Marvell® FastLinQ® Ethernet iSCSI Adapters and Ethernet FCoE Adapters

5740/57810/57800 Adapters and other 57*xx* and 57*xxx* Adapters

User's Guide

Third party information brought to
you courtesy of Dell.

Doc No. BC0054508-00 Rev. R

January 21, 2021

THIS DOCUMENT AND THE INFORMATION FURNISHED IN THIS DOCUMENT ARE PROVIDED “AS IS” WITHOUT ANY WARRANTY. MARVELL AND ITS AFFILIATES EXPRESSLY DISCLAIM AND MAKE NO WARRANTIES OR GUARANTEES, WHETHER EXPRESS, ORAL, IMPLIED, STATUTORY, ARISING BY OPERATION OF LAW, OR AS A RESULT OF USAGE OF TRADE, COURSE OF DEALING, OR COURSE OF PERFORMANCE, INCLUDING THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.

This document, including any software or firmware referenced in this document, is owned by Marvell or Marvell's licensors, and is protected by intellectual property laws. No license, express or implied, to any Marvell intellectual property rights is granted by this document. The information furnished in this document is provided for reference purposes only for use with Marvell products. It is the user's own responsibility to design or build products with this information. Marvell products are not authorized for use as critical components in medical devices, military systems, life or critical support devices, or related systems. Marvell is not liable, in whole or in part, and the user will indemnify and hold Marvell harmless for any claim, damage, or other liability related to any such use of Marvell products.

Marvell assumes no responsibility for the consequences of use of such information or for any infringement of patents or other rights of third parties that may result from its use. You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning the Marvell products disclosed herein. Marvell and the Marvell logo are registered trademarks of Marvell or its affiliates. Please visit www.marvell.com for a complete list of Marvell trademarks and guidelines for use of such trademarks. Other names and brands may be claimed as the property of others.

Copyright

Table of Contents

Preface

1 Functionality and Features

2 Configuring Teaming in Windows Server

3 Virtual LANs in Windows

8 VMware Driver Software

List of Figures

Figure **Page**

List of Tables

Preface

This section provides information about this guide's intended audience, content, document conventions, and laser safety information.

Intended Audience

This guide is intended for personnel responsible for installing and maintaining computer networking equipment.

What Is in This Guide

This guide describes the features, installation, and configuration of the Marvell FastLinQ 57840/57810/57800 and other 57*xx* and 57*xxx* Converged Network Adapters and Intelligent Ethernet Adapters.

Related Materials

For additional information, refer to the *Migration Guide: QLogic®/Broadcom NetXtreme I/II Adapters*, document number BC0054606-00. The migration guide presents an overview of Marvell's acquisition of specific Broadcom® Ethernet assets and its end-user impact, and was written in cooperation between Broadcom and Marvell.

Documentation Conventions

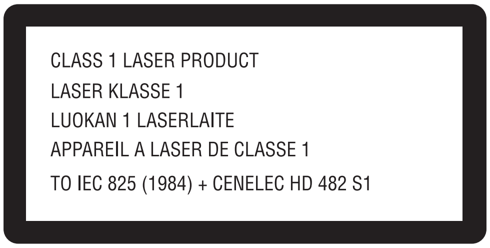
This guide uses the following documentation conventions:

- **NOTE** provides additional information.
- **CAUTION** without an alert symbol indicates the presence of a hazard that could cause damage to equipment or loss of data.
- **⚠ CAUTION** with an alert symbol indicates the presence of a hazard that could cause minor or moderate injury.
- **⚠ WARNING** indicates the presence of a hazard that could cause serious injury or death.

- Text in blue font indicates a hyperlink (jump) to a figure, table, or section in this guide, and links to Web sites are shown in underlined blue. For example:
  - ❑ Table 9-2 lists problems related to the user interface and remote agent.
  - ❑ See "Installation Checklist" on page 6.
  - ❑ For more information, visit www.marvell.com.
- Text in **bold** font indicates user interface elements such as a menu items, buttons, check boxes, or column headings. For example:
  - ❑ Click the **Start** button, point to **Programs**, point to **Accessories**, and then click **Command Prompt**.
  - ❑ Under **Notification Options**, select the **Warning Alarms** check box.
- Text in `Courier` font indicates a file name, directory path, or command line text. For example:
  - ❑ To return to the root directory from anywhere in the file structure, type **`cd /root`** and press ENTER.
  - ❑ Issue the following command: **`sh ./install.bin`**
- Key names and key strokes are indicated with UPPERCASE:
  - ❑ Press the CTRL+P keys.
  - ❑ Press the UP ARROW key.
- Text in *italics* indicates terms, emphasis, variables, or document titles. For example:
  - ❑ What are *shortcut keys*?
  - ❑ To enter the date type *mm/dd/yyyy* (where *mm* is the month, *dd* is the day, and *yyyy* is the year).
- Topic titles between quotation marks identify related topics either within this manual or in the online help, which is also referred to as *the help system* throughout this document.

Laser Safety Information

This product may use Class 1 laser optical transceivers to communicate over the fiber optic conductors. The U.S. Department of Health and Human Services (DHHS) does not consider Class 1 lasers to be hazardous. The International Electrotechnical Commission (IEC) 825 Laser Safety Standard requires labeling in English, German, Finnish, and French stating that the product uses Class 1 lasers. Because it is impractical to label the transceivers, the following label is provided in this manual.



CLASS 1 LASER PRODUCT
LASER KLASSE 1
LUOKAN 1 LASERLAITE
APPAREIL A LASER DE CLASSE 1
TO IEC 825 (1984) + CENELEC HD 482 S1

1 Functionality and Features

This chapter covers the following for the adapters:

- Functional Description
- "Features" on page 2
- "Supported Operating Environments" on page 6
- "Network Link and Activity Indication" on page 7

Functional Description

The Marvell 57*xx* and 57*xxx* adapter is a new class of gigabit Ethernet (GbE) and 10GbE converged network interface controller (C-NIC) that can simultaneously perform accelerated data networking and storage networking on a standard Ethernet network. The C-NIC offers acceleration for popular protocols used in the data center, such as:

- TCP offload engine (TOE) for accelerating TCP over 1GbE and 10GbE (on Windows Server OSs that support TOE)
- Internet small computer systems interface (iSCSI) offload for accelerating network storage access featuring centralized boot functionality (iSCSI boot)
- Fibre Channel over Ethernet (FCoE) offload and acceleration for Fibre Channel block storage

NOTE

Not all adapters support each listed protocol. Refer to the specific product data sheet for protocol support.

Enterprise networks that use multiple protocols and multiple network fabrics benefit from the C-NICs ability to combine data communications, storage, and clustering over a single Ethernet fabric by boosting server CPU processing performance and memory utilization while alleviating I/O bottlenecks.

The Marvell 57*xx* and 57*xxx* adapters include a 10Mbps, 100Mbps, 1000Mbps, or 10Gbps Ethernet MAC with both half-duplex and full-duplex capability and a 10Mbps, 100Mbps, 1000Mbps, or 10Gbps physical layer (PHY). The transceiver is fully compatible with the IEEE 802.3 standard for auto-negotiation of speed.

Using the Marvell teaming software, you can split your network into virtual LANs (VLANs), as well as group multiple network adapters together into teams to provide network load balancing and fault tolerance functionality.

- For detailed information about teaming, see Chapter 2 Configuring Teaming in Windows Server and Chapter 12 Marvell Teaming Services.
- For a description of VLANs, see Chapter 3 Virtual LANs in Windows.

Features

The following is a list of the Marvell 57*xx* and 57*xxx* adapter features. Some features may not be available on all adapters.

- TCP offload engine (TOE)
- iSCSI offload (see “iSCSI” on page 4)
- Fibre Channel over Ethernet (FCoE) (see “FCoE” on page 5)
- NIC partitioning (NPAR)
- Data center bridging (DCB):
  - Enhanced transmission selection (ETS; IEEE 802.1Qaz)
  - Priority-based flow control (PFC; IEEE 802.1Qbb)
  - Data center bridging capability exchange protocol (DCBX; CEE version 1.01)
- Single-chip solution:
  - 100/1000/10G triple-speed MAC
  - 1G/10G triple-speed MAC
  - Serializer/deserializer (SerDes) interface for optical transceiver connection
  - PCI Express® Gen3 x8 (10GbE 57840 only)
  - Zero copy capable hardware
- Other performance features:
  - TCP, IP, UDP checksum
  - TCP segmentation

    - Adaptive interrupts (see “Adaptive Interrupt Frequency” on page 6)
    - Receive side scaling (RSS)
- Manageability:
    - QLogic Control Suite (QCS) CLI diagnostic and configuration software (see “QLogic Control Suite CLI” on page 6)
    - QConvergeConsole (QCC) GUI diagnostics and configuration software for Linux® and Windows®
    - QCC PowerKit diagnostics and configuration software extensions to Microsoft® PowerShell® for Linux, VMware®, and Windows
    - QCC vSphere®/vCenter® GUI plug-in diagnostics and configuration software for VMware
    - QCC ESXCLI plug-in diagnostics and configuration software for VMware
    - Pre-boot comprehensive configuration management (CCM) configuration software
    - Pre-boot unified extensible firmware interface (UEFI) human interface infrastructure (HII) configuration software
    - Supports the PXE 2.0 specification
    - Wake on LAN (WoL) support
    - Universal management port (UMP) support
    - SMBus controller
    - Advanced configuration and power interface (ACPI) 1.1a compliant (multiple power modes) (see “Power Management” on page 5)
    - Intelligent platform management interface (IPMI) support
- Advanced network features:
    - Jumbo frames (up to 9,600 bytes). The OS and the link partner must support jumbo frames.
    - Virtual LANs
    - IEEE Std 802.3ad teaming
    - Smart Load Balancing™ teaming
    - Flow control (IEEE Std 802.3x)
    - LiveLink™ (supported in both the 32-bit and 64-bit Windows operating systems)
- Logical link control (LLC) (IEEE Std 802.2)

- High-speed on-chip reduced instruction set computer (RISC) processor (see "ASIC with Embedded RISC Processor" on page 6)
- Integrated 96KB frame buffer memory
- Quality of service (QoS)
- Serial gigabit media independent interface (SGMII), gigabit media independent interface (GMII), and media independent interface (MII) management interface
- 256 unique MAC unicast addresses
- Support for multicast addresses through a 128-bit hashing hardware function
- Support for VMDirectPath I/O over PCI physical functions

  Marvell 57*xx* and 57*xxx* Series Adapters support VMDirectPath I/O in Linux and ESX environments. VMDirectPath I/O is not supported in Windows environments. Marvell 57*xx* and 57*xxx* Series Adapters can be assigned to virtual machines for PCI pass-through operation. However, due to function level dependencies, all PCIe functions associated with an adapter must be assigned to the same virtual machine. Sharing PCIe physical functions across the hypervisor and/or one or more virtual machines is not supported.
- Serial Flash NVRAM memory
- JTAG support
- PCI Power Management Interface (v1.1)
- 64-bit base address register (BAR) support
- EM64T processor support
- iSCSI and FCoE boot support
- Virtualization:
  - ❑ Microsoft
  - ❑ VMware
  - ❑ Linux
  - ❑ XenServer®
- Single root I/O virtualization (SR-IOV)

iSCSI

The Internet Engineering Task Force (IETF) has standardized iSCSI. *SCSI* is a popular protocol that enables systems to communicate with storage devices, using block-level transfer (that is, address data stored on a storage device that is not a whole file). *iSCSI* maps the SCSI request/response application protocols and its standardized command set over TCP/IP networks.

Because iSCSI uses TCP as its sole transport protocol, it benefits from hardware acceleration of the TCP processing. However, iSCSI as a Layer 5 protocol has additional mechanisms beyond the TCP layer. iSCSI processing can also be offloaded, thereby reducing CPU utilization even further.

The Marvell 57*xx* and 57*xxx* adapters target best-system performance, maintain system flexibility to changes, and support current and future OS convergence and integration. Therefore, the adapter's iSCSI offload architecture is unique as evident by the split between hardware and host processing.

FCoE

FCoE allows Fibre Channel protocol to be transferred over Ethernet. FCoE preserves existing Fibre Channel infrastructure and capital investments. The following FCoE features are supported:

- Full stateful hardware FCoE offload
- Receiver classification of FCoE and FCoE initialization protocol (FIP) frames. FIP is used to establish and maintain connections.
- Receiver CRC offload
- Transmitter CRC offload
- Dedicated queue set for Fibre Channel traffic
- Data center bridging (DCB) provides lossless behavior with priority flow control (PFC)
- DCB allocates a share of link bandwidth to FCoE traffic with enhanced transmission selection (ETS)
- Supports Technical Committee T11 *Fibre Channel - Link Services (FC-LS)* specification; N_Port ID virtualization (NPIV) on Linux and Windows

Power Management

The adapter speed setting will link at the configured speed for WoL when the system is powered down.

NOTE

Dell® supports WoL on only one adapter in the system at a time.

For specific systems, see your system documentation for WoL support.

Adaptive Interrupt Frequency

The adapter driver intelligently adjusts host interrupt frequency based on traffic conditions to increase overall application throughput. When traffic is light, the adapter driver interrupts the host for each received packet, minimizing latency. When traffic is heavy, the adapter issues one host interrupt for multiple, back-to-back incoming packets, preserving host CPU cycles.

ASIC with Embedded RISC Processor

The core control for Marvell 57*xx* and 57*xxx* adapters resides in a tightly integrated, high-performance ASIC. The ASIC includes a RISC processor, which provides the flexibility to add new features to the adapter and conforms it to future network requirements through software downloads. This functionality also enables the adapter drivers to exploit the built-in host offload functions on the adapter as host operating systems are enhanced to take advantage of these functions.

QLogic Control Suite CLI

QLogic Control Suite (QCS) CLI provides useful information about each network adapter that is installed in your system. The QCS CLI utility also enables you to perform detailed tests, diagnostics, and analyses on each adapter, as well as to modify property values and view traffic statistics for each adapter.

Supported Operating Environments

The Marvell 57*xx* and 57*xxx* adapter has software support for the following operating systems:

- Microsoft Windows (32-bit and 64-bit extended)
- Linux (64-bit extended)
- ESXi™ Server (VMware)
- Citrix® XenServer
- Ubuntu

Network Link and Activity Indication

For copper-wire Ethernet connections, the state of the network link and activity is indicated by the LEDs on the RJ45 connector, as described in Table 1-1.

Table 1-1. Network Link and Activity Indicated by the RJ45 Port LEDs

Port LED	LED Appearance	Network State
Link LED	Off	No link (cable disconnected)
	Continuously illuminated	Link
Activity LED	Off	No network activity
	Blinking	Network activity

For fiber optic Ethernet connections and SFP+, the state of the network link and activity is indicated by a single LED located adjacent to the port connector, as described in Table 1-2.

Table 1-2. Network Link and Activity Indicated by the Port LED

LED Appearance	Network State
Off	No link (cable disconnected)
Continuously illuminated	Link
Blinking	Network activity

QLogic Control Suite also provides information about the status of the network link and activity.

2 Configuring Teaming in Windows Server

Teaming configuration in a Microsoft Windows Server® system includes an overview of load balancing and fault tolerance.

NOTE

This chapter describes teaming for adapters in Windows Server systems. For more information on a similar technology on Linux operating systems (called "channel bonding"), refer to your operating system documentation.

Load Balancing and Fault Tolerance

Teaming provides traffic load balancing and fault tolerance: redundant adapter operation in the event that a network connection fails. When multiple gigabit Ethernet network adapters are installed in the same system, they can be grouped into teams to create a virtual adapter.

A team can comprise two to eight network interfaces, and each interface can be designated as a primary interface or a standby interface. (Standby interfaces can be used only in a switch independent NIC teaming Smart Load Balancing and Failover type of team, and only one standby interface can be designated per SLB team.) If traffic is not identified on any of the adapter team member connections due to failure of the adapter, cable, switch port, or switch (where the teamed adapters are attached to separate switches), the load distribution is reevaluated and reassigned among the remaining team members. If all of the primary adapters are down, the hot standby adapter becomes active. Existing sessions are maintained and there is no impact on the user.

NOTE

Although you can create a team with one adapter, Marvell does not recommend this practice because it defeats the purpose of teaming. A team consisting of one adapter is automatically created when setting up VLANs on a single adapter, and this should be the only time that you create a team with one adapter.

Types of Teams

The available types of teams for the Windows family of operating systems are:

- Smart Load Balancing and Failover
- Link Aggregation (802.3ad)
- Generic Trunking (FEC/GEC)/802.3ad-Draft Static
- SLB (Auto-Fallback Disable)

Smart Load Balancing and Failover

Smart Load Balancing and Failover is the Broadcom® implementation of switch-independent NIC teaming load balancing based on IP flow. This feature supports balancing IP traffic across multiple adapters (team members) in a bidirectional manner. In this type of team, all adapters in the team have separate MAC addresses. This team type provides automatic fault detection and dynamic failover to another team member or to a hot standby member. Failover is performed independently of Layer 3 protocol (IP, IPX, and NetBIOS Extended User Interface [NetBEUI]); rather, it works with existing Layer 2 and 3 switches. No switch configuration (such as trunk, link aggregation) is necessary for this type of team to work.

NOTE

- If you do not enable LiveLink when configuring SLB teams, Marvell recommends that you either disable Spanning Tree Protocol (STP) or enable Port Fast at the switch or port. This practice minimizes the downtime due to spanning tree loop determination when failing over. LiveLink mitigates such issues.
- TCP/IP is fully balanced and Internetwork Packet Exchange (IPX) balances only on the transmit side of the team; other protocols are limited to the primary adapter.
- If a team member is linked at a higher speed than another, most of the traffic is handled by the adapter with the higher speed rate.

Link Aggregation (802.3ad)

The Link Aggregation mode supports link aggregation and conforms to the IEEE 802.3ad (LACP) specification. Configuration software allows you to dynamically configure the adapters that you want to participate in a specific team. If the link partner is not correctly configured for 802.3ad link configuration, errors are detected and noted. With this mode, all adapters in the team are configured to receive packets for the same MAC address. The team link partner determines the load-balancing scheme for inbound packets. In this mode, at least one of the link partners must be in active mode.

NOTE

The static and dynamic Link Aggregation (switch dependent) team type is not supported on ports with NIC partitioning (NPAR) mode enabled or iSCSI-offload enabled. Some switches support FCoE-offload in dynamic LACP teaming mode. Consult your switch documentation for more information.

