 1124
Idle Time 889
IEEE 802.1q 431
IEEE 802.1Q tag 1055
IEEE Compliance 1123
IEEE Standard 802.3ad 481
IGMP
  viewing which interfaces are IGMP-enabled 463
implicit deny 101
Interface modes
  Layer 2 473
  Layer 3 473
Interface Range Macros 495
Interface types
  100/1000 Ethernet 474
  10-Gigabit Ethernet 474
  1-Gigabit Ethernet 474
  Ethernet 469
  Loopback 474
  management 474
  Management Ethernet interface 473
  Port Channel 474
  SONET 469
  VLAN 474
interface types
  null interface 474
interfaces
  auto negotiation setting 508
  clearing counters 514
  commands allowed when part of a port channel 484
  configuring secondary IP addresses 517
  determining configuration 475
  member of port channel 486
  viewing Layer 3 interfaces 471
  viewing only Layer 2 interfaces 511
Inter-VLAN routing 479
  considerations 479
IP ACLs
  applying ACL for loopback 119
  applying IP ACL to interface 115
  configuring extended IP ACL 111
  configuring extended IP ACL for TCP 112
  configuring extended IP ACL for UDP 112
  configuring filter without sequence number 113
  configuring standard IP ACL 108, 110
  deleting a filter 109, 110
  extended IP ACLs 102, 111
  standard IP ACL 102
types 102
  viewing configuration 109
IP addresses
  assigning IP address to interface 475
  assigning to interface 517
  assigning to port channel 488
  assigning to VLAN 1060
  composition 515
  configuring static routes 518
IP fragmentation 503
IP hashing scheme 491
ip local-proxy-arp command 800
IP MTU
  configuring 506
  maximum size 503
IP prefix lists
  "permit all" statement 122
  applying to OSPF routes 124
  applying to RIP routes 124
  configuring filter 121
  configuring filters without seq command 122
  definition 120
  deleting filter 122, 123
  implementation 121
permit default route only 122
rules 121, 225
using the le and ge parameters 120
IP routing
  VLANs 1055
  ip scp topdir 899
  ip ssh authentication-retries 899
  ip ssh connection-rate-limit 899
  ip ssh hostbased-authentication enable 900
  ip ssh password-authentication enable 900
  ip ssh pub-key-file 900
  ip ssh rhostsfile 900
  ip ssh rsa-authentication 900
  ip ssh rsa-authentication enable 900
  ip ssh server command 898
  IP traffic load balancing 491
  IP version 4 515
IPsec 762
IS-IS
  area address 570
defaults 574
dynamic hostname exchange 573
  Intermediate System to Intermediate System 569
  Level 1 570
  Level 1-2 570
  Level 2 570
  NET 570
  N-selector 570
  system address 570
IS-IS
  graceful-restart 579
  redistribute OSPF 215, 589, 590
  IS-IS TLVs 573
  ISO/IEC 10589 571
  isolated port 794
  Isolated VLAN 794

J
Jumpstart mode
default in BMP 2.0 256

L
LAG hash algorithm 484, 486, 488
LAG. See Port Channels.
Layer 2 mode
  configuring 474
Layer 2 protocols
  configuring 474
Layer 3 mode
  enable traffic 475
  Layer 3 protocols
  configuring 475
Level 1
  definition 570
  using NET 570
Level 1-2
  definition 570
Level 2
  definition 570
  using NET 570
line card, auto negotiation 508
Link Aggregation Group 481
link debounce interface 499
Link Debounce Timer 500, 503
Link Layer Discovery Protocol (LLDP) 639
link MTU
  configuring 506
  maximum size 503
Link State Advertisements. See LSAs.
Link State PDUs. See LSPs.
LLDP 639
LLDP-MED 643
load balancing 488
load-balance
  command combinations 490
criteria 488
hash algorithm 488
LOCAL_PREF attribute
  changing default value 222
  changing value for a router 222
logging
  changing settings 57
  consolidating messages 60
  including timestamp 60
  UNIX system logging facility 58
Loopback interface
  configuring 480
  defaults 474
definition 480
deleting interface 480
viewing configuration 480
Loopback, Configuring ACLs to 119
LSAs 720
  AS Boundary 727
  AS External 728
  Network 727
  Network Summary 727
  NSSA External 728
  Opaque Area-local 727
  Opaque Link-local 728
  Router 727
  types supported 727
Index | 1139

