Understanding Preboot Networking on Dell PowerEdge Servers

This paper explains how the NIC and BIOS work together to configure and provide PXE, iSCSI, and FCoE boot support in UEFI and Legacy boot modes.

William Moyes
September 2014
# Table of contents

Executive summary ................................................................................................................................. 4
1  Understanding network boot ...................................................................................................................... 5
   1.1  Core concepts .................................................................................................................................. 5
   1.2  Legacy BIOS boot mode ................................................................................................................ 5
   1.3  UEFI boot mode ........................................................................................................................... 6
2  NIC settings in a hybrid BIOS ................................................................................................................ 8
3  Differences in UEFI mode between 12\textsuperscript{th} and 13\textsuperscript{th} generation PowerEdge servers ................................................................. 9
   3.1  UEFI behavior on the 12\textsuperscript{th} generation PowerEdge servers .............................................. 9
   3.2  UEFI behavior on the 13\textsuperscript{th} generation PowerEdge servers .............................................. 9
      3.2.1  Deferred network stack startup .......................................................................................... 9
      3.2.2  Network Settings menu ..................................................................................................... 10
A  Acronyms and Terminology .................................................................................................................. 12
Executive summary

The PowerEdge family of servers provide system administrators a great degree of flexibility in configuring network boot. However, the multitude of options may be confusing unless you have an understanding of how preboot networking operates in the UEFI boot and Legacy BIOS boot modes.

This white paper explains the preboot networking theory of operation for the Dell 12th and 13th generation PowerEdge servers. Specifically this white paper:

- Serves as an introduction to both UEFI and BIOS preboot networking.
- Explains the vital differences between UEFI and Legacy BIOS boot modes.
- Describes the differences in UEFI preboot networking between the 12th and 13th generation PowerEdge servers.

Note: This white paper does not replace the BIOS Setup Guide for Dell 13th Generation PowerEdge Servers, nor does this white paper cover configuring a datacenter’s infrastructure to network boot a PowerEdge server.

Key takeaways:

- Dell 12th and 13th generation PowerEdge servers use a hybrid BIOS that supports both UEFI and Legacy BIOS boot modes.
- In the Legacy BIOS boot mode, the NIC owns all aspects of networking and network booting. The NIC vendor’s setup pages control which features are enabled and how they work. By default, the first NDC or LOM port has PXE enabled.
- In the UEFI boot mode, the BIOS owns the networking stack, similar to the operating system.
  - 12th generation PowerEdge servers automatically create an IPv4 PXE boot option for every network interface visible to the BIOS. Newer BIOS’s contain a setup option to create IPv6 PXE boot options instead.
  - 13th generation PowerEdge servers add a Networking Settings setup page when in UEFI boot mode to control UEFI network boot behavior. By default, an IPv4 PXE boot option is created for the first NDC port.
  - Some CNA/HBA NIC’s can be configured to emulate a disk controller instead of, or in addition to providing a NIC interface. To the BIOS these emulated disks look like any other RAID or disk controller. The emulation settings are controlled by the NIC’s setup pages.
1 Understanding network boot

1.1 Core concepts

There are two approaches to network booting - Program Download and Disk Emulation.

Program Download involves retrieving a single executable over the network and executing it. The downloaded program is usually not a complete OS, but rather a bootstrap program (such as GRUB or WDS NBP). The downloaded program uses the network interface to retrieve the remainder of the operating system over the network. PXE is an example of Program Download.

Disk Emulation uses software, or a combination of hardware and software to emulate a physical disk. Reads and writes to the disk are converted into network packets. From the BIOS's perspective there is very little difference between booting using network Disk Emulation versus booting from a RAID adapter. Examples of Disk Emulation include FCoE and iSCSI.

1.2 Legacy BIOS boot mode

The original system BIOS had no concept of networking or network booting. However, the system BIOS would scan for and execute code contained in an adapter card’s ROMs. NIC vendors leveraged this capability to add network boot features to their products. Out of necessity NIC vendors placed the entire preboot network stack into the NIC’s ROM. Today it is not uncommon for an adapter card’s ROM to carry Legacy BIOS boot mode support for PXE, iSCSI, and FCoE.