Generic Trunking (FEC/GEC)/802.3ad-Draft Static

The Generic Trunking (FEC/GEC)/802.3ad-Draft Static type of team is very similar to the Link Aggregation (802.3ad) type of team in that all adapters in the team are configured to receive packets for the same MAC address. The Generic Trunking (FEC/GEC)/802.3ad-Draft Static) type of team, however, does not provide LACP or marker protocol support. This team type supports a variety of environments in which the adapter link partners are statically configured to support a proprietary trunking mechanism. For instance, this type of team could be used to support Lucent® OpenTrunk™ or Cisco® Fast EtherChannel (FEC). Basically, the Generic Trunking team type is a light version of the Link Aggregation (802.3ad) team type. This approach is much simpler because it does contain a formalized link aggregation control protocol (LACP). As with the other types of teams, the creation of teams and the allocation of physical adapters to various teams is done statically through user configuration software.

The Generic Trunking (FEC/GEC/802.3ad-Draft Static) type of team supports load balancing and failover for both outbound and inbound traffic.

NOTE

Generic Trunking (FEC/GEC/802.3ad Draft Static) team type is not supported for ports with NPAR mode or FCoE-Offload or iSCSI-Offload enabled.

SLB (Auto-Fallback Disable)

The SLB (Auto-Fallback Disable) type of team is identical to the Smart Load Balancing and Failover type of team, with the following exception: When the standby member is active, if a primary member comes back on line, the team continues using the standby member, rather than switching back to the primary member.

All primary interfaces in a team participate in load-balancing operations by sending and receiving a portion of the total traffic. Standby interfaces take over in the event that all primary interfaces have lost their links.

Failover teaming provides redundant adapter operation (fault tolerance) in the event that a network connection fails. If the primary adapter in a team is disconnected because of failure of the adapter, cable, or switch port, the secondary team member becomes active, redirecting both inbound and outbound traffic originally assigned to the primary adapter. Sessions are maintained, causing no impact to the user.

Limitations of Smart Load Balancing and Failover and SLB (Auto-Fallback Disable) Types of Teams

Smart Load Balancing (SLB) is a protocol-specific scheme. The level of support for IP is listed in Table 2-1.

Table 2-1. Smart Load Balancing

Operating System Protocol	Failover and Fallback—All Dell IP	Failover and Fallback—Multivendor IP
Windows Server 2016 and 2019	Yes	Yes
Azure Stack HCI	Yes	Yes
Operating System Protocol	**Load Balance—All Dell IP**	**Load Balance —Multivendor IP**
Windows Server 2016 and 2019	Yes	Yes
Azure Stack HCI	Yes	Yes

The Smart Load Balancing type of team works with all Ethernet switches without having to configure the switch ports to any special trunking mode. Only IP traffic is load-balanced in both inbound and outbound directions. IPX traffic is load-balanced in the outbound direction only. Other protocol packets are sent and received through one primary interface only. Failover for non-IP traffic is supported only for Dell network adapters. The Generic Trunking type of team requires the Ethernet switch to support some form of port trunking mode (for example, Cisco's Gigabit EtherChannel or other switch vendor's Link Aggregation mode). The Generic Trunking type of team is protocol-independent, and all traffic should be load-balanced and fault-tolerant.

NOTE

If you do not enable LiveLink when configuring SLB teams, Marvell recommends that you either disable STP or enable Port Fast at the switch. This practice minimizes the downtime due to the spanning tree loop determination when failing over. LiveLink mitigates such issues.

Teaming and Large Send Offload and Checksum Offload Support

Large send offload (LSO) and checksum offload are enabled for a team only when all of the members support and are configured for the feature.

3 Virtual LANs in Windows

This chapter provides information about VLANs in Windows for teaming.

- VLAN Overview
- "Adding VLANs to Teams" on page 16

VLAN Overview

Virtual LANs (VLANs) allow you to split your physical LAN into logical parts, to create logical segmentation of work groups, and to enforce security policies for each logical segment. Each defined VLAN behaves as its own separate network with its traffic and broadcasts isolated from the others, increasing bandwidth efficiency within each logical group. Up to 64 VLANs (63 tagged and 1 untagged) can be defined for each Marvell adapter on your server using the NIC teaming driver (through the QCC GUI or QCS CLI), depending on the amount of memory available in your system. See the respective Linux, VMware, or Windows documentation for more information on the in-OS NIC bonding/teaming services.

The VLAN definitions are created as follows:

Windows Server 2012 and later	Windows in-OS NIC teaming service
Linux	in-OS NIC bonding services
VMware	in-OS NIC teaming services

VLANs can be added to a team/bond to allow multiple VLANs with different VLAN IDs. A virtual adapter is created for each VLAN added.

Although VLANs are commonly used to create individual broadcast domains and separate IP subnets, it is sometimes useful for a server to have a simultaneous presence on more than one VLAN. Marvell adapters support multiple VLANs on a per-port or per-team basis, allowing very flexible network configurations.

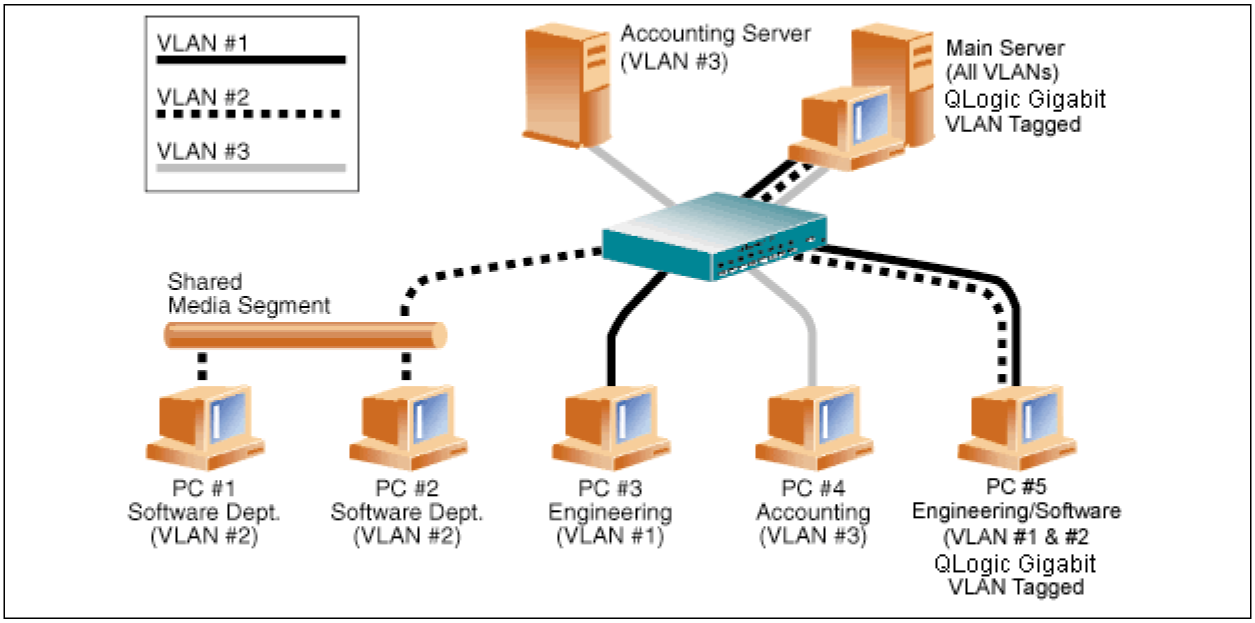

VLAN #1
VLAN #2
VLAN #3
Accounting Server (VLAN #3)
Main Server (All VLANs) QLogic Gigabit VLAN Tagged
Shared Media Segment
PC #1 Software Dept. (VLAN #2)
PC #2 Software Dept. (VLAN #2)
PC #3 Engineering (VLAN #1)
PC #4 Accounting (VLAN #3)
PC #5 Engineering/Software (VLAN #1 & #2 QLogic Gigabit VLAN Tagged

Figure 3-1. Example of Servers Supporting Multiple VLANs with Tagging

Figure 3-1 shows an example network that uses VLANs. In this example network, the physical LAN consists of a switch, two servers, and five clients. The LAN is logically organized into three different VLANs, each representing a different IP subnet. Table 3-1 describes the features of this network.

Table 3-1. Example VLAN Network Topology

Component	Description
VLAN #1	An IP subnet consisting of the Main Server, PC #3, and PC #5. This subnet represents an engineering group.
VLAN #2	Includes the Main Server, PC #1 and PC #2 through shared media segment, and PC #5. This VLAN is a software development group.
VLAN #3	Includes the Main Server, the Accounting Server, and PC #4. This VLAN is an accounting group.

Table 3-1. Example VLAN Network Topology (Continued)

Component	Description
Main Server	A high-use server that needs to be accessed from all VLANs and IP subnets. The Main Server has a Marvell adapter installed. All three IP subnets are accessed through the single physical adapter interface. The server is attached to one of the switch ports, which is configured for VLANs #1, #2, and #3. Both the adapter and the connected switch port have tagging turned on. Because of the tagging VLAN capabilities of both devices, the server is able to communicate on all three IP subnets in this network, but continues to maintain broadcast separation between all of them.
Accounting Server	Available to VLAN #3 only. The Accounting Server is isolated from all traffic on VLANs #1 and #2. The switch port connected to the server has tagging turned off.
PCs #1 and #2	Attached to a shared media hub that is then connected to the switch. PCs #1 and #2 belong only to VLAN #2, and are logically in the same IP subnet as the Main Server and PC #5. The switch port connected to this segment has tagging turned off.
PC #3	A member of VLAN #1, PC #3 can communicate only with the Main Server and PC #5. Tagging is not enabled on PC #3 switch port.
PC #4	A member of VLAN #3, PC #4 can only communicate with the servers. Tagging is not enabled on PC #4 switch port.
PC #5	A member of both VLANs #1 and #2, PC #5 has a Marvell adapter installed. It is connected to switch port #10. Both the adapter and the switch port are configured for VLANs #1 and #2 and have tagging enabled.

NOTE

VLAN tagging is only required to be enabled on switch ports that create trunk links to other switches, or on ports connected to tag-capable end-stations, such as servers or workstations with Marvell adapters.

For Hyper-V®, create VLANs in the vSwitch-to-VM connection instead of in a team, to allow VM live migrations to occur without having to ensure the future host system has a matching team VLAN setup.

Adding VLANs to Teams

Each Marvell adapter team supports up to 64 VLANs (63 tagged and 1 untagged). Note that only Marvell adapters and Alteon® AceNIC adapters can be part of a team with VLANs. With multiple VLANs on an adapter, a server with a single adapter can have a logical presence on multiple IP subnets. With multiple VLANs in a team, a server can have a logical presence on multiple IP subnets and benefit from load balancing and failover.

NOTE

You can configure adapters that are members of a failover team to also support VLANs. Because VLANs are not supported for an Intel LOM, if an Intel LOM is a member of a failover team, you cannot configure VLANs for that team.

4 Installing the Hardware

This chapter applies to Marvell 57*xx* and 57*xxx* add-in network interface cards. Hardware installation covers the following:

- System Requirements
- "Safety Precautions" on page 19
- "Preinstallation Checklist" on page 19
- "Installation of the Add-In NIC" on page 20

NOTE

Service Personnel: This product is intended only for installation in a Restricted Access Location (RAL).

System Requirements

Before you install Marvell 57*xx* and 57*xxx* adapters, verify that your system meets the hardware and operating system requirements described in this section.

Hardware Requirements

- IA32- or EMT64-based computer that meets operating system requirements
- One open PCI Express slot. Depending on the PCI Express support on your adapter, the slot may be one of these types:
  - ❑ PCI Express 1.0a x1
  - ❑ PCI Express 1.0a x4
  - ❑ PCI Express Gen2 x8
  - ❑ PCI Express Gen3 x8

  Full dual-port 10Gbps bandwidth is supported on PCI Express Gen2 x8 or faster slots.
- 128MB RAM (minimum)

Operating System Requirements

NOTE

Because the *Dell Update Packages Version xx.xx.xxx User's Guide* is not updated in the same cycle as this Ethernet adapter user's guide, consider the operating systems listed in this section as the most current.

This section describes the requirements for each supported OS.

General

The following host interface is required:

- PCI Express v1.0a, x1 (or greater) Host Interface

Microsoft Windows

One of the following versions of Microsoft Windows:

- Windows Server 2019
- Windows Server 2016
- Azure Stack HCI

Linux

One of the following versions of Linux:

- Red Hat Enterprise Linux (RHEL) 8.3
- RHEL 8.2
- RHEL 7.9
- RHEL 7.8
- SUSE Linux Enterprise Server (SLES) 15 SP2
- SLES 15 SP1

VMware ESXi

One of the following versions of vSphere® ESXi:

- VMware ESXi 7.0 U1
- VMware ESXi 6.7 U3

Citrix XenServer

The following versions of Hypervisor:

- Hypervisor 8.2 LTSR
- Hypervisor 7.2 CU2 LTSR

Ubuntu

- Ubuntu 20.04

Safety Precautions

⚠ WARNING

The adapter is being installed in a system that operates with voltages that can be lethal. Before you open the case of your system, observe the following precautions to protect yourself and to prevent damage to the system components:

- Remove any metallic objects or jewelry from your hands and wrists.
- Make sure to use only insulated or nonconducting tools.
- Before you touch internal components, verify that the system is powered OFF and is unplugged.
- Install or remove adapters in a static-free environment. The use of a properly grounded wrist strap or other personal antistatic devices and an antistatic mat is strongly recommended.

Preinstallation Checklist

1. Verify that your system meets the hardware and software requirements listed under “System Requirements” on page 17.
2. Verify that your system is using the latest BIOS.

NOTE

If you acquired the adapter software on a disk or from the Dell support Web Site (http://support.dell.com), verify the path to the adapter driver files.

3. If your system is active, shut it down.
4. When system shutdown is complete, turn off the power and unplug the power cord.
5. Remove the adapter from its shipping package and place it on an antistatic surface.
6. Check the adapter for visible signs of damage, especially on the edge connector. Never attempt to install a damaged adapter.

Installation of the Add-In NIC

The following instructions apply to installing the Marvell 57*xx* and 57*xxx* adapters (add-in NIC) in most systems. Refer to the manuals that were supplied with your system for details about performing these tasks on your specific system.

Installing the Add-In NIC

1. Review Safety Precautions and Preinstallation Checklist. Before you install the adapter, ensure that the system power is OFF, the power cord is unplugged from the power outlet, and that you are following proper electrical grounding procedures.
2. Open the system case and select the slot based on the adapter, which may be of type PCIe® 1.0a x1, PCIe 1.0a x4, PCIe Gen2 x8, PCIe Gen3 x8, or other appropriate slot. A lesser width adapter can be seated into a greater width slot (x8 in an x16), but a greater width adapter cannot be seated into a lesser width slot (x8 in an x4). If you do not know how to identify a PCI Express slot, refer to your system documentation.
3. Remove the blank cover-plate from the slot that you selected.
4. Align the adapter connector edge with the PCI Express connector slot in the system.
5. Applying even pressure at both corners of the card, push the adapter card into the slot until it is firmly seated. When the adapter is properly seated, the adapter port connectors are aligned with the slot opening, and the adapter faceplate is flush against the system chassis.

   CAUTION

   Do not use excessive force when seating the card, as this may damage the system or the adapter. If you have difficulty seating the adapter, remove it, realign it, and try again.

6. Secure the adapter with the adapter clip or screw.
7. Close the system case and disconnect any personal antistatic devices.

Connecting the Network Cables

The Marvell 57*xx* and 57*xxx* adapters have either an RJ45 connector used for attaching the system to an Ethernet copper-wire segment or a fiber optic connector for attaching the system to an Ethernet fiber optic segment.

NOTE

This section does not apply to blade servers.

Copper Wire

To connect a copper wire:

1. Select an appropriate cable. Table 4-1 lists the copper cable requirements for connecting to 100 and 1000BASE-T and 10GBASE-T ports.

Table 4-1. 100/1000BASE-T and 10GBASE-T Cable Specifications

Port Type	Connector	Media	Maximum Distance
100/1000BASE-T [a]	RJ45	CAT-5 [b] UTP	100m (328ft)
10GBASE-T	RJ45	CAT-6 [c] UTP CAT-6a and CAT-7 [c] UTP	40m (131ft) 100m (328CAT-ft)

[a] 1000BASE-T signaling requires four twisted pairs of CAT-5 balanced cabling, as specified in ISO/IEC 11801:2002 and ANSI/EIA/TIA-568-B.

[b] CAT-5 is the minimum requirement. CAT-5e, CAT-6, CAT-6a, and CAT-7 are fully supported.

[c] 10GBASE-T signaling requires four twisted pairs of CAT-6 or CAT-6A (augmented CAT-6) balanced cabling, as specified in ISO/IEC 11801:2002 and ANSI/TIA/EIA-568-B.

2. Connect one end of the cable to the RJ45 connector on the adapter.
3. Connect the other end of the cable to an RJ45 Ethernet network port.

Fiber Optic

To connect a fiber optic cable:

1. Select an appropriate cable. Table 4-2 lists the fiber optic cable requirements for connecting to 1000 and 2500BASE-X ports. See also the tables in "Supported SFP+ Modules Per NIC" on page 253.

Table 4-2. 1000/2500BASE-X Fiber Optic Specifications

Port Type	Connector	Media	Maximum Distance
1000BASE-X	Small form factor (SFF) transceiver with LC™ connection system (Infineon® part number V23818-K305-L57)	Multimode fiber (MMF) System optimized for 62.5/50µm graded index fiber	550m (1804 ft)
2500BASE-X [a]	Small form factor (SFF) transceiver with LC™ connection system (Finisar® part number FTLF8542E2KNV)	Multimode fiber (MMF) System optimized for 62.5/50µm graded index fiber	550m (1804 ft)

[a] Electricals are leveraged from IEEE 802.3ae-2002 (XAUI). 2500BASE-X is a term used by Marvell to describe 2.5Gbp (3.125GBd) operation.

2. Connect one end of the cable to the fiber optic connector on the adapter.
3. Connect the other end of the cable to an fiber optic Ethernet network port.

5 Manageability

Information about manageability includes:

- CIM
- "Host Bus Adapter API" on page 24

CIM

The common information model (CIM) is an industry standard defined by the Distributed Management Task Force (DMTF). Microsoft implements CIM on Windows Server platforms. Marvell supports CIM on Windows Server and Linux platforms.