LSPs 570

M
MAC hashing scheme 491
management interface 474
 accessing 477
configuring a management interface 477
 configuring IP address 477
definition 476
IP address consideration 477
management interface, switch 473
max age 870, 1009
MBGP 236
Member VLAN (FRRP) 421
MIB Location 1134
minimum oper up links in a port channel 487
mirror, port 785, 1039
 remote port mirroring 1040
monitor interfaces 496
MSDP 663
MT IS-IS 571
MT IS-IS TLVs 573
MTU
 configuring MTU values for Port Channels 506
 configuring MTU values for VLANs 506
definition 503
IP MTU
 configuring 506
 maximum size 503
link MTU
 configuring 506
 maximum size 503
MTU Size, Configuring on an Interface 506
MULTI_EXIT_DISC attribute
 changing 222
default value 193
Multi-Topology IS-IS 571

N
NET 570
 area address 570
 length 570
 N-selector 570
 system address 570
Network Entity Title. See NET.
NIC teaming 626
no-more 38
no-more parameter 38
NSAP addresses 570
NTP
 configuring authentication 1028
 configuring source address 1028
null 474
null interface
 available command 480
definition 480
 entering the interface 480
 information 474

O
Open Shortest Path First 720
OSFP Adjacency with Cisco Routers 731
OSPF 720
 backbone area 736
 changing authentication parameters 745
 changing interface parameters 744
 configuring a passive interface 741
 configuring a stub area 739
 configuring a stub-router advertisement 740
 configuring network command 736
 configuring virtual links 748
 debugging OSPF 752, 773
default 732
disabling OSPF 734, 736
 enabling routing 733
 maximum metric in LSAs 740
 redistributing routes 749
restarting OSPF 734, 736
 router ID 737
 using loopback interfaces 738
 using prefix lists 749
 viewing configuration of neighboring router 751, 772
 viewing interface areas 737
 viewing virtual link configuration 748
OSPFv23
 redistribute routes 758
OSPFv3
 authentication 762
 configuration 754
 configure a passive interface 757
default route 758
 graceful restart 759
 viewing configuration of graceful restart 760
Output of show interfaces switchport 918