When the NIC’s legacy ROM is executed by the BIOS, the ROM typically displays a copyright banner and a hot-key prompt. If the hot-key is pressed the ROM activates the setup menu to allow the administrator to configure the NIC’s behavior.

![A typical Legacy ROM banner](image)

Before the ROM completes running, the ROM replaces the BIOS’s service routines for either the disk services (INT13) or boot vector (INT19) with a pointer to itself. This allows the ROM to emulate a bootable

---

1 The exact behavior of the ROM will vary from one NIC vendor to another.
2 Modern add-in devices, like those found in the 12th and 13th generation PowerEdge servers cooperate with the BIOS for control of disk and boot vector to allow the BIOS to create and maintain a Boot Order list.
disk, or to regain control at the end of POST so the card’s ROM can download and execute a boot program instead of booting from a local disk.

1.3 UEFI boot mode

The industry is moving towards replacing the legacy BIOS interfaces with UEFI (Unified Extensible Firmware Interface). UEFI defines the interfaces between the platform, add-in cards, and the OS. UEFI provides better boot-times, greater flexibility, and fewer limitations when compared to legacy BIOS.

The 13th generation PowerEdge server represents Dell’s third generation of UEFI support in the PowerEdge family. The 11th, 12th, and 13th generation PowerEdge servers use a hybrid BIOS that supports both legacy boot mode for backwards compatibility in addition to providing UEFI boot mode.

The UEFI specification defines:

- a general purpose TCP/IP network stack for pre-OS use (IPv4, IPv6, TCP, UDP, DHCP, and so on).
- interfaces for leveraging VLAN support.
- APIs for PXE booting UEFI operating systems over both IPv4 and IPv6 networks along with support for many other networking technologies.

The UEFI spec provides system vendors latitude with respect to which features are supported natively by their BIOS’s. Dell chose to incorporate the TCP/IP network stack, VLAN, and UEFI PXE boot support directly into the 12th and 13th generation PowerEdge BIOS3.

Since the UEFI network stack is part of the BIOS instead of the NIC’s ROM, the setup options that control the behavior of UEFI network booting are found in the system’s setup menu, not the NIC’s.

The built-in networking stack is of little value unless the NIC provides a driver to give the UEFI network stack a means to send and receive packets. Most modern NICs contain both a legacy ROM image and a UEFI driver in their ROMs. The BIOS automatically scans for and loads the UEFI driver for each card. On the 12th and 13th generation PowerEdge servers the scan occurs on every boot regardless of the boot mode. Unlike legacy ROMs, UEFI drivers load without displaying a banner.

UEFI drivers and the BIOS communicate with each other using standardized APIs called UEFI protocols. Examples of protocols include NII/UNDI4 for network communication, Blocklo for disk storage, DriverHealth for device status, and Human Interface Infrastructure (HII) for configuration. An instance of a protocol is called an interface.

---

3 Early 12th generation PowerEdge server BIOS’s only supported IPv4. BIOS updates released after August 2013 add IPv6 support. All 13th generation PowerEdge servers support IPv6.

4 Technically the NII is the UEFI protocol, the UNDI is something else, but that is beyond the scope of this white paper.
A NIC’s UEFI driver produces one NII/UNDI interface for each physical network port on the adapter as seen in the illustration above. The BIOS’s network stack uses the NII/UNDI interfaces as a means to send and receive network packets\(^5\). The NIC’s UEFI driver can create setup pages using the HII protocols.

**Note:** Some network adapters (specifically CNAs) can ‘act’ like disk controllers instead of, or in addition to acting like network adapters. The UEFI drivers for these devices expose a SCSI or Blockio protocol instead of or in addition to a NII/UNDI protocol. The behavior of a CNA may be user selectable from the NIC’s setup menu.

Since the disk emulation feature of a CNA is under control of the NIC, all settings related to configuring the network aspects of the emulated disk are handled by the NIC’s setup pages. A CNA’s emulated disk appears to BIOS as 3rd party disk controller.