The Marvell implementation of CIM provides various classes to provide information to users through CIM client applications. Note that the Marvell CIM data provider provides data only, and users can choose their preferred CIM client software to browse the information exposed by the Marvell CIM provider.

The Marvell CIM provider provides information through the `QLGC_NetworkAdapter` class, which provides network adapter information pertaining to a group of adapters including Marvell and other vendors' controllers.

To inspect or monitor these events, use either the Event Viewer provided by Windows Server platforms or the CIM. The Marvell CIM provider also provides event information through the CIM generic event model. These events are `__InstanceCreationEvent`, `__InstanceDeletionEvent`, and `__InstanceModificationEvent`, and are defined by CIM. CIM requires the client application to register the events from the client application, using queries as the following examples to receive events properly:

```
SELECT * FROM __InstanceModificationEvent
where TargetInstance ISA “QLGC_NetworkAdapter”
SELECT * FROM __InstanceModificationEvent
where TargetInstance ISA "QLGC_ExtraCapacityGroup"
SELECT * FROM __InstanceCreationEvent
where TargetInstance ISA "QLGC_NetworkAdapter"
SELECT * FROM __InstanceDeletionEvent
where TargetInstance ISA "QLGC_NetworkAdapter"
```

```
SELECT * FROM __InstanceCreationEvent
where TargetInstance ISA "QLGC_ActsAsSpare"
SELECT * FROM __InstanceDeletionEvent
where TargetInstance ISA "QLGC_ActsAsSpare"
```

For detailed information about these events, see the CIM documentation:

http://www.dmtf.org/sites/default/files/standards/documents/DSP0004V2.3_final.pdf

Marvell also implements the SMI-S, which defines CIM management profiles for storage systems.

Host Bus Adapter API

Marvell supports the SNIA Common HBA API on Windows and Linux operating systems. The Common HBA API is an application program interface for the management of Fibre Channel Host Bus Adapters.

6 Boot Agent Driver Software

This chapter covers how to set up MBA in both client and server environments:

- Overview
- "Setting Up MBA in a Client Environment" on page 26
- "Setting Up MBA in a Linux Server Environment" on page 32

Overview

Marvell 57*xx* and 57*xxx* adapters support pre-execution environment (PXE), remote program load (RPL), iSCSI, and bootstrap protocol (BOOTP). Marvell's Multi-Boot Agent (MBA) is a software module that allows your network computer to boot with the images provided by remote servers across the network. The Marvell MBA driver complies with the PXE 2.1 specification and is released with split binary images. These images reside in the adapter's firmware and provide flexibility to users in different environments where the motherboard may or may not have built-in base code.

The MBA module operates in a client/server environment. A network consists of one or more boot servers that provide boot images to multiple computers through the network. The Marvell implementation of the MBA firmware module has been tested successfully in the following environments:

- **Linux Red Hat PXE Server**. Marvell PXE clients are able to remotely boot and use network resources (NFS mount, and so forth) and to perform Linux installations. In the case of a remote boot, the Linux universal driver binds seamlessly with the Marvell Universal Network Driver Interface (UNDI) and provides a network interface in the Linux remotely-booted client environment.
- **Intel APITEST**. The Marvell PXE driver passes all API compliance test suites.
- **Windows Deployment Services (WDS)**. To extend functionalities beyond basic network connectivity when loading an operating system through Microsoft WDS, generate a WinPE (3.0 or later) image using the EVBD or Network Driver Interface Specification (NDIS) driver.

Setting Up MBA in a Client Environment

Setting up MBA in a client environment involves the following steps:

1. Configuring the MBA Driver.
2. Setting Up the BIOS for the boot order.

Configuring the MBA Driver

This section pertains to configuring the MBA driver (located in the adapter firmware) on add-in NIC models of the Marvell network adapter. For configuring the MBA driver on LOM models of the Marvell network adapter, check your system documentation, as this driver resides in the system BIOS.

NOTE

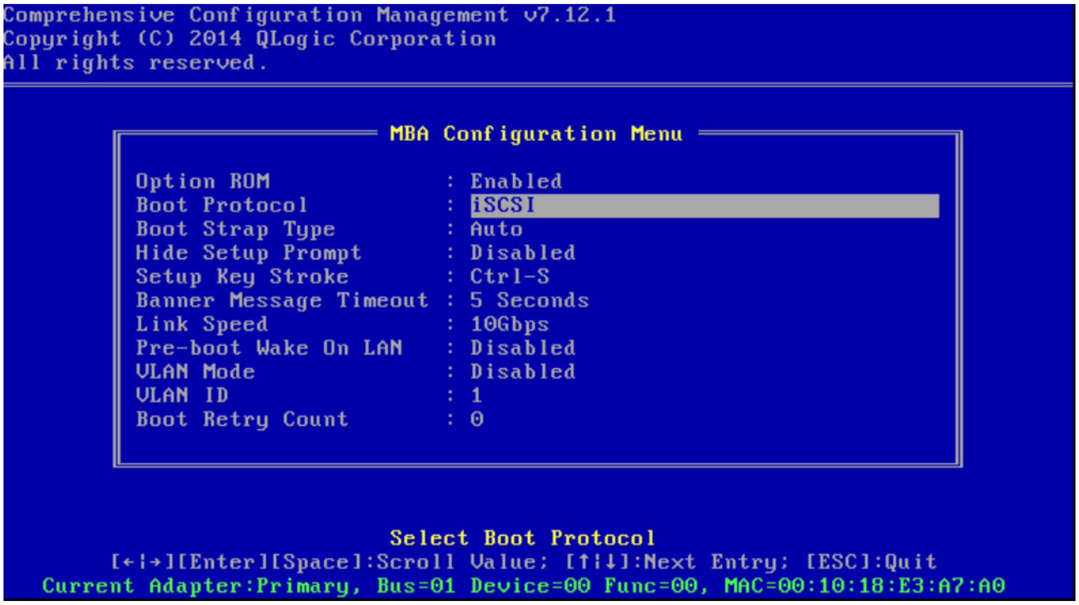
You can use Marvell's Comprehensive Configuration Management (CCM) utility or the unified extensible firmware interface (UEFI) to configure the MBA driver one adapter at a time as described in the following procedure.

CCM is available only when the system is in legacy mode; it is not available in UEFI boot mode. The UEFI device configuration pages are available in both modes.

Using Comprehensive Configuration Management

To use CCM to configure the MBA driver:

1. Restart the system.
2. Press the CTRL+ S keys within four seconds after you are prompted to do so. A list of adapters appears.
   a. Select the adapter to configure, and then press the ENTER key. The Main Menu appears.
   b. Select **MBA Configuration** to view the **MBA Configuration Menu**, as shown in Figure 6-1.



```
Comprehensive Configuration Management v7.12.1
Copyright (C) 2014 QLogic Corporation
All rights reserved.

                      MBA Configuration Menu

        Option ROM                : Enabled
        Boot Protocol             : iSCSI
        Boot Strap Type           : Auto
        Hide Setup Prompt         : Disabled
        Setup Key Stroke          : Ctrl-S
        Banner Message Timeout    : 5 Seconds
        Link Speed                : 10Gbps
        Pre-boot Wake On LAN      : Disabled
        VLAN Mode                 : Disabled
        VLAN ID                   : 1
        Boot Retry Count          : 0

                      Select Boot Protocol
      [←¦→][Enter][Space]:Scroll Value: [↑¦↓]:Next Entry: [ESC]:Quit
   Current Adapter:Primary, Bus=01 Device=00 Func=00, MAC=00:10:18:E3:A7:A0
```

Figure 6-1. CCM MBA Configuration Menu

3. To access the **Boot Protocol** item, press the UP ARROW and DOWN ARROW keys. If other boot protocols besides **Preboot Execution Environment (PXE)** are available, press RIGHT ARROW or LEFT ARROW to select the boot protocol of choice: **FCoE** or **iSCSI**.

NOTE

For iSCSI and FCoE boot-capable LOMs, set the boot protocol through the BIOS. See your system documentation for more information.

NOTE

If you have multiple adapters in your system and you are unsure which adapter you are configuring, press the CTRL+F6 keys, which causes the port LEDs on the adapter to start blinking.

4. To move to and change the values for other menu items, as needed, press the UP ARROW, DOWN ARROW, LEFT ARROW, and RIGHT ARROW keys.
5. To save the settings, press the F4 key.
6. When you are finished, press the ESC key.

Using UEFI

To use UEFI to configure the MBA driver:

1. Restart the system.
2. Enter the system BIOS **System Setup**'s **Device Settings** configuration menu (see Figure 6-2).

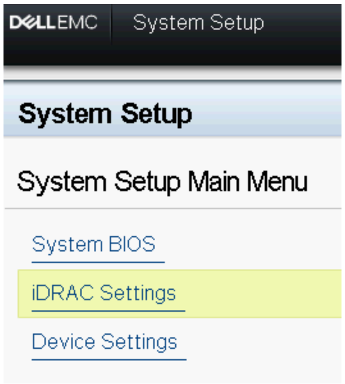
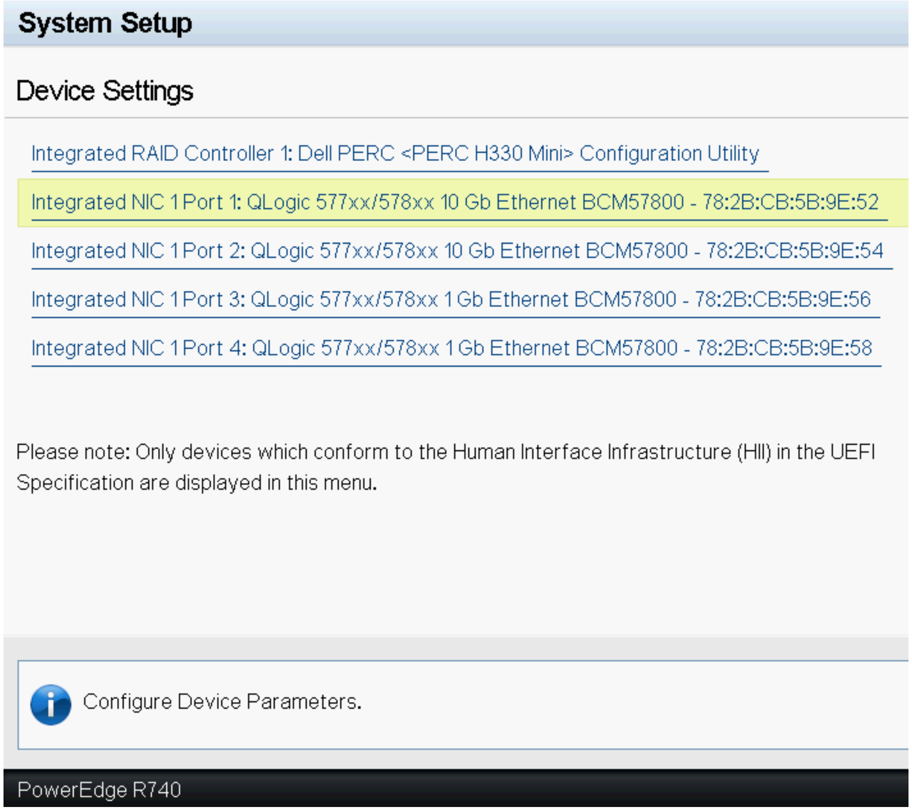


Figure 6-2. System Setup, Device Settings

3. Select the device on which you want to change MBA settings (see Figure 6-3).



System Setup

Device Settings

Integrated RAID Controller 1: Dell PERC <PERC H330 Mini> Configuration Utility

Integrated NIC 1 Port 1: QLogic 577xx/578xx 10 Gb Ethernet BCM57800 - 78:2B:CB:5B:9E:52

Integrated NIC 1 Port 2: QLogic 577xx/578xx 10 Gb Ethernet BCM57800 - 78:2B:CB:5B:9E:54

Integrated NIC 1 Port 3: QLogic 577xx/578xx 1 Gb Ethernet BCM57800 - 78:2B:CB:5B:9E:56

Integrated NIC 1 Port 4: QLogic 577xx/578xx 1 Gb Ethernet BCM57800 - 78:2B:CB:5B:9E:58

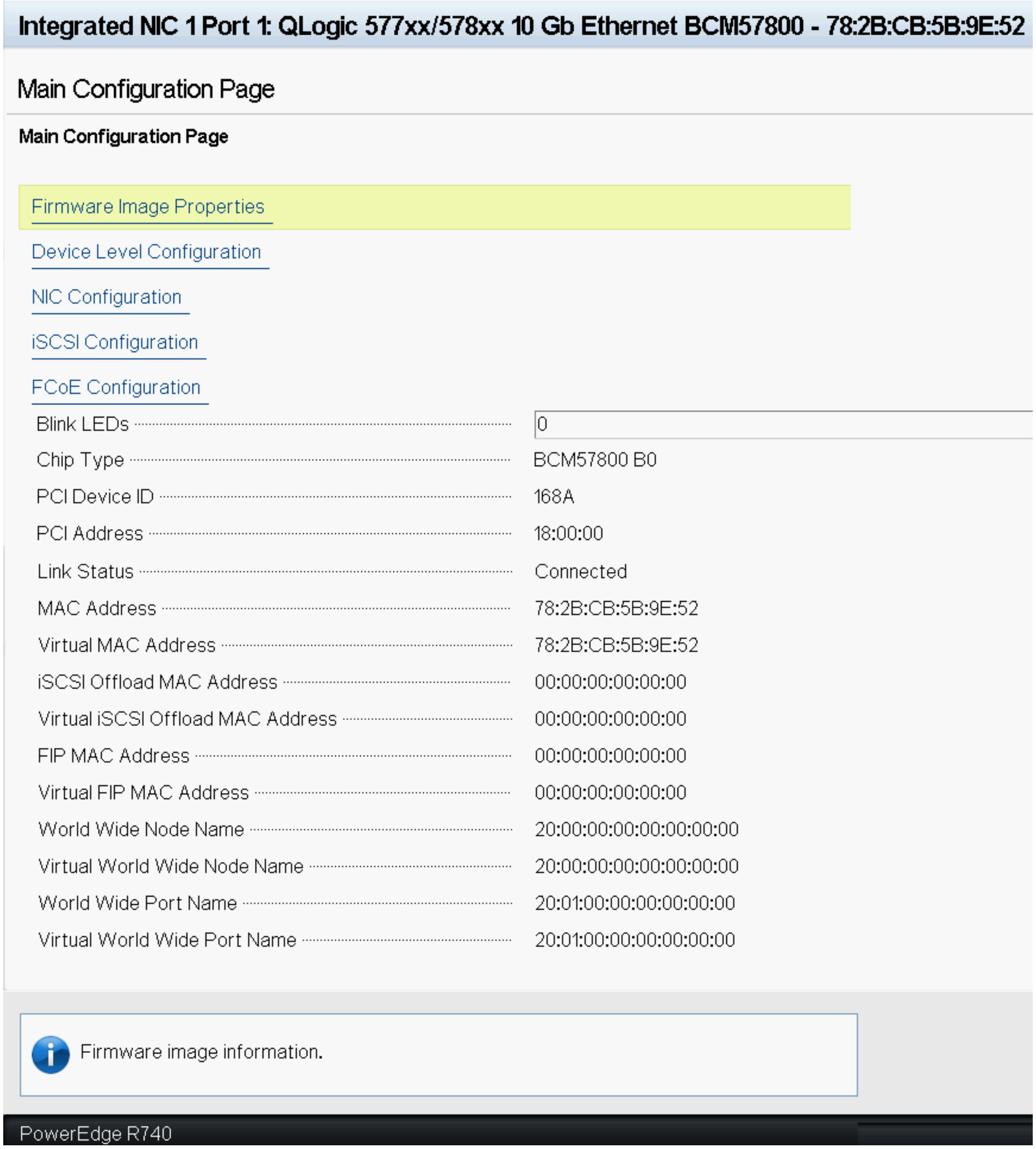
Please note: Only devices which conform to the Human Interface Infrastructure (HII) in the UEFI Specification are displayed in this menu.

Configure Device Parameters.

PowerEdge R740

Figure 6-3. Device Settings

4. On the **Main Configuration Page**, select **NIC Configuration** (see Figure 6-4).



Integrated NIC 1 Port 1: QLogic 577xx/578xx 10 Gb Ethernet BCM57800 - 78:2B:CB:5B:9E:52

Main Configuration Page

Main Configuration Page

Firmware Image Properties

Device Level Configuration

NIC Configuration

iSCSI Configuration

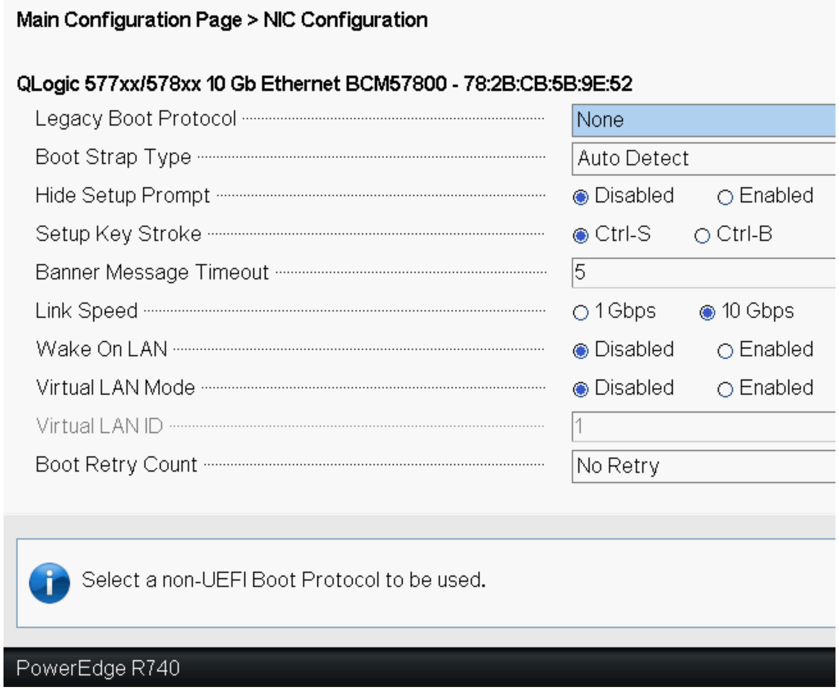
FCoE Configuration

Blink LEDs	0
Chip Type	BCM57800 B0
PCI Device ID	168A
PCI Address	18:00:00
Link Status	Connected
MAC Address	78:2B:CB:5B:9E:52
Virtual MAC Address	78:2B:CB:5B:9E:52
iSCSI Offload MAC Address	00:00:00:00:00:00
Virtual iSCSI Offload MAC Address	00:00:00:00:00:00
FIP MAC Address	00:00:00:00:00:00
Virtual FIP MAC Address	00:00:00:00:00:00
World Wide Node Name	20:00:00:00:00:00:00:00
Virtual World Wide Node Name	20:00:00:00:00:00:00:00
World Wide Port Name	20:01:00:00:00:00:00:00
Virtual World Wide Port Name	20:01:00:00:00:00:00:00

Firmware image information.