P
packet-based hashing 488
passwords
 configuring password 884
PDU 570
port channel
definition 481
port channel (LAG), configure 483
port channel, minimum oper up links 487
Port Channels
  configuring MTU values 506
  member of VLANs 1056
Port channels
  benefits 481
  defaults 474
port channels
  adding physical interface 484
  assigning IP address 488
  commands allowed on individual interfaces 484
  configuring 483
  containing 100/1000 and GE interfaces 482
  IP routing 488
  placing in Layer 2 mode 483
  reassigning interfaces 486
port cost 870, 1009
port mirror 785, 1039
  remote 1040
Port Monitoring Commands
  Important Points to Remember 785
port priority 870, 1009
port types (private VLAN) 794
port-based VLANs 1055
  assigning IP address 1060
  benefits 1055
  creating VLAN 1056
  definition 1055
  deleting VLAN 1057
  enabling tagging 1058
  interface requirements 1055
  IP routing 1055
  number supported 1055
  remove interface 1059
  remove tagged interface 1058
  tag frames 1058
  tagged interface
    member of multiple VLANs 1058
Portfast 871, 1010
PPP 473
Prefix list. See IP Prefix list.
Primary VLAN 794
Private VLAN (PVLAN) 793
private-vlan mapping secondary-vlan command 795
Privilege Level 890
privilege levels
  and CLI commands 884
  definition 883
  number of levels available 883
privilege level 0 definition 883
privilege level 1 definition 883
privilege level 15 definition 883
promiscuous port 794
Protocol Data Units. See PDU.
Proxy ARP
  default 524
Q
QoS
  dot1p queue numbers 818
  dot1p-priority values 818
  purpose of input policies 826
  rate limit outgoing traffic 821
QoS (Quality of Service) chapter 815
QSFP port splitting 498
Quality of Service (QoS) chapter 815
R
RADIUS
  changing an optional parameter 892
  configuration requirements 888
  configuring global communication parameter 892
  debugging RADIUS transactions 893, 895
  definition 888
  deleting a server host 892
  specifying a server host 891, 896
  viewing RADIUS configuration 893
RADIUS authentication 883
RADIUS Authentication and Authorization 889
radius-server host command 882
rate-interval command 511
redistribute
  ROUTER ISIS 215, 589, 590
remote port mirroring 1040
RFC 1058 841
RFC 1493 1004
RFC 1858 897
RFC 2138 888
RFC 2338 1098
RFC 2453 841
RFC 3128 897
RFC 791 515, 516
RFC 959 61
RFC Compliance 1124
RIP
  adding routes 845
  auto summarization default 842
  changing RIP version 846
  configuring interfaces to run RIP 844
Tag Header 1055
definition 1055
Ethernet frame 1056
tagged interfaces
member of multiple VLANs 1058
TCP Tiny and Overlapping Fragment Attack, Protection Against 897
TDR (Time Domain Reflectometer) 497
Telnet 896
Telnet Daemon, Enabling and Disabling 903
Time Domain Reflectometer (TDR) 497
Time to Live (TTL) 655
Trace list 904
Trace lists
configuring a trace list 905
configuring filter without sequence number 907
configuring trace list for TCP 905
configuring trace list for UDP 906
trunk port 794
TTL 655

U
user level
definition 883
user name
configuring user name 884
username command 886

V
virtual IP addresses 1102
Virtual LANs. See VLAN.
Virtual Router Identifier. See VRID.
Virtual Router Redundancy Protocol. See VRRP.
virtual-ip
Important Things to Remember 477
VLAN
isolated 794
primary 794
private
base 794
secondary
community 794
isolated 794
VLAN configuration, automatic 431
VLAN Protocol Identifier 1056
VLAN types 793
VLAN types (private VLAN) 793
VLANs 474, 1053
adding a port channel 487
adding interface 1055
assigning IP address 1060
benefits 1053
configuring MTU values 506
defaults 1054
definition 1053
enabling tagging 1058
FTP 61, 1060
hybrid ports 1058
IP routing 479, 1055
Layer 2 interfaces 1057
port-based 1055
removing a port channel 487
removing tagged interface 1055, 1058
SNMP 1060
tagged interfaces 1056, 1057
TFTP 1060
untagged interfaces 1057
viewing configured 1057
VLSM 515
VLSM (Variable Length Subnet Masks) 841
VLT
ports
orphan 1069
VRRP 1097
advertisement interval 1107
benefits 1099
changing advertisement interval 1107
configuring priority 1104
configuring simple authentication 1105
definition 1097
disabling preempt 1106
MAC address 1097
monitoring interface 1109
simple authentication 1105
transmitting VRRP packets 1102
virtual IP addresses 1102
virtual router 1101
VRID 1097, 1101
VTY lines
access class configuration 910
access classes and TACACS+ servers 895
assigning access classes by username 910
deny all, deny incoming subnet access-class application 911
deny10 ACLs, support for remote authentication and authorization 896
line authentication, support for 911
local authentication and authorization 910
local authentication and authorization, local database
source of access class 910
radius authentication, support for 911
remote authentication and authorization 895
remote authentication and authorization, 10.0.0.0 subnets 896
remote authentication and local authorization 911
TACACS+ authentication, support for local authorization 911

W
When to Use CAM Profiling 276