Only network adapters that expose an NII/UNDI protocol are visible to the UEFI network stack. Lifecycle Controller and BIOS network features (such as PXE booting, and FTP firmware updates) require a NII/UNDI protocol.

\(^5\) It is also possible for the NIC to directly produce a SNP or MNP, but most NIC’s keep it simple by creating a NII/UNDI.
NIC settings in a hybrid BIOS

Regardless of which boot mode the system is set to, the BIOS uses one of the two UEFI aware setup engines. By default, the graphical setup engine is used, however serial console redirection can force the text mode setup engine to be used. Both setup engines provide the administrator with the same capabilities, and both support the UEFI Human Interface Infrastructure (HII).

HII allows a UEFI driver to insert its configuration pages into the System Setup. On Dell PowerEdge servers these pages can be found under the Device Settings menu. Even though these pages are displayed by System Setup, the NIC’s configuration is still physically stored on the NIC itself, and the settings only impact the NIC’s hardware and software.

Dell has worked with NIC vendors to ensure feature parity between the NIC vendor’s legacy hotkey based setup and the setup they offer using UEFI. Certain vendors have also eliminated their legacy hotkey menus altogether.

**Note:** A common mistake is attempting to configure UEFI PXE boot by making changes to a NIC’s settings using the Device Settings menu or using the NIC’s hot-key based setup pages.

Remember, the NIC’s setup pages only control the behavior of the NIC and the NIC’s software. In UEFI the NIC does not own PXE booting or the network stack like it did in legacy mode. Most NIC’s use wording in their setup pages to indicate this such as “Legacy Boot Protocol”.

The situation becomes more confusing in the case of FCoE and iSCSI. Some NIC’s like CNAs will emulate a disk controller in UEFI mode, but many do not. For further information refer to the NIC’s documentation.

NICs that retain their legacy ROM based setup hotkey menus share the same design and limitations mentioned above, namely the configuration is physically stored on the NIC itself. The settings stored on the NIC impact the NIC itself and its onboard software. However, in legacy boot mode the NIC’s software stack is the only software that is relevant for network booting.
3 Differences in UEFI mode between 12\textsuperscript{th} and 13\textsuperscript{th} generation PowerEdge servers

Legacy network boot behavior is largely unchanged between the 12\textsuperscript{th} and 13\textsuperscript{th} generation of PowerEdge servers. One notable change is that legacy NIC ROMs no longer run when in UEFI mode on a 13\textsuperscript{th} generation PowerEdge servers. Therefore, on 13\textsuperscript{th} generation PowerEdge servers the hotkey menus is not visible when in UEFI Boot mode.

3.1 UEFI behavior on the 12\textsuperscript{th} generation PowerEdge servers

The BIOS on the 12\textsuperscript{th} generation PowerEdge servers connect the UEFI network stack to every available network interface on every boot. It also creates an IPv4 PXE boot option for every network interface in the system’s boot sequence list. This means that a server with 8 network ports will have 8 PXE boot options in its boot list.

Older BIOS’s only supported IPv4, however, the latest BIOS’s for 12\textsuperscript{th} generation PowerEdge servers includes added support for IPv6. A setup option is available under System BIOS -> Boot Settings -> UEFI Boot Settings menu to convert the IPv4 PXE boot options into IPv6 PXE boot options.

\textbf{Note:} The IPv4/IPv6 PXE boot setting is the last option on the page therefore you may need to scroll to the bottom of the page to see it.

3.2 UEFI behavior on the 13\textsuperscript{th} generation PowerEdge servers

The BIOS on the 13\textsuperscript{th} generation PowerEdge servers has a more flexible, extensible, and faster approach to network booting.

3.2.1 Deferred network stack startup

The BIOS utilizes UEFI’s support for deferred driver startup. Instead of starting the whole network stack on every NIC on every boot, initialization of the NIC port is deferred until it is actually used.

As an example, a system with 8 NIC ports that boots from the hard disk does not initialize any of the network interfaces during the boot process since none of the NIC ports are needed to boot from the hard disk. However, if that same system is rebooted and F12 is pressed to PXE boot then the network stack only starts on the NIC’s that the BIOS attempted to PXE boot from.