Figure 6-4. Main Configuration Page

5. In the NIC Configuration page (see Figure 6-5), use the **Legacy Boot Protocol** drop-down menu to select the boot protocol of choice, if boot protocols other than **Preboot Execution Environment (PXE)** are available. If available, other boot protocols include **iSCSI** and **FCoE**. The 57800's fixed speed, 1GbE ports support only PXE and iSCSI remote boot.



Main Configuration Page > NIC Configuration

QLogic 577xx/578xx 10 Gb Ethernet BCM57800 - 78:2B:CB:5B:9E:52

Legacy Boot Protocol	None	
Boot Strap Type	Auto Detect	
Hide Setup Prompt	● Disabled	○ Enabled
Setup Key Stroke	● Ctrl-S	○ Ctrl-B
Banner Message Timeout	5	
Link Speed	○ 1 Gbps	● 10 Gbps
Wake On LAN	● Disabled	○ Enabled
Virtual LAN Mode	● Disabled	○ Enabled
Virtual LAN ID	1	
Boot Retry Count	No Retry	

Select a non-UEFI Boot Protocol to be used.

PowerEdge R740

Figure 6-5. NIC Configuration

NOTE

For iSCSI and FCoE boot-capable LOMs, the boot protocol is set through the BIOS. See your system documentation for more information.

6. Press the UP ARROW, DOWN ARROW, LEFT ARROW, and RIGHT ARROW keys to move to and change the values for other menu items, as needed.
7. Select **Back** to go to **Main** menu
8. Select **Finish** to save and exit.

Setting Up the BIOS

To boot from the network with the MBA, make the MBA enabled adapter the first bootable device under the BIOS. This procedure depends on the system BIOS implementation. Refer to the user manual for the system for instructions.

Setting Up MBA in a Linux Server Environment

The Red Hat Enterprise Linux distribution has PXE Server support. It allows users to remotely perform a complete Linux installation over the network. The distribution comes with the boot images *boot kernel* (vmlinuz) and *initial ram disk* (initrd), which are located on the Red Hat disk#1:

```
/images/pxeboot/vmlinuz
/images/pxeboot/initrd.img
```

Refer to the Red Hat documentation for instructions on how to install PXE Server on Linux.

The `Initrd.img` file distributed with Red Hat Enterprise Linux, however, does not have a Linux network driver for the Marvell 57*xx* and 57*xxx* adapters. This version requires a driver disk for drivers that are not part of the standard distribution. You can create a driver disk for the Marvell 57*xx* and 57*xxx* adapters from the image distributed with the installation CD. Refer to the Linux `Readme.txt` file for more information.

7 Linux Driver Software

Information about the Linux driver software includes:

Introduction

This section discusses the Linux drivers for the Marvell 57*xx* and 57*xxx* network adapters listed in Table 7-1.

Table 7-1. Marvell 57xx and 57xxx Linux Drivers

Linux Driver	Description
bnx2	Linux driver for the 57*xx* 1Gb network adapters. This driver directly controls the hardware and is responsible for sending and receiving Ethernet packets on behalf of the Linux host networking stack. The driver also receives and processes device interrupts, both on behalf of itself (for Layer 2 networking) and on behalf of the C-NIC driver (for iSCSI offload).

Table 7-1. Marvell 57xx and 57xxx Linux Drivers (Continued)

Linux Driver	Description
bnx2x	Linux driver for the 57*xxx* 1Gb/10Gb network adapters. This driver directly controls the hardware and is responsible for sending and receiving Ethernet packets on behalf of the Linux host networking stack. The driver also receives and processes device interrupts, both on behalf of itself (for Layer 2 networking) and on behalf of the bnx2fc (FCoE) and C-NIC drivers.
cnic	The C-NIC driver provides the interface between Marvell's upper-layer protocol (for example, storage) drivers and Marvell's 57*xx* and 57*xxx* 1Gb and 10Gb network adapters. The C-NIC module works with the bnx2 and bnx2x network drives in the downstream and the bnx2fc (FCoE) and bnx2i (iSCSI) drivers in the upstream.
bnx2i	Linux iSCSI Host Bus Adapter driver to enable iSCSI offload on the 57*xx* and 57*xxx* 1Gb and 10Gb network adapters.
bnx2fc	Linux FCoE kernel mode driver used to provide a translation layer between the Linux SCSI stack and the Marvell FCoE firmware and hardware. In addition, the driver interfaces with the networking layer to transmit and receive encapsulated FCoE frames on behalf of the Open-FCoE `libfc/libfcoe` for FIP and device discovery.

Limitations

The Linux drivers have the limitations described in the following:

- bnx2 Driver Limitations
- bnx2x Driver Limitations
- bnx2i Driver Limitations
- bnx2fc Driver Limitations

bnx2 Driver Limitations

The current version of the driver has been tested on 2.4.*x* kernels, starting from the 2.4.24 kernel, and all 2.6.*x* and 3.*x* kernels. The driver may not compile on kernels older than 2.4.24. Testing is concentrated on i386 and x86_64 architectures. Only limited testing has been done on some other architectures.

Minor changes to some source files and the Makefile may be needed on some kernels. Additionally, the Makefile does not compile the C-NIC driver on kernels older than 2.6.31.

iSCSI offload is supported only on 2.6.31 and later kernels.

RHEL5.4 and later has special backported code to support the C-NIC driver; these distributions are supported.

bnx2x Driver Limitations

The current version of the driver has been tested on 2.6.*x* kernels, starting from the 2.6.9 kernel. The bnx2x driver may not compile on kernels older than 2.6.9. Testing is concentrated on i386 and x86_64 architectures. Only limited testing has been done on some other architectures. Minor changes to some source files and the `makefile` may be needed on some kernels.

bnx2i Driver Limitations

The current version of the driver has been tested on 2.6.*x* kernels, starting from the 2.6.18 kernel. The bnx2i driver may not compile on older kernels. Testing is concentrated on i386 and x86_64 architectures.

bnx2fc Driver Limitations

The current version of the driver has been tested on 2.6.*x* kernels, starting from the 2.6.32 kernel, which is included in RHEL 6.1 distribution. The bnx2fc driver may not compile on older kernels. Testing was limited to i386 and x86_64 architectures.

Packaging

The Linux drivers are released in the following packaging formats:

Dynamic Kernel Module Support (DKMS) Packages

- `netxtreme2-`*`version`*`.dkms.noarch.rpm`
- `netxtreme2-`*`version`*`.dkms.src.rpm`

Kernel Module Packages (KMP)

- **SLES**
  - ❑ `netxtreme2-kmp-[kernel]-`*`version`*`.i586.rpm`
  - ❑ `netxtreme2-kmp-[kernel]-`*`version`*`.x86_64.rpm`
- **Red Hat**
  - ❑ `kmod-kmp-netxtreme2-{kernel]-`*`version`*`.i686.rpm`
  - ❑ `kmod-kmp-netxtreme2-{kernel]-`*`version`*`.x86_64.rpm`

The QLogic Control Suite (QCS) CLI management utility is also distributed as an RPM package (`QCS-{`*`version`*`}.{`*`arch`*`}.rpm`).

Source Packages

Identical source files to build the driver are included in both RPM and TAR source packages. The supplemental TAR file contains additional utilities such as patches and driver diskette images for network installation.

The following is a list of included files:

- **`netxtreme2-version.src.rpm`**: RPM package with 57*xx* and 57*xxx* bnx2, bnx2x, cnic, bnx2fc, bnx2ilibfc, and libfcoe driver source.
- **`netxtreme2-version.tar.gz`**: TAR zipped package with 57*xx* and 57*xxx* bnx2, bnx2x, cnic, bnx2fc, bnx2i, libfc, and libfcoe driver source.
- **`iscsiuio-version.tar.gz`**: iSCSI user space management tool binary.

The Linux driver has a dependency on Open-FCoE userspace management tools as the front-end to control FCoE interfaces. The package names of the Open-FCoE tools are *fcoe-utils* and *open-fcoe*.

Installing Linux Driver Software

Procedures for installing the Linux driver software include:

- Installing the Source RPM Package
- Building the Driver from the Source TAR File
- Installing the Binary DKMS RPM Driver Package
- Installing the Binary KMOD and KMP Driver Package

NOTE

If a bnx2x, bnx2i, or bnx2fc driver is loaded and the Linux kernel is updated, the driver module must be recompiled if the driver module was installed using the source RPM or the TAR package. This requirement does not apply to the source DKMS RPM.

Installing the Source RPM Package

The following are guidelines for installing the driver source RPM package.

Prerequisites:

- Linux kernel source
- C compiler

To install and configure the source RPM package:

1. Install the source RPM package:

```
rpm -ivh netxtreme2-<version>.src.rpm
```

2. Change the directory to the RPM path and build the binary RPM for your kernel.

NOTE

For RHEL 8, install the `kernel-rpm-macros` and `kernel-abi-whitelists` package before building the binary RPM.

For RHEL:

```
cd ~/rpmbuild
rpmbuild -bb SPECS/netxtreme2.spec
```

For SLES:

```
cd /usr/src/packages
rpmbuild -bb SPECS/netxtreme2.spec
```

3. Install the newly compiled RPM:

   ```
   rpm -ivh RPMS/<arch>/netxtreme2-<version>.<arch>.rpm
   ```

   The `--force` option may be needed on some Linux distributions if conflicts are reported.

4. For FCoE offload, install the Open-FCoE utility.

   For RHEL 7.5 and later, the version of `fcoe-utils` or `open-fcoe` included in your distribution is sufficient and no out of box upgrades are provided.

   Where available, installation with yum will automatically resolve dependencies. Otherwise, required dependencies can be located on your OS installation media.

5. For SLES 15 and later, turn on the FCoE and link layer discover protocol agent daemon (lldpad) services for FCoE offload, and just lldpad for iSCSI-offload-TLV as follows:

   ```
   chkconfig lldpad on
   chkconfig fcoe on
   ```

6. Inbox drivers are included with all of the supported operating systems. Rebooting is the simplest means to ensure the newly installed drivers are loaded.

7. For FCoE offload, after rebooting, create configuration files for all FCoE ethX interfaces:

   ```
   cd /etc/fcoe
   cp cfg-ethx cfg-<ethX FCoE interface name>
   ```

   NOTE

   Note that your distribution might have a different naming scheme for Ethernet devices (that is, pXpX or emX instead of ethX).

8. For FCoE offload or iSCSI-offload-TLV, modify `/etc/fcoe/cfg-<interface>` by setting `DCB_REQUIRED=yes` to `DCB_REQUIRED=`**no**.

9. Turn on all ethX interfaces.

   ```
   ifconfig <ethX> up
   ```

10. For SLES, use YaST (an installation and configuration tool for openSUSE and the SUSE Linux Enterprise distributions) to configure your Ethernet interfaces to automatically start at boot by setting a static IP address or enabling DHCP on the interface.

11. For FCoE offload and iSCSI-offload-TLV, disable lldpad on Marvell Converged Network Adapter interfaces. This step is required because Marvell utilizes an offloaded DCBX client.

    ```
    lldptool set-lldp –i <ethX> adminStatus=disasbled
    ```

12. For FCoE offload and iSCSI-offload-TLV, make sure `/var/lib/lldpad/lldpad.conf` is created and each `<ethX>` block does not specify `adminStatus`, or if specified, it is set to **0** (`adminStatus=0`) as follows.

    ```
     lldp :
    {
      eth5 :
      {
        tlvid00000001 :
        {
          info = "04BC305B017B73";
        };
        tlvid00000002 :
        {
          info = "03BC305B017B73";
    ```

```
    };
  };
```

13. For FCoE offload and iSCSI-offload-TLV, restart lldpad service to apply new settings.

    ```
    service lldpad restart
    ```

14. For FCOE offload, restart FCoE service to apply new settings.

    ```
    service fcoe restart
    ```

Installing the KMP Package

NOTE

The examples in this procedure refer to the bnx2x driver, but also apply to the bxn2fc and bnx2i drivers.

To install the KMP package:

1. Install the KMP package:

   ```
   rpm -ivh <file>
   rmmod bnx2x
   ```

2. Load the driver as follows:

   ```
   modprobe bnx2x
   ```

Building the Driver from the Source TAR File

NOTE

The examples used in this procedure refer to the bnx2x driver, but also apply to the bnx2i and bnx2fc drivers.

To build the driver from the TAR file:

1. Create a directory and extract the TAR files to the directory:

   ```
   tar xvzf netxtreme2-version.tar.gz
   ```

2. Build the driver `bnx2x.ko` (or `bnx2x.o`) as a loadable module for the running kernel:

   ```
   cd netxtreme2-version
   make
   ```

3. Test the driver by loading it (first unload the existing driver, if necessary):

   `rmmod bnx2x` (or **`bnx2fc`** or **`bnx2i`**)

   `insmod bnx2x/src/bnx2x.ko` (or **`bnx2fc/src/bnx2fc.ko`**, or **`bnx2i/src/bnx2i.ko`**)

4. For iSCSI offload and FCoE offload, load the C-NIC driver (if applicable):

   ```
   insmod cnic.ko
   ```

5. Install the driver and man page:

   ```
   make install
   ```

   NOTE

   See the RPM instructions in the preceding for the location of the installed driver.

6. Install the user daemon (iscsiuio).

Refer to "Load and Run Necessary iSCSI Software Components" on page 42 for instructions on loading the software components required to use the Marvell iSCSI offload feature.

To configure the network protocol and address after building the driver, refer to the manuals supplied with your operating system.

Installing the Binary DKMS RPM Driver Package

Dynamic Kernel Module Support (DKMS) is designed to simplify the rebuilding of modules whenever you upgrade the kernel. To upgrade, create a framework where a kernel-dependent module source can reside.

To install the binary DKMS RPM driver package:

1. Download the binary DKMS RPM (`dkms-`*`version`*`.noarch.rpm`):

   http://linux.dell.com/dkms/

2. Install the binary DKMS RPM package by issuing the following command:

   ```
   rpm -ivh dkms-version.noarch.rpm
   ```

3. Install the DKMS RPM driver package by issuing the following command:

   ```
   rpm -ivh netxtreme2-version dkms.noarch.rpm
   ```

Verify that your network adapter supports iSCSI by checking the message log. If the message `bnx2i: dev eth0 does not support iSCSI` appears in the message log after loading the bnx2i driver, iSCSI is not supported. This message may not appear until the interface is opened, as with:

```
ifconfig eth0 up
```

4. To use iSCSI, refer to "Load and Run Necessary iSCSI Software Components" on page 42 to load the necessary software components. For more information, go to:

   http://linux.dell.com

Installing the Binary KMOD and KMP Driver Package

To install the binary Kernel Modules (KMOD) and KMP driver package:

1. Install the KMOD and KMP RPM driver package:

   ❑ SUSE:

   ```
   netxtreme2-kmp-default-<driver ver>_<kernel>-<rel>.<dist
   maj.min>.<arch>.rpm
   ```

   ❑ Red Hat:

   ```
   kmod-netxtreme2-<driver ver>.<dist maj.min>.<arch>.rpm
   ```

2. Verify that your network adapter supports iSCSI by checking the message log. If the message, `bnx2i: dev eth0 does not support iSCSI`, appears in the message log after loading the bnx2i driver, iSCSI is not supported. This message may not appear until the interface is opened, as with:

   ```
   ifconfig eth0 up
   ```

3. To use iSCSI, refer to "Load and Run Necessary iSCSI Software Components" on page 42 to load the necessary software components. For more information, go to:

   http://linux.dell.com

Load and Run Necessary iSCSI Software Components

The Marvell iSCSI Offload software suite consists of three kernel modules and a user daemon. Required software components can be loaded either manually or through system services.

1. Unload the existing driver, if necessary. To do so manually, issue the following command:

   ```
   rmmod bnx2i
   ```

2. Load the iSCSI driver. To do so manually, issue one of the following commands:

   ```
   insmod bnx2i.ko
   modprobe bnx2i
   ```

Unloading or Removing the Linux Driver

- Unloading or Removing the Driver from an RPM Installation
- Removing the Driver from a TAR Installation

Unloading or Removing the Driver from an RPM Installation

NOTE

- The examples used in this procedure refer to the bnx2x driver, but also apply to the bnx2fc and bnx2i drivers.
- On 2.6 kernels, it is not necessary to bring down the eth# interfaces before unloading the driver module.
- If the C-NIC driver is loaded, unload the C-NIC driver before unloading the bnx2x driver.
- Prior to unloading the bnx2i driver, disconnect all active iSCSI sessions to targets.

To unload the driver, enter `ifconfig` to bring down all eth# interfaces opened by the driver, and then issue the following command:

```
rmmod bnx2x
```

NOTE

The preceding command also removes the C-NIC module.

If the driver was installed using RPM, issue the following command to remove it:

```
rpm -e netxtreme2
```

Removing the Driver from a TAR Installation

NOTE

The examples used in this procedure refer to the bnx2x driver, but also apply to the bnx2fc and bnx2i drivers.

If the driver was installed using make install from the TAR file, manually delete the `bnx2x.ko` driver file from the operating system. See "Installing the Source RPM Package" on page 36 for the location of the installed driver.

Uninstalling QCS with the RPM Package

To uninstall the QCS CLI and/or associated RPC agent with the Linux RPM package, issue the following command:

```
% rpm -e <package_name>.rpm
```

Where **`<package_name>`** is one of the following:

QCS CLI	`QCS-CLI-<version>-<arch>.rpm`
RPC agent	`qlnxremote-<version>.<arch>.rpm`

Patching PCI Files (Optional)

NOTE

The examples used in this procedure refer to the bnx2x driver, but also apply to the bnx2fc and bnx2i drivers.

For hardware detection utilities such as Red Hat kudzu to properly identify bnx2x supported devices, you may need to update several files containing PCI vendor and device information. Apply the updates by running the scripts provided in the supplemental TAR file. For example, on Red Hat Enterprise Linux, apply the updates by issuing the following commands:

```
./patch_pcitbl.sh  /usr/share/hwdata/pcitable pci.updates
/usr/share/hwdata/pcitable.new bnx2
./patch_pciids.sh /usr/share/hwdata/pci.ids pci.updates
/usr/share/hwdata/pci.ids.new
```

Next, back up the old files and rename the new files for use.

```
cp /usr/share/hwdata/pci.ids /usr/share/hwdata/old.pci.ids
cp /usr/share/hwdata/pci.ids.new /usr/share/hwdata/pci.ids
cp /usr/share/hwdata/pcitable /usr/share/hwdata/old.pcitable
cp /usr/share/hwdata/pcitable.new /usr/share/hwdata/pcitable
```

Network Installations

For network installations through NFS, FTP, or HTTP (using a network boot disk or PXE), you may need a driver disk that contains the bnx2x driver. The driver disk includes images for the most recent Red Hat and SUSE versions. You can compile boot drivers for other Linux versions by modifying the `Makefile` and the make environment. Additional information is available from the Red Hat Web site:

http://www.redhat.com

Setting Values for Optional Properties

Optional properties exist for the different drivers:

- bnx2 Driver Parameter
- bnx2x Driver Parameters
- bnx2i Driver Parameters
- bnx2fc Driver Parameter
- cnic Driver Parameters

For additional information about the drivers, see the associated README files.

bnx2 Driver Parameter

The `disable_msi` parameter can be supplied as a command line argument to the insmod or modprobe command for bnx2.

When set to `1` (enabled), this parameter disables MSI and MSI-X and uses the legacy INTx mode.

Marvell recommends setting the `disable_msi` parameter to `1` to always disable MSI/MSI-X on all QLogic adapters in the system. Issue one of the following commands:

```
insmod bnx2.ko disable_msi=1
```

```
modprobe bnx2 disable_msi=1
```

This parameter can also be set in the `modprobe.conf` file. See the man page for more information.

bnx2x Driver Parameters

Parameters for the bnx2x driver are described in the following sections.

int_mode

Use the optional parameter `int_mode` to force using an interrupt mode other than MSI-X. By default, the driver tries to enable MSI-X if it is supported by the kernel. If MSI-X is not attainable, the driver tries to enable MSI if it is supported by the kernel. If MSI is not attainable, the driver uses the legacy INTx mode.

To force using the legacy INTx mode on all 57*xx* and 57*xxx* network adapters in the system, set the `int_mode` parameter to `1` as shown in the following:

```
vmkload_mod bnx2x int_mode=1
```

To force using MSI mode on all 57*xx* and 57*xxx* network adapters in the system, set the `int_mode` parameter to `2` as shown in the following:

```
vmkload_mod bnx2x int_mode=2
```

disable_tpa

Use the optional parameter `disable_tpa` to disable the transparent packet aggregation (TPA) feature. By default, the driver aggregates TCP packets.

To disable the TPA feature on all 57*xx* and 57*xxx* network adapters in the system, set the `disable_tpa` parameter to `1`:

```
insmod bnx2x.ko disable_tpa=1
```

or

```
modprobe bnx2x disable_tpa=1
```

dropless_fc

The `dropless_fc` parameter is set to 1 (by default) to enable a complementary flow control mechanism on 57*xxx* adapters. The normal flow control mechanism is to send pause frames when the on-chip buffer (BRB) is reaching a specific level of occupancy, which is a performance-targeted flow control mechanism. On 57*xxx* adapters, you can enable a complementary flow control mechanism to send pause frames when one or more of the host receive buffers are exhausted.

`dropless_fc` is a “zero packet drop” targeted flow control mechanism.

Set the `dropless_fc` parameter to 1 to enable the drop-less flow control mechanism feature on all 57*xxx* adapters in the system.

```
insmod bnx2x.ko dropless_fc=1
```

or

```
modprobe bnx2x dropless_fc=1
```

autogreen

The autogreen parameter forces the specific AutoGrEEEN behavior. AutoGrEEEn is a proprietary, pre-IEEE standard Energy Efficient Ethernet (EEE) mode supported by some 1000BASE-T and 10GBASE-T RJ45 interfaced switches.

By default, the driver uses the NVRAM configuration settings per port. When this module parameter is set, it can override the NVRAM configuration settings to force AutoGrEEEN to either the active (1) or inactive (2) state. The default value of 0 sets the port to use the NVRAM settings.

native_eee

The native_eee parameter can force specific IEEE 802.3az Energy Efficient Ethernet (EEE) behavior, which is supported on some 1000BASE-T and 10GBASE-T RJ45 interfaced switches.

By default, the driver uses the NVRAM configuration settings per port. If this parameter is set, it can force EEE to be enabled, and the value will be used as the idle time (1-FFFFFh or 1,048,575) required before entering transmit LPI.

Set native_eee to -1 to forcefully disable EEE. Set native_eee to 0 (default) to use the NVRAM settings.

num_queues

The num_queues parameter forces the number of RSS queues and overrides the default value, which is equal to number of CPU cores.

pri_map

On earlier versions of Linux that do not support tc-mqprio, use the optional parameter pri_map to map the VLAN PRI value or the IP DSCP value to a different or the same class of service (CoS) in the hardware. This 32-bit parameter is evaluated by the driver as eight values of 4 bits each. Each nibble sets the required hardware queue number for that priority.

For example, set the pri_map parameter to 0x22221100 to map priority 0 and 1 to CoS 0, map priority 2 and 3 to CoS 1, and map priority 4–7 to CoS 2. In another example, set the pri_map parameter to 0x11110000 to map priority 0–3 to CoS 0, and map priority 4–7 to CoS 1.

tx_switching

The `tx_switching` parameter sets the L2 Ethernet send direction to test each transmitted packet. If the packet is intended for the transmitting NIC port, it is hair-pin looped back by the adapter.

This parameter is relevant only in multifunction (NPAR) mode, especially in virtualized environments.

full_promiscous

The `full_promiscous` parameter extends the existing promiscuous mode settings to accept all unmatched unicast packets on the interface.

By default, this parameter is disabled (set to `0`).

fairness_threshold

The `fairness_threshold` parameter enables firmware thresholds for physical functions (PFs) in multifunction (MF) mode where more than one PF is configured on a single, physical Ethernet port.

By default, this parameter is disabled (set to `0`).

poll

This optional debug parameter is used for timer-based polling.

mrrs

The `mrrs` optional debug parameter overrides the maximum read request size (MRRS) of the hardware. Valid values are in the range of 0–3.

use_random_vf_mac

When this parameter is enabled (set to `1`), all created VFs will have a random forced MAC.

By default, this parameter is disabled (set to `0`).

debug

The debug parameter sets the default message level (msglevel) on all adapters in the system at one time.

To set the message level for a specific adapter, issue the `ethtool -s` command.

bnx2i Driver Parameters

Optional parameters `en_tcp_dack`, `error_mask1`, and `error_mask2` can be supplied as command line arguments to the `insmod` or `modprobe` command for bnx2i.

error_mask1 and error_mask2

Use the `error_mask` (Configure firmware iSCSI error mask #) parameters to configure a specific iSCSI protocol violation to be treated either as a warning or a fatal error. All fatal iSCSI protocol violations will result in session recovery (ERL 0). These are bit masks.

Defaults: All violations are treated as errors.

CAUTION

Do not use `error_mask` if you are not sure about the consequences. These values are to be discussed with the Marvell development team on a case-by-case basis. This parameter is just a mechanism to work around iSCSI implementation issues on the target side and without proper knowledge of iSCSI protocol details, users are advised not to experiment with these parameters.

en_tcp_dack

The `en_tcp_dack` parameter enables and disables the TCP delayed ACK feature on offloaded iSCSI connections.

Defaults: TCP delayed ACK is ENABLED. For example:

```
insmod bnx2i.ko en_tcp_dack=0
```

or

```
modprobe bnx2i en_tcp_dack=0
```

time_stamps

The `time_stamps` parameter enables and disables the TCP time stamp feature on offloaded iSCSI connections.

Defaults: the TCP time stamp option is disabled. For example:

```
insmod bnx2i.ko time_stamps=1
```

or

```
modprobe bnx2i time_stamps=1
```

sq_size

Use the `sq_size` parameter to choose send queue size for offloaded connections and SQ size determines the maximum SCSI commands that can be queued. SQ size also has a bearing on the quantity of connections that can be offloaded; as QP size increases, the quantity of connections supported decreases. With the default values, the BCM5708 adapters can offload 28 connections.

Default: 128

Range: 32 to 128

Note that Marvell validation is limited to a power of 2; for example, 32, 64, and 128.

rq_size

Use the `rq_size` parameter to choose the size of asynchronous buffer queue size per offloaded connections. RQ size is not required greater than 16 as it is used to place iSCSI ASYNC/NOP/REJECT messages and SCSI sense data.

Default: 16

Range: 16 to 32

Note that Marvell validation is limited to a power of 2; for example, 16 or 32.

event_coal_div

The `event_coal_div` (event coalescing divide factor) parameter is a performance-tuning parameter that moderates the rate of interrupt generation by the iSCSI firmware.

Default: 2

Valid values: 1, 2, 4, 8

last_active_tcp_port

The `last_active_port` parameter is a status parameter that indicates the last TCP port number used in the iSCSI offload connection.

Default: N/A

Valid values: N/A

This parameter is read-only.

ooo_enable

The ooo_enable (enable TCP out-of-order) parameter feature enables and disables TCP out-of-order RX handling feature on offloaded iSCSI connections.

Default: TCP out-of-order feature is ENABLED. For example:

```
insmod bnx2i.ko ooo_enable=1
```

or

```
modprobe bnx2i ooo_enable=1
```

bnx2fc Driver Parameter

You can supply the optional parameter debug_logging as a command line argument to the insmod or modprobe command for bnx2fc.

debug_logging

The bit mask to enable debug logging enables and disables driver debug logging.

Default: None. For example:

```
insmod bnx2fc.ko debug_logging=0xff
```

or

```
modprobe bnx2fc debug_logging=0xff
```

I/O level debugging = 0x1

Session level debugging = 0x2

HBA level debugging = 0x4

ELS debugging = 0x8

Misc debugging = 0x10

Max debugging = 0xff

cnic Driver Parameters

To set the qcnic driver parameters, issue one of the following commands:

```
#esxcli system module parameters set -m qcnic -p Param=Value
#esxcfg-module -s <param>=<value> qcnic
```

cnic_debug

The cnic_debug parameter sets the driver debug message level. Valid values are in the range of 0h–8000000h. The default value is 0h.

cnic_dump_kwqe_enable

The `cnic_dump_kwe_en` parameter enables and disables single work-queue element message (kwqe) logging. By default, this parameter is set to `1` (disabled).

Driver Defaults

Default settings for the drivers are described in the following sections:

- bnx2 Driver Defaults
- bnx2x Driver Defaults

bnx2 Driver Defaults

Speed: Autonegotiation with all speeds advertised

Flow Control: Autonegotiation with RX and TX advertised

MTU: 1500 (range is 46–9000)

RX Ring Size: 255 (range is 0–4080)

RX Jumbo Ring Size: 0 (range is 0–16320) adjusted by the driver based on MTU and RX Ring Size

TX Ring Size: 255 (range is (MAX_SKB_FRAGS+1)–255). MAX_SKB_FRAGS varies on different kernels and different architectures. On a 2.6 kernel for x86, MAX_SKB_FRAGS is 18.

Coalesce RX Microseconds: 18 (range is 0–1023)

Coalesce RX Microseconds IRQ: 18 (range is 0–1023)

Coalesce RX Frames: 6 (range is 0–255)

Coalesce RX Frames IRQ: 6 (range is 0–255)

Coalesce TX Microseconds: 80 (range is 0–1023)

Coalesce TX Microseconds IRQ: 80 (range is 0–1023)

Coalesce TX Frames: 20 (range is 0–255)

Coalesce TX Frames IRQ: 20 (range is 0–255)

Coalesce Statistics Microseconds: 999936 (approximately 1 second) (range is 0–16776960 in increments of 256)

MSI: Enabled (if supported by the 2.6 kernel and the interrupt test passes)

TSO: Enabled (on 2.6 kernels)

WoL: Initial setting based on the NVRAM setting

bnx2x Driver Defaults

Speed: Autonegotiation with all speeds advertised

Flow control: Autonegotiation with RX and TX advertised

MTU: 1500 (range is 46–9600)

RX Ring Size: 4078 (range is 0–4078)

TX Ring Size: 4078 (range is (MAX_SKB_FRAGS+4)–4078). MAX_SKB_FRAGS varies on different kernels and different architectures. On a 2.6 kernel for x86, MAX_SKB_FRAGS is 18.

Coalesce RX Microseconds: 25 (range is 0–3000)

Coalesce TX Microseconds: 50 (range is 0–12288)

Coalesce Statistics Microseconds: 999936 (approximately 1 second) (range is 0–16776960 in increments of 256)

MSI-X: Enabled (if supported by the 2.6 kernel and the interrupt test passes)

TSO: Enabled

WoL: Disabled

Driver Messages

The following are the most common sample messages that may be logged in the `/var/log/messages` file. Issue the `dmesg -n <level>` command to control the level at which messages appear on the console. Most systems are set to level 6 by default. To see all messages, set the level higher.

- bnx2x Driver Messages
- bnx2i Driver Messages
- bnx2fc Driver Messages

bnx2x Driver Messages

The bnx2x driver messages include the following.

Driver Sign On

```
QLogic 57xx and 57xxx 10 Gigabit Ethernet Driver bnx2x v1.6.3c
(July 23, 2007)
```

C-NIC Driver Sign On (bnx2 only)

```
QLogic 57xx and 57xxx cnic v1.1.19 (Sep 25, 2007)
```

NIC Detected

```
eth#: QLogic 57xx and 57xxx xGb (B1)
PCI-E x8 found at mem f6000000, IRQ 16, node addr 0010180476ae

cnic: Added CNIC device: eth0
```

Link Up and Speed Indication

```
bnx2x: eth# NIC Link is Up, 10000 Mbps full duplex
```

Link Down Indication

```
bnx2x: eth# NIC Link is Down
MSI-X Enabled Successfully
bnx2x: eth0: using MSI-X
```

bnx2i Driver Messages

The bnx2i driver messages include the following.

BNX2I Driver Sign-on

```
QLogic 57xx and 57xxx iSCSI Driver bnx2i v2.1.1D (May 12, 2015)
```

Network Port to iSCSI Transport Name Binding

```
bnx2i: netif=eth2, iscsi=bcm570x-050000
bnx2i: netif=eth1, iscsi=bcm570x-030c00
```

Driver Completes Handshake with iSCSI Offload-enabled C-NIC Device

```
bnx2i [05:00.00]: ISCSI_INIT passed
```

NOTE

This message is displayed only when the user attempts to make an iSCSI connection.

Driver Detects iSCSI Offload Is Not Enabled on the C-NIC Device

```
bnx2i: iSCSI not supported, dev=eth3
bnx2i: bnx2i: LOM is not enabled to offload iSCSI connections,
dev=eth0
bnx2i: dev eth0 does not support iSCSI
```

Exceeds Maximum Allowed iSCSI Connection Offload Limit

```
bnx2i: alloc_ep: unable to allocate iscsi cid
bnx2i: unable to allocate iSCSI context resources
```

Network Route to Target Node and Transport Name Binding Are Two Different Devices

```
bnx2i: conn bind, ep=0x... ($ROUTE_HBA) does not belong to hba
$USER_CHOSEN_HBA
```

Where `ROUTE_HBA` is the net device on which the connection was offloaded based on route information, and `USER_CHOSEN_HBA` is the Host Bus Adapter to which the target node is bound (using iSCSI transport name)

Target Cannot Be Reached on Any of the C-NIC Devices

```
bnx2i: check route, cannot connect using cnic
```

Network Route Is Assigned to Network Interface, Which Is Down

```
bnx2i: check route, hba not found
```

SCSI-ML Initiated Host Reset (Session Recovery)

```
bnx2i: attempting to reset host, #3
```

C-NIC Detects iSCSI Protocol Violation - Fatal Errors

```
bnx2i: iscsi_error - wrong StatSN rcvd
bnx2i: iscsi_error - hdr digest err
bnx2i: iscsi_error - data digest err
bnx2i: iscsi_error - wrong opcode rcvd
bnx2i: iscsi_error - AHS len > 0 rcvd
bnx2i: iscsi_error - invalid ITT rcvd
bnx2i: iscsi_error - wrong StatSN rcvd
bnx2i: iscsi_error - wrong DataSN rcvd
bnx2i: iscsi_error - pend R2T violation
bnx2i: iscsi_error - ERL0, UO
bnx2i: iscsi_error - ERL0, U1
bnx2i: iscsi_error - ERL0, U2
bnx2i: iscsi_error - ERL0, U3
bnx2i: iscsi_error - ERL0, U4
bnx2i: iscsi_error - ERL0, U5
bnx2i: iscsi_error - ERL0, U
bnx2i: iscsi_error - invalid resi len
bnx2i: iscsi_error - MRDSL violation
```

```
bnx2i: iscsi_error - F-bit not set
bnx2i: iscsi_error - invalid TTT
bnx2i: iscsi_error - invalid DataSN
bnx2i: iscsi_error - burst len violation
bnx2i: iscsi_error - buf offset violation
bnx2i: iscsi_error - invalid LUN field
bnx2i: iscsi_error - invalid R2TSN field
bnx2i: iscsi_error - invalid cmd len1
bnx2i: iscsi_error - invalid cmd len2
bnx2i: iscsi_error - pend r2t exceeds MaxOutstandingR2T value
bnx2i: iscsi_error - TTT is rsvd
bnx2i: iscsi_error - MBL violation
bnx2i: iscsi_error - data seg len != 0
bnx2i: iscsi_error - reject pdu len error
bnx2i: iscsi_error - async pdu len error
bnx2i: iscsi_error - nopin pdu len error
bnx2i: iscsi_error - pend r2t in cleanup
bnx2i: iscsi_error - IP fragments rcvd
bnx2i: iscsi_error - IP options error
bnx2i: iscsi_error - urgent flag error
```

C-NIC Detects iSCSI Protocol Violation—Non-FATAL, Warning

```
bnx2i: iscsi_warning - invalid TTT
bnx2i: iscsi_warning - invalid DataSN
bnx2i: iscsi_warning - invalid LUN field
```

NOTE

You must configure the driver to consider a specific violation to treat as warning and not as a critical error.