If a pre-boot UEFI application desires networking support on a particular interface the application can perform a ConnectController() on the NICs UEFI handle. For example the UEFI Shell will connect all drivers to all handles when the ‘connect –r’ command is issued. This is standard UEFI behavior.
3.2.2 Network Settings menu

The Network Settings menu appears under BIOS settings when the BIOS on a 13th generation PowerEdge server boots in an UEFI mode. This setup page allows you to tailor the UEFI network boot options for the system.

The BIOS on the 13th generation PowerEdge servers allow you to enable and configure up to four PXE boot options. As seen in Figure 3, the phrase ‘PXE Device’ describes these boot options even though the boot options are a software construct.

Figure 3  UEFI Network Settings Page on a 13th generation PowerEdge Server
Understanding Preboot Networking on Dell PowerEdge Servers

Figure 4  PXE Settings Page on a 13th generation PowerEdge Server

The settings page for each PXE device allows you to select the physical network interface (NIC and port), IP protocol, and VLAN settings. The enabled PXE boot options appear in the UEFI boot sequence just like any disk or other bootable item.

You can access the UEFI boot sequence in the Boot Settings > Boot Option Setting menu. The UEFI boot sequence lists all active bootable targets including disks, Optical Drives, bootable USB devices, and PXE boot devices. From this menu you can customize the order in which the boot options are attempted. Note that changes made in the Network Settings menu may not be reflected in the Boot Settings menu until after the system is rebooted.

In the example below, two PXE ‘devices’ were enabled. The first uses the integrated NIC’s Port 1 in IPv4 mode, and the second uses the same physical port in IPv6 PXE mode. You can confirm the behavior of each PXE device by examining the settings in the Network Settings menu. The NIC and port information is also displayed in the boot sequence list.

Figure 5  UEFI Boot Sequence list showing two PXE boot ‘devices’

Other scenarios are also supported, including falling back to a different NIC, or attempting to PXE boot the same port but with a different VLAN tag.
**Acronyms and Terminology**

**BIOS**: Basic Input/Output System. The computer’s BIOS contains programs stored on a flash memory chip. The BIOS initializes the systems and loads the operating system.

**CNA**: Converged Network Adapter. A NIC that provides both general networking and HBA capabilities for protocols such as iSCSI or FCoE.

**FCoE**: Fibre Channel over Ethernet. A network protocol that uses Ethernet physical interfaces and framing to convey Fibre Channel messages.

**GRUB**: Erich Boleyn’s ‘GRand Unified Bootloader’ for Linux. A bootstrap application used in many Linux distributions.

**HBA**: Host Bus Adapter. A device that connects a systems host bus (for example a PCIe slot) to SCSI, Fibre Channel, or a similar medium.

**iSCSI**: A protocol to carry SCSI disk commands over TCP/IP.

**Legacy BIOS**: The de facto 16-bit interface between the OS and the platform firmware. It is based upon the original PC BIOS.

**LOM**: LAN on Motherboard. Network interfaces that are integrated directly into the systems motherboard, as opposed to a plug-in NIC.

**NDC**: Network Daughter Card. Network interfaces that exist on a daughter card as opposed to being integrated directly into the motherboard or added using a plug-in NIC. Many PowerEdge servers use a NDC instead of integrating the LAN support into the motherboard (see LOM).

**NIC**: Network Interface Card. Typically a plug-in PCIe network adapter, but may also refer to NDC or LOM’s.

**POST**: Power-On Self-Test. When the computer is turned on, the POST tests various system components, such as RAM, the hard-disk drives, and the keyboard. POST is synonymous with the BIOS’s startup process.

**PXE**: Preboot eXecution Environment. A specification that allows booting from the network by identifying and then downloading the correct boot image for a system.

**UEFI**: Unified Extensible Firmware Interface. An industry standard specification that defines a software interface between an operating system and platform firmware. UEFI is replacing Legacy BIOS as the software interface to the OS.

**VLAN**: Virtual Local Area Network. A technology that uses IEEE 802.1Q tagging to identify packets on a single physical connection as belonging to separate virtual networks.

**WDS NBP**: Windows Deployment Services Network Bootstrap Program