Driver Puts a Session Through Recovery

```
conn_err - hostno 3 conn 03fbcd00, iscsi_cid 2 cid a1800
```

Reject iSCSI PDU Received from the Target

```
bnx2i - printing rejected PDU contents
[0]: 1 ffffffa1 0 0 0 0 20 0
[8]: 0 7 0 0 0 0 0 0
[10]: 0 0 40 24 0 0 ffffff80 0
[18]: 0 0 3 ffffff88 0 0 3 4b
```

```
[20]: 2a 0 0 2 ffffffc8 14 0 0
[28]: 40 0 0 0 0 0 0 0
```

Open-iSCSI Daemon Handing Over Session to Driver

```
bnx2i: conn update - MBL 0x800 FBL 0x800MRDSL_I 0x800 MRDSL_T
0x2000
```

bnx2fc Driver Messages

The bnx2fc driver messages include the following.

BNX2FC Driver Signon

```
QLogic FCoE Driver bnx2fc v0.8.7 (Mar 25, 2011)
```

Driver Completes Handshake with FCoE Offload Enabled C-NIC Device

```
bnx2fc [04:00.00]: FCOE_INIT passed
```

Driver Fails Handshake with FCoE Offload Enabled C-NIC Device

```
bnx2fc: init_failure due to invalid opcode
bnx2fc: init_failure due to context allocation failure
bnx2fc: init_failure due to NIC error
bnx2fc: init_failure due to completion status error
bnx2fc: init_failure due to HSI mismatch
```

No Valid License to Start FCoE

```
bnx2fc: FCoE function not enabled <ethX>
bnx2fC: FCoE not supported on <ethX>
```

Session Failures Due to Exceeding Maximum Allowed FCoE Offload Connection Limit or Memory Limits

```
bnx2fc: Failed to allocate conn id for port_id <remote port id>
bnx2fc: exceeded max sessions..logoff this tgt
bnx2fc: Failed to allocate resources
```

Session Offload Failures

```
bnx2fc: bnx2fc_offload_session - Offload error
<rport> not FCP type. not offloading
<rport> not FCP_TARGET. not offloading
```

Session Upload Failures

```
bnx2fc: ERROR!! destroy timed out
bnx2fc: Disable request timed out.  destroy not set to FW
bnx2fc: Disable failed with completion status <status>
bnx2fc: Destroy failed with completion status <status>
```

Unable to Issue ABTS

```
bnx2fc: initiate_abts: tgt not offloaded
bnx2fc: initiate_abts: rport not ready
bnx2fc: initiate_abts: link is not ready
bnx2fc: abort failed, xid = <xid>
```

Unable to Recover the IO Using ABTS (Due to ABTS Timeout)

```
bnx2fc: Relogin to the target
```

Unable to Issue I/O Request Due to Session Not Ready

```
bnx2fc: Unable to post io_req
```

Drop Incorrect L2 Receive Frames

```
bnx2fc: FPMA mismatch... drop packet
bnx2fc: dropping frame with CRC error
```

Host Bus Adapter and lport Allocation Failures

```
bnx2fc: Unable to allocate hba
bnx2fc: Unable to allocate scsi host
```

NPIV Port Creation

```
bnx2fc: Setting vport names, <WWNN>, <WWPN>
```

Teaming with Channel Bonding

With the Linux drivers, you can team adapters together using the bonding kernel module and a channel bonding interface. For more information, see the Channel Bonding information in your operating system documentation.

Statistics

Detailed statistics and configuration information can be viewed using the ethtool utility. See the ethtool man page for more information.

Linux iSCSI Offload

iSCSI offload information for Linux includes the following:

- Open iSCSI User Applications
- User Application iscsiuio
- Bind iSCSI Target to Marvell iSCSI Transport Name
- VLAN Configuration for iSCSI Offload (Linux)
- Making Connections to iSCSI Targets
- Maximum Offload iSCSI Connections
- Linux iSCSI Offload FAQ

Open iSCSI User Applications

Install and run the inbox Open-iSCSI initiator programs from the DVD. For details, refer to "Packaging" on page 35.

User Application iscsiuio

Install and run the iscsiuio daemon before attempting to create iSCSI connections. The driver cannot establish connections to the iSCSI target without the daemon's assistance.

To install and run the iscsiuio daemon:

1. Install the iscsiuio source package as follows:

   ```
   # tar -xvzf iscsiuio-<version>.tar.gz
   ```

2. CD to the directory where iscsiuio is extracted as follows:

   ```
   # cd iscsiuio-<version>
   ```

3. Compile and install as follows:

   ```
   # ./configure
   # make
   # make install
   ```

4. Ensure that the iscsiuio version matches the source package as follows:

   ```
   # iscsiuio -v
   ```

5. Start iscsiuio as follows:

   ```
   # iscsiuio
   ```

Bind iSCSI Target to Marvell iSCSI Transport Name

By default, the Open-iSCSI daemon connects to discovered targets using software initiator (`transport name = 'tcp'`). Users who want to offload iSCSI connection onto C-NIC device should explicitly change transport binding of the iSCSI iface. Perform the binding change using the iscsiadm CLI utility as follows,

```
iscsiadm -m iface -I <iface_file_name> -n iface.transport_name -v
bnx2i -o update
```

Where the `iface` file includes the following information for SLES:

```
iface.net_ifacename = ethX
iface.iscsi_ifacename = <name of the iface file>
iface.hwaddress = xx:xx:xx:xx:xx:xx
iface.ipaddress = XX.XX.XX.XX
iface.transport_name = bnx2i
```

Ensure that the `iface.hwaddress` is in lowercase format.

If you want to switch back to use the software initiator, enter the following:

```
iscsiadm -m iface -I <iface_file_name> -n iface.transport_name -v
tcp -o update
```

Where the `iface` file includes the following information:

```
iface.net_ifacename = ethX
iface.iscsi_ifacename = <name of the iface file>
iface.transport_name = tcp
```

VLAN Configuration for iSCSI Offload (Linux)

iSCSI traffic on the network may be isolated in a VLAN to segregate it from other traffic. When this is the case, you must make the iSCSI interface on the adapter a member of that VLAN.

To configure the iSCSI VLAN, add the VLAN ID in the `iface` file for iSCSI. In the following example, the VLAN ID is set to 100.

```
#Begin Record 6.2.0-873.2.el6
Iface.iscsi_ifacefile name = <>
Iface.ipaddress = 0.0.0.0
Iface.hwaddress = <>
Iface.trasport_name = bnx2i
Iface.vlan_id = 100
Iface.vlan_priority  = 0
Iface.iface_num = 100
Iface.mtu = 0
Iface.port = 0
#END Record
```

NOTE

Although not strictly required, Marvell recommends configuring the same VLAN ID on the `iface.iface_num` field for `iface` file identification purposes.

Making Connections to iSCSI Targets

Refer to Open-iSCSI documentation for a comprehensive list of `iscsiadm` commands. The following is a sample list of commands to discovery targets and to create iSCSI connections to a target.

Add Static Entry

```
iscsiadm -m node -p <ipaddr[:port]> -T
iqn.2007-05.com.qlogic:target1 -o new -I <iface_file_name>
```

iSCSI Target Discovery Using `sendtargets`

```
iscsiadm -m discovery --type sendtargets -p <ipaddr[:port]> -I
<iface_file_name>
```

Login to Target Using `iscsiadm` Command

```
iscsiadm --mode node --targetname <iqn.targetname> --portal
<ipaddr[:port]> --login
```

List All Drives Active in the System

```
fdisk -l
```

Maximum Offload iSCSI Connections

With default driver parameters set, which includes 128 outstanding commands, bnx2i can offload 128 connections on Marvell 5771*x* adapters.

This quantity is not a hard limit, but just a simple on-chip resource allocation math. The bnx2i can offload more connections by reducing the shared queue size, which in turn limits the maximum outstanding tasks on a connection. See “Setting Values for Optional Properties” on page 44 for information on `sq_size` and `rq_size`. When the maximum allowed connection offload limit is reached, the driver logs the following message to `syslog`:

```
bnx2i: unable to allocate iSCSI context resources
```

Linux iSCSI Offload FAQ

- Not all Marvell 57*xx* and 57*xxx* adapters support iSCSI offload.
- The iSCSI session will not recover after a hot remove and hot plug.
- For Microsoft Multipath I/O (MPIO) to work properly, you must enable iSCSI `noopout` on each iSCSI session. For procedures on setting up `noop_out_interval` and `noop_out_timeout` values, refer to Open-iSCSI documentation.
- In the scenario where multiple C-NIC devices are in the system and the system is booted through the Marvell iSCSI boot solution, ensure that the `iscsi` node under `/etc/iscsi/nodes` for the boot target is bound to the NIC that is used for booting.

8 VMware Driver Software

This chapter covers the following for the VMware driver software:

- Introduction
- "Packaging" on page 63
- "Download, Install, and Update Drivers" on page 64
- "FCoE Support" on page 83
- "iSCSI Support" on page 86

NOTE

Information in this chapter applies primarily to the currently supported VMware versions: ESXi 6.7 and ESXi 7.0. ESXi 6.7 uses native drivers for all protocols.

Introduction

This section describes the VMware ESXi drivers for the Marvell 57*xx* and 57*xxx* PCIe 1/10GbE network adapters. It provides information on downloading, installing, and updating VMware drivers, describes the driver parameters and defaults, provides information on unloading and removing drivers, and describes driver messages.

The VMware ESXi drivers are listed in Table 8-1.

Table 8-1. Marvell 57xx and 57xxx VMware Drivers

VMware Driver	Description
bnx2	VMware legacy driver for the 57*xx* 1Gb network adapters. This driver directly controls the hardware and is responsible for sending and receiving Ethernet packets on behalf of the VMware host networking stack. The driver also receives and processes device interrupts, both on behalf of itself (for Layer 2 networking) and on behalf of the C-NIC driver (for iSCSI offload).

Table 8-1. Marvell 57xx and 57xxx VMware Drivers (Continued)

VMware Driver	Description
bnx2x	VMware legacy driver for the 57*xxx* 1/10Gb network adapters. This driver directly controls the hardware and is responsible for sending and receiving Ethernet packets on behalf of the VMware host networking stack. The driver also receives and processes device interrupts, both on behalf of itself (for Layer 2 networking) and on behalf of the C-NIC driver (for FCoE offload and iSCSI offload).
cnic	VMware C-NIC legacy driver. This driver provides the interface between Marvells upper-layer protocol (for example, storage) legacy drivers and Marvell's 57*xx* and 57*xxx* 1/10Gb network adapters. The C-NIC module works with the bnx2 and bnx2x legacy network drives in the downstream and the bnx2fc (FCoE) and bnx2i (iSCSI) legacy drivers in the upstream.
bnx2i	VMware iSCSI offload HBA legacy driver. This driver enables iSCSI offload on the 57*xx* and 57*xxx* 1Gb/10Gb network adapters.
bnx2fc	VMware FCoE offload HBA legacy driver. This driver enables FCoE offload on the 57712/578*xx* 10Gb converged network adapters.
qflge	VMware native driver for the 57*xx* 1Gb network adapters. This driver directly controls the hardware and is responsible for sending and receiving Ethernet packets on behalf of the VMware host networking stack.
qfle3	VMware native driver for the 57xxx 1Gb/10Gb network adapters. This driver directly controls the hardware and is responsible for sending and receiving Ethernet packets on behalf of the VMware host networking stack.
qfle3i	VMware iSCSI offload HBA native driver. This driver enables iSCSI offload on the 57*xx* and 57*xxx* 1/10Gb network adapters.
qfle3f	VMware FCoE offload HBA native driver. This driver enables FCoE offload on the 57712/578*xx* 10Gb adapters. This driver automatically starts the FCoE initialization process; you do not need to perform any manual steps.

Packaging

The driver package offline bundle Depot Zip file is inside the downloaded zip file. Consequently, you need to unzip the downloaded file (from VMware) to get to the applicable offline bundle Depot Zip file package before copying it to your VMware server.

The VMware driver is released in the packaging formats shown in Table 8-2.

Table 8-2. VMware Driver Packaging

Format	Drivers
Compressed ZIP	`QLG-qcnic-6.7-offline_bundle-<version>.zip` (native ESXi 6.7)
Compressed ZIP	`QLG-qcnic-7.0-offline_bundle-<version>.zip` (native ESXi 7.0)

Download, Install, and Update Drivers

To download, install, or update the VMware ESXi drivers for 57*xx* and 57*xxx* 10GbE network adapters, see http://www.vmware.com/support. This package is double zipped—unzip the package before copying it to the ESXi host.

Driver Parameters

The following sections describe the parameters for these drivers:

- bnx2 Driver Parameters
- bnx2x Driver Parameters
- cnic Driver Parameters
- bnx2i Driver Parameters
- bnx2fc Driver Parameter
- qcnic Driver Parameters
- qfle3 Driver Parameters
- qfle3i Driver Parameters
- qfle3f Driver Parameters

bnx2 Driver Parameters

disable_msi

The `disable_msi` parameter can be supplied as a command line argument to the insmod or modprobe command for bnx2.

When set to `1` (enabled), this parameter disables MSI and MSI-X and uses the legacy INTx mode.

Marvell recommends setting the `disable_msi` parameter to `1` to always disable MSI/MSI-X on all QLogic adapters in the system. Issue one of the following commands:

```
insmod bnx2.ko disable_msi=1

modprobe bnx2 disable_msi=1
```

This parameter can also be set in the `modprobe.conf` file. See the man page for more information.

bnx2x Driver Parameters

You can supply several optional parameters as a command line argument to the `vmkload_mod` command. Set these parameters by issuing the `esxcfg-module` command. For more information, issue the command: **`esxcfg-module -h`**.

int_mode

Use the optional parameter `int_mode` to force using an interrupt mode other than MSI-X. By default, the driver tries to enable MSI-X if it is supported by the kernel. If MSI-X is not attainable, the driver tries to enable MSI if it is supported by the kernel. If MSI is not attainable, the driver uses the legacy INTx mode.

To force using the legacy INTx mode on all 57*xx* and 57*xxx* network adapters in the system, set the `int_mode` parameter to `1` as shown in the following:

```
vmkload_mod bnx2x int_mode=1
```

To force using MSI mode on all 57*xx* and 57*xxx* network adapters in the system, set the `int_mode` parameter to `2` as shown in the following:

```
vmkload_mod bnx2x int_mode=2
```

disable_tpa

Use the optional parameter `disable_tpa` to disable the transparent packet aggregation (TPA) feature. By default, the driver aggregates TCP packets.

To disable the TPA feature on all 57*xx* and 57*xxx* network adapters in the system, set the `disable_tpa` parameter to `1`:

```
vmkload_mod bnx2x disable_tpa=1
```

Use ethtool to disable TPA (LRO) for a specific network adapter.

dropless_fc

The dropless_fc parameter is set to 1 (by default) to enable a complementary flow control mechanism on 57*xxx* adapters. The normal flow control mechanism is to send pause frames when the on-chip buffer (BRB) is reaching a specific level of occupancy, which is a performance-targeted flow control mechanism. On 57*xxx* adapters, you can enable a complementary flow control mechanism to send pause frames when one or more of the host receive buffers are exhausted.

dropless_fc is a “zero packet drop” targeted flow control mechanism.

Set the dropless_fc parameter to 1 to enable the drop-less flow control mechanism feature on all 57*xxx* adapters in the system.

```
vmkload_mod bnx2x dropless_fc=1
```

autogreen

The autogreen parameter forces the specific AutoGrEEEN behavior. AutoGrEEEn is a proprietary, pre-IEEE standard Energy Efficient Ethernet (EEE) mode supported by some 1000BASE-T and 10GBASE-T RJ45 interfaced switches.

By default, the driver uses the NVRAM configuration settings per port. When this module parameter is set, it can override the NVRAM configuration settings to force AutoGrEEEN to either the active (1) or inactive (2) state. The default value of 0 sets the port to use the NVRAM settings.

native_eee

The native_eee parameter can force specific IEEE 802.3az Energy Efficient Ethernet (EEE) behavior, which is supported on some 1000BASE-T and 10GBASE-T RJ45 interfaced switches.

By default, the driver uses the NVRAM configuration settings per port. If this parameter is set, it can force EEE to be enabled, and the value will be used as the idle time (1-FFFFFh or 1,048,575) required before entering transmit LPI.

Set native_eee to -1 to forcefully disable EEE. Set native_eee to 0 (default) to use the NVRAM settings.

num_queues

The num_queues parameter forces the number of RSS queues and overrides the default value, which is equal to number of CPU cores.

pri_map

On earlier versions of Linux that do not support `tc-mqprio`, use the optional parameter `pri_map` to map the VLAN PRI value or the IP DSCP value to a different or the same class of service (CoS) in the hardware. This 32-bit parameter is evaluated by the driver as eight values of 4 bits each. Each nibble sets the required hardware queue number for that priority.

For example, set the `pri_map` parameter to `0x22221100` to map priority 0 and 1 to CoS 0, map priority 2 and 3 to CoS 1, and map priority 4–7 to CoS 2. In another example, set the `pri_map` parameter to `0x11110000` to map priority 0–3 to CoS 0, and map priority 4–7 to CoS 1.

tx_switching

The `tx_switching` parameter sets the L2 Ethernet send direction to test each transmitted packet. If the packet is intended for the transmitting NIC port, it is hair-pin looped back by the adapter.

This parameter is relevant only in multifunction (NPAR) mode, especially in virtualized environments.

full_promiscous

The `full_promiscous` parameter extends the existing promiscuous mode settings to accept all unmatched unicast packets on the interface.

By default, this parameter is disabled (set to `0`).

fairness_threshold

The `fairness_threshold` parameter enables firmware thresholds for physical functions (PFs) in multifunction (MF) mode where more than one PF is configured on a single, physical Ethernet port.

By default, this parameter is disabled (set to `0`).

poll

This optional debug parameter is used for timer-based polling.

MRSS

The `mrrs` optional debug parameter overrides the maximum read request size (MRRS) of the hardware. Valid values are in the range of 0–3.

use_random_vf_mac

When this parameter is enabled (set to 1), all created VFs will have a random forced MAC.

By default, this parameter is disabled (set to 0).

debug

The debug parameter sets the default message level (msglevel) on all adapters in the system at one time.

To set the message level for a specific adapter, issue the `ethtool -s` command.

RSS

Use the optional `RSS` parameter to specify the quantity of receive side scaling queues. `RSS= -1` disables RSS queues.

max_vfs

Use the optional parameter `max_vfs` to enable a specific quantity of virtual functions. Values for `max_vfs` can be 1 to 64, or set `max_vfs=0` (default) to disable all virtual functions.

enable_vxlan_ofld

Use the optional parameter `enable_vxlan_ofld` to enable VXLAN task offloads with TX TSO and TX CSO.

enable_default_queue_filters

Use the optional parameter `enable_default_queue_filters` to enable the classification filters on the default queue. The hardware supports a total of 512 classification filters that are equally divided among the ports of an adapter. For example, a quad-port adapter has 128 filters per port. For NPAR configuration, filters are applied on the default queue to support traffic switching between the partitions belonging to the same physical port.

When the quantity of filters exceeds the hardware limits, the message, `Rx filters on NetQ Rx Queue 0 exhausted` appears in the vmkernel logs. The message indicates that the hardware filter limit was reached and no further entries can be added. You can disable filters on the default queue by setting the `enable_default_queue_filters` parameter to 0, which disables traffic switching between the partitions.

enable_live_grcdump

Use the `enable_live_grcdump parameter` to indicate which firmware dump is collected for troubleshooting. Valid values are:

Value	Description
0x0	Disable live global register controller (GRC) dump
0x1	Enable parity/live GRC dump (default)
0x2	Enable transmit timeout GRC dump
0x4	Enable statistics timeout GRC dump

The default setting is appropriate for most situations. Do not change the default value unless requested by the support team.

cnic Driver Parameters

To set the qcnic driver parameters, issue one of the following commands:

```
#esxcli system module parameters set -m qcnic -p Param=Value
#esxcfg-module -s <param>=<value> qcnic
```

cnic_debug

The `cnic_debug` parameter sets the driver debug message level. Valid values are in the range of 0h–8000000h. The default value is 0h.

cnic_dump_kwqe_enable

The `cnic_dump_kwe_en` parameter enables and disables single work-queue element message (kwqe) logging. By default, this parameter is set to `1` (disabled).

bnx2i Driver Parameters

Optional parameters `en_tcp_dack`, `error_mask1`, and `error_mask2` can be supplied as command line arguments to the `insmod` or `modprobe` command for bnx2i.

error_mask1 and error_mask2

Use the `error_mask` (Configure firmware iSCSI error mask #) parameters to configure a specific iSCSI protocol violation to be treated either as a warning or a fatal error. All fatal iSCSI protocol violations will result in session recovery (ERL 0). These are bit masks.

Defaults: All violations are treated as errors.

CAUTION

Do not use `error_mask` if you are not sure about the consequences. These values are to be discussed with the Marvell development team on a case-by-case basis. This parameter is just a mechanism to work around iSCSI implementation issues on the target side and without proper knowledge of iSCSI protocol details, users are advised not to experiment with these parameters.

en_tcp_dack

The `en_tcp_dack` parameter enables and disables the TCP delayed ACK feature on offloaded iSCSI connections.

Defaults: TCP delayed ACK is ENABLED. For example:

```
insmod bnx2i.ko en_tcp_dack=0
```

or

```
modprobe bnx2i en_tcp_dack=0
```

time_stamps

The `time_stamps` parameter enables and disables the TCP time stamp feature on offloaded iSCSI connections.

Defaults: the TCP time stamp option is disabled. For example:

```
insmod bnx2i.ko time_stamps=1
```

or

```
modprobe bnx2i time_stamps=1
```

sq_size

Use the sq_size parameter to choose send queue size for offloaded connections and SQ size determines the maximum SCSI commands that can be queued. SQ size also has a bearing on the quantity of connections that can be offloaded; as QP size increases, the quantity of connections supported decreases. With the default values, the BCM5708 adapters can offload 28 connections.

Default: 128

Range: 32 to 128

Note that Marvell validation is limited to a power of 2; for example, 32, 64, and 128.

rq_size

Use the rq_size parameter to choose the size of asynchronous buffer queue size per offloaded connections. RQ size is not required greater than 16 as it is used to place iSCSI ASYNC/NOP/REJECT messages and SCSI sense data.

Default: 16

Range: 16 to 32

Note that Marvell validation is limited to a power of 2; for example, 16 or 32.

event_coal_div

The event_coal_div (event coalescing divide factor) parameter is a performance-tuning parameter that moderates the rate of interrupt generation by the iSCSI firmware.

Default: 2

Valid values: 1, 2, 4, 8

Event Coalescing Divide Factor is a performance-tuning parameter used to moderate the rate of interrupt generation by the iSCSI firmware.

Default: 2

Valid values: 1, 2, 4, 8

last_active_tcp_port

The last_active_port parameter is a status parameter that indicates the last TCP port number used in the iSCSI offload connection.

Default: N/A

Valid values: N/A

This parameter is read-only.

ooo_enable

The ooo_enable (enable TCP out-of-order) parameter feature enables and disables TCP out-of-order RX handling feature on offloaded iSCSI connections.

Default: TCP out-of-order feature is ENABLED. For example:

```
insmod bnx2i.ko ooo_enable=1
```

or

```
modprobe bnx2i ooo_enable=1
```

bnx2fc Driver Parameter

You can supply the optional parameter debug_logging as a command line argument to the insmod or modprobe command for bnx2fc.

debug_logging

The bit mask to enable debug logging enables and disables driver debug logging.

Default: None. For example:

```
insmod bnx2fc.ko debug_logging=0xff
```

or

```
modprobe bnx2fc debug_logging=0xff
```

I/O level debugging = 0x1

Session level debugging = 0x2

HBA level debugging = 0x4

ELS debugging = 0x8

Misc debugging = 0x10

Max debugging = 0xff

qcnic Driver Parameters

To set the qcnic driver parameters, issue one of the following commands:

```
#esxcli system module parameters set -m qcnic -p Param=Value
#esxcfg-module -s <param>=<value> qcnic
```

cnic_debug

The cnic_debug parameter sets the driver debug message level. Valid values are in the range of 0h–8000000h. The default value is 0h.

cnic_dump_kwqe_en

The `cnic_dump_kwe_en` parameter enables and disables single work-queue element message (kwqe) logging. By default, this parameter is set to `1` (disabled).

qfle3 Driver Parameters

For a list of valid parameters, issue one of the following commands:

```
# esxcli system module parameters list -m qfle3
# esxcfg-module -i qfle3
```

To change a parameter, issue one of the following commands:

```
#esxcli system module parameters set -m qedentv -p Param=Value
#esxcfg-module -s Param=Value qfle3
```

debug_mask

Set the `debug_mask` module parameter only for debug purposes, as the additional logging will flood numerous messages. Marvell does not recommend setting this parameter for regular driver use.

The valid values for `debug_mask` are:

```
0x00000001     /* load and unload    */
0x00000002     /* interrupt handling */
0x00000004     /* slowpath handling  */
0x00000008     /* stats updates      */
0x00000010     /* packet transmit    */
0x00000020     /* packet receive     */
0x00000040     /* phy/link handling  */
0x00000080     /* not used    */
0x00000100     /* dumping mbuf info  */
0x00000200     /* register access    */
0x00000400     /* lro processing     */
0x00000800     /* uplink debug       */
0x00001000     /* qeueu debug       */
0x00002000     /* hw debug       */
0x00004000     /* cmp debug  */
0x00008000     /* start process debug  */
0x00010000     /* debug assert       */
0x00020000     /* debug poll       */
0x00040000     /* debug TXSG       */
0x00080000     /* debug crash       */
```

```
0x00100000     /* debug vlan       */
0x00200000     /* state machine      */
0x00400000     /* nvm access      */
0x00800000     /* SRIOV        */
0x01000000     /* mgmt interface    */
0x02000000     /* CNIC */
0x04000000     /* DCB */

0xFFFFFFFF     /* all enabled */
```

enable_fwdump

The `enable_fwdump` parameter enable and disables the firmware dump file. Set to `1` to enable the firmware dump file. Set to `0` (default) to disable the firmware dump file.

enable_lro

The `enable_Iro` parameter enables and disables the TPA (LRO) feature. Set to `0` to disable TPA. Set to `1` (default) to enable TPA.

hw_vlan

The `hw_vlan` parameter enables and disables VLAN removal/insertion by hardware. Set to `0` to disable VLAN removal/insertion. Set to `1` (default) to enable VLAN removal/insertion.

intr_mode

The `intr_mode` parameter sets the interrupt mode:

Value	Mode
0	Auto (default)
1	IRQ
2	MSI
3	MSI-X

mtu

This parameter specifies the MTU when the driver is loaded. Valid values are in the range of `0-9000`. (default: `1500`)

offload_flags

This parameter specifies the offload flags:

Value	Flag
1	CSO
2	TSO
4	VXLAN offload
8	Geneve offload
15	Default. All tunneled offloads (CSO, TSO, VXLAN, Geneve) are enabled.

rx_filters

The `rx_filters` parameter defines the number of receive filters per NetQueue. Set to `1` to use the default number of receive filters based on availability. Set to `0` to disable use of multiple receive filters. Set to a value in the range of `1`, `2`, `3`, and so on to force the number of receive filters used for NetQueue. The default is `-1`.

rxqueue_nr

The `rxqueue_nr` parameter sets the number of receive queues. Set to `0` (default) for Auto. Set to a number in the range of `1-8` for the number of fixed queues. The default is `4` queues.

rxring_bd_nr

The `rxring_bd_nr` parameter sets the number of receive buffer descriptors (BDs). The minimum value is 4,096 (default). The maximum value is 16,384. Values are round up to nearest power of two.

txqueue_nr

The `txqueue_nr` parameter sets the number of transmit queues. Set to `0` for Auto. Set to a value in the range of `1-8` for the number of fixed queues. The default is `4` queues.

txring_bd_nr

The `txring_bd_nr` parameter sets the number of transmit BDs. the minimum value is `4,096` (default). The maximum value is `16,384`. Values are round up to nearest power of two.

RSS

The `RSS` parameter sets the number of RSS queues. Set to `0` (default) to allow VMware to automatically control the number of RSS queues used by VXLAN tunneled traffic and host traffic. Set to a value in the range of `1-4` to indicate a fixed queue number.

DRSS

The DRSS parameter sets the number of RSS queues associated with the default queue. The minimum number of RSS queues is 2; the maximum number is 4. To disable this parameter, set it to 0 (default).

This parameter is used for VXLAN gateways, where multiple unknown MAC addresses may be received by the default queue.

rss_engine_nr

The rss_engine_nr parameter sets the number of RSS engines. Valid values are 0 (Disabled) or 1-4 (fixed number of RSS engines). The default is 4 RSS engines.

enable_vxlan_filters

The enable_vxlan_filters parameter enables and disables the VXLAN receive filters.

A VXLAN filter comprises the inner MAC address, the outer MAC address, and the VXLAN Network Identifier (VNI). This filter is used to create NetQueues for a VXLAN traffic flow.

Set to 0 (default) to disable VXLAN receive filters. Set to 1 to enable VXLAN receive filters.

dropless_fc

The dropless_fc parameter is set to 1 (by default) to enable a complementary flow control mechanism on 57*xxx* adapters. The normal flow control mechanism is to send pause frames when the on-chip buffer (BRB) is reaching a specific level of occupancy, which is a performance-targeted flow control mechanism. On 57*xxx* adapters, you can enable a complementary flow control mechanism to send pause frames when one or more of the host receive buffers are exhausted.

dropless_fc is a “zero packet drop” targeted flow control mechanism.

Set the dropless_fc parameter to 1 to enable the drop-less flow control mechanism feature on all 57*xxx* adapters in the system.

max_vfs

The max_vfs parameter indicates the number of virtual functions (VFs) to be enabled for each PCI function. Valid values are in the range of 0-164. A value of 0 disables this feature. A value in the range of 1-64 indicates the number of VFs to enable. The actual maximum VF count depends on the on 57*xxx* adapter hardware.

auto_recovery

The auto_recovery parameter enables or disables automatic recovery of the interface after detecting a hardware error. Set auto_recovery to 1 (default) to enable automatic recovery of the interface.

psod_on_error

The `psod_on_error` parameter indicates if the host panics when the interface detects an error. The default setting is 0 (the host does not panic). Set this parameter to `1` for the host to panic when the interface detects an error.

qfle3i Driver Parameters

For a list of qlfe3i driver parameters, issue one of the following commands:

```
# esxcli system module parameters list -m qfle3i
# esxcfg-module -i qfle3i
```

To change a parameter's value, issue one of the following commands:

```
#esxcli system module parameters set -m qfle3i -p <param>=<value>
#esxcfg-module -s <parameter>=<value> qfle3i
```

qfle3i_chip_cmd_max

The `qlfe3i_chip_cmd_max` parameter sets the maximum I/Os queued to the 57*xx* and 57*xxx* adapters. The default is `24`.

qfle3i_esx_mtu_max

The `qfle3i_esx_mtu_max` parameter sets the maximum MTU size supported for offload sessions. Valid values are in the range of `1500-9000`. The default is `9000`.

qfle3i_max_sectors

The `qfle3i_max_sectors` parameter sets the maximum sectors supported by the driver. Valid values are in the range of `64-256`. Set this parameter to `-1` for default values of `256` (10Gb) and `127` for 1Gb.

qfle3i_max_task_pgs

The `qfle3i_max_task_pgs` parameter sets the maximum number of pages (per connection) for iSCSI tasks. Valid values are in the range of `2-8`. The default value is `2`.

qfle3i_nopout_when_cmds_active

The `qfle3i_nopout_when_cmds_active` parameter sends an iSCSI NOP Out PDU even when the connection is active (not idle). Valid values are in the range of `2-8`. The default value is `1`.

cmd_cmpl_per_work

The `qfle3i_cmd_cmpl_per_work` parameter sets the number of command queue entries (CQEs) processed per work. The default value is `256`.

en_hba_poll

The en_hba_poll parameter sets the adapter poll timer. The default value is 0.

en_tcp_dack

The en_tcp_dack parameter enables TCP delayed ACK. Enabling TCP delayed ACK helps improve network performance by combining several ACKs in a single response. The default value is 1 (enabled).

Certain iSCSI targets do not handle ACK piggybacking. If this parameter is enabled on these types of targets, the host cannot login to the target. If this even occurs, Marvell recommends disabling this parameter.

error_mask1, error_mask2

Use the error_mask (Configure firmware iSCSI error mask #) parameters to configure a specific iSCSI protocol violation to be treated either as a warning or a fatal error. All fatal iSCSI protocol violations will result in session recovery (ERL 0). These are bit masks.

Defaults: All violations are treated as errors.

CAUTION

Do not use error_mask if you are not sure about the consequences. These values are to be discussed with the Marvell development team on a case-by-case basis. This parameter is just a mechanism to work around iSCSI implementation issues on the target side and without proper knowledge of iSCSI protocol details, users are advised not to experiment with these parameters.

event_coal_div

The event_coal_div parameter sets the event coalescing divide factor. The default value is 1.

event_coal_min

The event_coal_min parameter sets the minimum number of event-coalescing commands. The default is 24.

ooo_enable

The ooo_enable (enable TCP out-of-order) parameter feature enables and disables TCP out-of-order RX handling feature on offloaded iSCSI connections. Set to 0 to disable this support. Set to 1 (default) to enable this support.

qfle3i_debug_level

The qfle3i_debug_level parameter is a bit mask that enables and disables debug logs. The default is 0 (disabled).

The following debug logs can be masked:

Log	Value (h)
DEFAULT_LEVEL	001
Initialization	002
Conn Setup	004
TMF	008
iSCSI NOP	010
CNIC IF	020
ITT CLEANUP	040
CONN EVT	080
SESS Recovery	100
Internal	200
IO Path	400
APP INTERFACE	800

rq_size

Use the rq_size parameter to choose the size of asynchronous buffer queue size per offloaded connections. RQ size is not required greater than 16 as it is used to place iSCSI ASYNC/NOP/REJECT messages and SCSI sense data.

Default: 16

Range: 16 to 32

Note that Marvell validation is limited to a power of 2; for example, 16 or 32.

sq_size

Use the sq_size parameter to choose send queue size for offloaded connections and SQ size determines the maximum SCSI commands that can be queued. SQ size also has a bearing on the quantity of connections that can be offloaded; as QP size increases, the quantity of connections supported decreases. With the default values, the BCM5708 adapters can offload 28 connections.

Default: 128

Range: 32 to 128

Note that Marvell validation is limited to a power of 2; for example, 32, 64, and 128.

tcp_buf_size

The `tcp_buf_size` parameter sets the TCP send and receive buffer size. The default is `64 × 1,024`.

time_stamps

The `time_stamps` parameter enables and disables TCP time stamps. Set to `0` to disable time stamps. Set to `1` (default) to enable time stamps.

qfle3f Driver Parameters

To view all the qlfe3f parameters, issue one of the following commands:

```
# esxcli system module parameters list -m qfle3f
# esxcfg-module -i qfle3f
```

To set a parameter, issue one of the following commands:

```
#esxcli system module parameters set -m qfle3f -p Param=Value
#esxcfg-module -s Param=Value qfle3f
```

qfle3f_debug_level

The `qfle_3f_debug_level` parameter enables additional messaging from the driver. Set to `0` (default) to disable additional messaging. Set to `1` to enable additional messaging.

qfle3f_devlOSs_tmo

The `qfle3f_devlOSs_tmo` parameter sets the remote LUN device loss time-out value (in seconds). The default is `20` seconds. Valid values are in the range of 1–120 seconds.

qfle3f_max_luns

The `qfle3f_max_luns` parameter adjusts the maximum number of LUNs supported by the driver. The default value is FFFFh (65,535 LUNs).

qfle3f_queue_depth

The `qfle3f_queue_depth` parameter adjusts the maximum queue depth per LUN. By default, the OS settings are used.

qfle3f_enable_r_a_tov

The `qfle3f_enable_r_a_tov` parameter enables or disables a user-defined R_A_TOV. Set to `0` to disable R_A_TOV. Set to `1` (default) to enable R_A_TOV.

qfle3f_r_a_tov

When the `qfle3f_enable_r_a_tov` parameter is set to `1`, the `qfle3f_r_a_tov` parameter sets the value of a user-defined R_A_TOV. The default value is `10`.

qfle3f_autodiscovery

The `qfle3f_autodiscovery` parameter controls auto-FCoE discovery during system boot. Set to `0` (default) to disable auto-FCoE discovery. Set to `1` to enable auto-FCoE discovery.

qfle3f_create vmkMgmt_Entry

The `qfle3f_createvmkMgmt_Entry` parameter creates the vmkMgmt interface. Set to `0` if the vmkMgmt interface will not be used. Set to `1` (default) to create the vmkMgmt interface.

Driver Defaults

The following sections list the defaults for the Ethernet drivers.

bnx2

Defaults for the bnx2 VMware ESXi driver are listed in Table 8-3.

Table 8-3. bnx2 Driver Defaults

Parameter	Default
Speed	Autonegotiation with all speeds advertised
Flow Control	Autonegotiation with Rx and Tx advertised
MTU	1500 (range 46–9000)
Rx Ring Size	255 (range 0–4080)
Rx Jumbo Ring Size	0 (range 0–16320) automatically adjusted by the driver based on MTU and Rx Ring Size
Tx Ring Size	255 (range (MAX_SKB_FRAGS+1) – 255) MAX_SKB_FRAGS varies on different kernels and different architectures. On a 2.6/3.x kernel for x86, MAX_SKB_FRAGS is 18.
Number of RSS Channels	Varies depending on the number of CPUs (range 1–8).
Number of TSS Channels	Varies depending on the number of CPUs (range 1–8).
Coalesce Rx µsecs	18 (range 0–1023)
Coalesce Rx µsecs IRQ	18 (range 0–1023)
Coalesce Rx frames	12 (range 0–255)

Table 8-3. bnx2 Driver Defaults (Continued)

Parameter	Default
Coalesce Rx frames IRQ	2 (range 0–255)
Coalesce Tx µsecs	80 (range 0–1023)
Coalesce Tx µsecs IRQ	18 (range 0–1023)
Coalesce Tx frames	20 (range 0–255)
Coalesce Tx frames IRQ	2 (range 0–255)
Coalesce stats µsecs	999936 (approximately 1 second) (range 0–16776960 in 256 increments)
MSI/MSI-X	Enabled (if supported by 2.6/3.*x* kernel and interrupt test passes)
TSO	Enabled on 2.6/3.*x* kernels
WoL	Initial setting based on NVRAM's setting.

qfle3

Defaults for the qlfe3 VMware ESXi driver are listed in Table 8-4.

Table 8-4. qfle3 Driver Defaults

Parameter	Default
Firmware Dump File	Disabled
TPA (LRO)	Enabled
VLAN Removal/Insertion by Hardware	Enabled
Interrupt Mode	Auto
MTU	1500 (range 0–9,000) (ESXi 7.0 range 0–9,190)
Offload Flags	15
Number of RSS Queues	Auto
Number of RX Filters per NetQueue	–1 (range 0–…)
Number of Rx Queues	Auto
Number of Rx BD Buffers	4,096 (16,384 maximum)
Number of Tx Queues	4 (range 1–8)

Table 8-4. qfle3 Driver Defaults (Continued)

Parameter	Default
Number of Tx BD Buffers	4,096 (16,384 maximum)
Number of RSS Queues for Default Queue	0 (Disabled) (2 minimum; 4 maximum)
Number of RSS Engines	4 (range 0–4)
VXLAN Filters	Disabled
Pause on Exhausted Host Ring	Disabled
Number of VFs per PCI Function	0 (Disabled) (range 1–64)

Unloading and Removing Drivers

The following sections describe how to remove the Ethernet drivers.

qfle3

To remove the driver package, issue the following command:

```
#esxcli software vib remove --vibname <vib-name>
```

For example:

```
esxcli software vib remove --vibname qfle3
```

To unload the driver temporarily, issue the following command:

```
#vmkload_mod -u qfle3
```

FCoE Support

This section describes the contents and procedures associated with installation of the VMware software package for supporting Marvell FCoE C-NICs.

Drivers

Marvell 57712/578*xx* FCoE drivers include the bnx2x and the bnx2fc.

- The **bnx2x** driver manages all PCI device resources (registers, host interface queues, and so on.) and also acts as the Layer 2 VMware low-level network driver for Marvell's 57*xx* and 57*xxx* 10G device. This driver directly controls the hardware and is responsible for sending and receiving Ethernet packets on behalf of the VMware host networking stack. The bnx2x driver also receives and processes device interrupts, both on behalf of itself (for Layer 2 networking) and on behalf of the bnx2fc (FCoE protocol) and C-NIC drivers.

- The **bnx2fc** Marvell VMware FCoE driver is a kernel mode driver used to provide a translation layer between the VMware SCSI stack and the Marvell FCoE firmware and hardware. In addition, the driver interfaces with the networking layer to transmit and receive encapsulated FCoE frames on behalf of the Open-FCoE libfc and libfcoe for FIP and device discovery.

Supported Distributions

The FCoE and DCB feature set is supported on VMware ESXi 6.0 and later.

Enabling FCoE

To enable FCoE hardware offload on the C-NIC using the legacy bnx2fc driver:

1. Determine the ports that are FCoE-capable by issuing the following command:

   ```
   # esxcli fcoe nic list
   ```

   Output example:

   ```
   vmnic4
   User Priority: 3
   Source MAC: FF:FF:FF:FF:FF:FF
   Active: false
   Priority Settable: false
   Source MAC Settable: false
   VLAN Range Settable: false
   VN2VN Mode Enabled: false
   ```

2. Enable the FCoE interface as follows:

   ```
   # esxcli fcoe nic discover -n vmnicX
   ```

   Where ***X*** is the interface number determined in Step 1**.**

3. Verify that the interface is working as follows:

   ```
   # esxcli fcoe adapter list
   ```

   Output example:

   ```
   vmhba34
   Source MAC: bc:30:5b:01:82:39
   FCF MAC: 00:05:73:cf:2c:ea
   VNPort MAC: 0e:fc:00:47:04:04
   Physical NIC: vmnic7
   User Priority: 3
   VLAN id: 2008
   ```

```
VN2VN Mode Enabled: false
```

The output of this command should show a valid FCoE forwarder (FCF) MAC, VNPort MAC, Priority, and VLAN ID for the fabric that is connected to the C-NIC.

You can also issue the following command to verify that the interface is working properly:

```
# esxcfg-scsidevs -a
```

Output example:

```
vmhba34 bnx2fc link-up fcoe.1000<mac address>:2000<mac address>
vmhba35 bnx2fc link-up fcoe.1000<mac address>:2000<mac address>
```

NOTE

The label `Software FCoE` is a VMware term used to describe initiators that depend on the inbox FCoE libraries and utilities. Marvell's FCoE solution is a fully state, connection-based, hardware offload solution designed to significantly reduce the CPU burden encumbered by a non-offload software initiator.

The native qfle3f driver automatically starts the FCoE initialization and need not follow these steps.

Installation Check

To verify the correct installation of the driver and to ensure that the host port is seen by the switch, follow these steps.

To verify the correct installation of the driver:

1. Verify that the host port shows up in the switch fabric login (FLOGI) database by issuing the one of the following commands:

   `show flogi database` (for a Cisco FCF)

   `fcoe -loginshow` (for a Brocade FCF)

2. If the host WWPN does not appear in the FLOGI database, provide driver log messages for review.

Limitations

FCoE support has the following limitations:

- NPIV is not supported on ESXi by the legacy bnx2fc driver, due to dependencies on supporting (libfc, libfcoe) components and modules. The native qfle3f driver supports NPIV.
- Non-offload FCoE is not supported with offload-capable Marvell devices. Only the full hardware offload path is supported.

iSCSI Support

Marvell provides the bnx2i driver to support iSCSI. The Marvell 57*xx* and 57*xxx* iSCSI driver, bnx2i, is a Marvell VMware iSCSI Host Bus Adapter driver. Similar to bnx2fc, bnx2i is a kernel mode driver used to provide a translation layer between the VMware SCSI stack and the Marvell iSCSI firmware and hardware. The bnx2i functions under the Open-iSCSI framework.

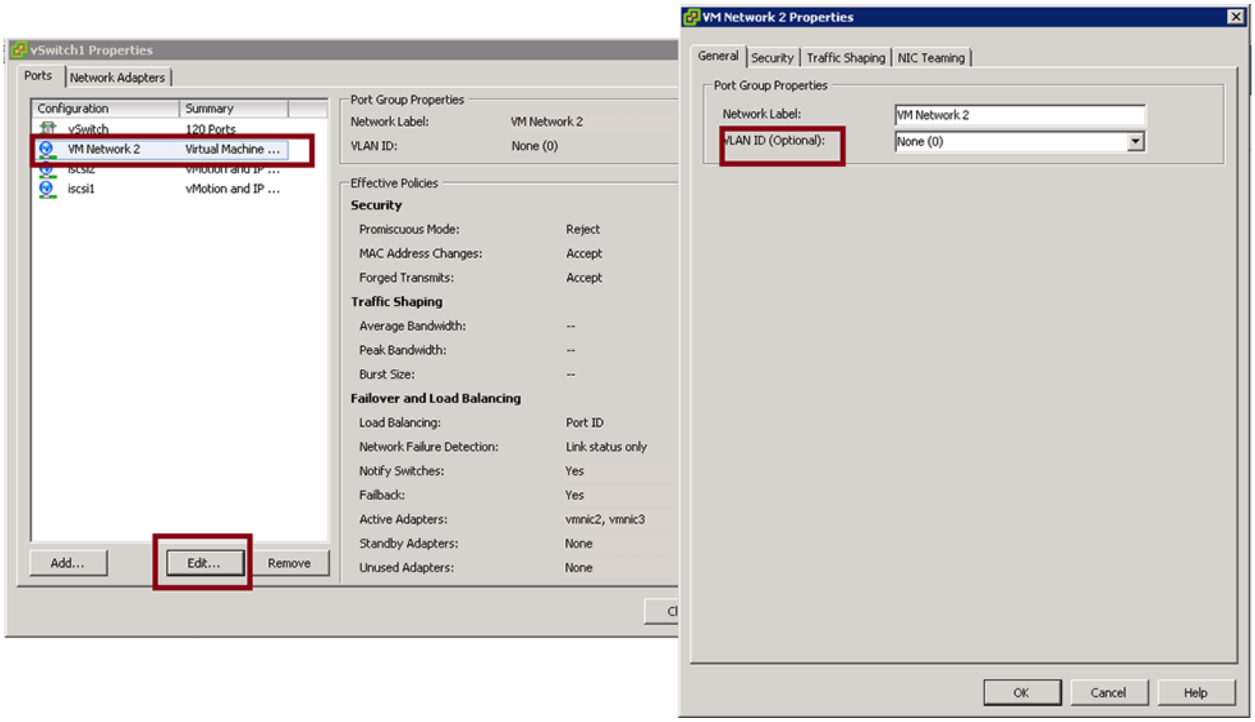
VLAN Configuration for iSCSI Offload (VMware)

iSCSI traffic on the network may be isolated in a VLAN to segregate it from other traffic. When this is the case, you must make the iSCSI interface on the adapter a member of that VLAN.

To configure the VLAN using the V-Sphere client (GUI):

1. Select the ESXi host.
2. Click the **Configuration** tab.
3. On the Configuration page, select the **Networking** link, and then click **Properties**.
4. On the selected vSwitch Properties, Ports page, click the virtual switch or port groups, and then click **Edit**.

5. (Optional) On the VM Network Properties, General page, assign a VLAN number in the **VLAN ID** box. Figure 8-1 and Figure 8-2 show examples.



Configuration	Summary
vSwitch	120 Ports
VM Network 2	Virtual Machine ...
iscsi2	vMotion and IP ...
iscsi1	vMotion and IP ...

Port Group Properties

Network Label: VM Network 2

VLAN ID: None (0)

Effective Policies

Security

Promiscuous Mode: Reject

MAC Address Changes: Accept

Forged Transmits: Accept

Traffic Shaping

Average Bandwidth: --

Peak Bandwidth: --

Burst Size: --

Failover and Load Balancing

Load Balancing: Port ID

Network Failure Detection: Link status only

Notify Switches: Yes

Failback: Yes

Active Adapters: vmnic2, vmnic3

Standby Adapters: None

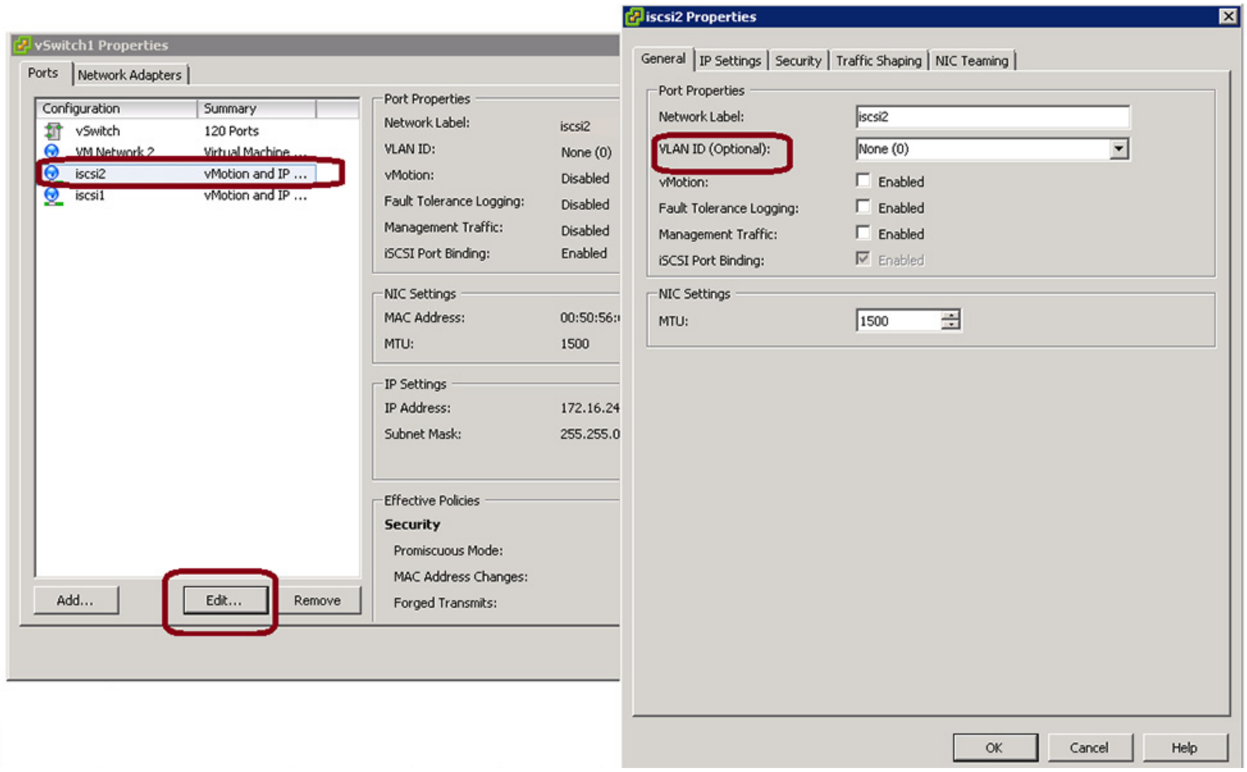
Unused Adapters: None

Port Group Properties

Network Label: VM Network 2

VLAN ID (Optional): None (0)

Figure 8-1. VM Network Properties: Example 1



Configuration	Summary
vSwitch	120 Ports
VM Network 2	Virtual Machine ...
iscsi2	vMotion and IP ...
iscsi1	vMotion and IP ...

Port Properties

Network Label:	iscsi2
VLAN ID:	None (0)
vMotion:	Disabled
Fault Tolerance Logging:	Disabled
Management Traffic:	Disabled
iSCSI Port Binding:	Enabled

NIC Settings

MAC Address:	00:50:56:
MTU:	1500

IP Settings

IP Address:	172.16.24
Subnet Mask:	255.255.0

Effective Policies

Security

Promiscuous Mode:

MAC Address Changes:

Forged Transmits:

Port Properties

Network Label:	iscsi2
VLAN ID (Optional):	None (0)
vMotion:	□ Enabled
Fault Tolerance Logging:	□ Enabled
Management Traffic:	□ Enabled
iSCSI Port Binding:	☑ Enabled

NIC Settings

MTU: 1500

Figure 8-2. VM Network Properties: Example 2

6. Configure the VLAN on VMkernel.

9 Windows Driver Software

Windows driver software information includes the following:

- Supported Drivers
- “Installing the Driver Software” on page 90
- “Modifying the Driver Software” on page 94
- “Repairing or Reinstalling the Driver Software” on page 94
- “Removing the Device Drivers” on page 95
- “Viewing or Changing the Properties of the Adapter” on page 95
- “Setting Power Management Options” on page 95
- “Configuring the Communication Protocol to Use with QCC GUI, QCC PowerKit, and QCS CLI” on page 97

Supported Drivers

The Windows drivers are listed in Table 9-1.

Table 9-1. Marvell 57xx and 57xxx Windows Drivers

Windows Driver	Description
bxVBD	Windows (system device) virtual bus driver (VBD) for the 57*xx* 1Gb network adapters. This driver directly controls the hardware.
eVBD	Windows (system device) VBD for the 57*xxx* 1/10Gb network adapters. This driver directly controls the hardware.
bxND	Windows (NDIS) Ethernet driver for the 57*xx* and 57*xxx* 1/10Gb network adapters.
bxOIS	Windows (storage) iSCSI offload driver for the 57*xx* and 57*xxx* 1/10Gb network adapters.
bxFCoE	Windows (storage) FCoE offload driver for the 57712 and 578*xx* 10Gb network adapters.

Installing the Driver Software

NOTE

These instructions are based on the assumption that your Marvell 57*xx* and 57*xxx* adapters were not factory installed. If your controller was installed at the factory, the driver software has been installed for you.

When Windows first starts after a hardware device (such as a Marvell 57*xxx* adapter) has been installed, or after the existing device driver has been removed, the operating system automatically detects the hardware and prompts you to install the driver software for that device.

The two methods of driver installation include:

- Graphical interactive installation mode (see "Using the Installer" on page 90)
- Command-line silent mode for unattended installation (see "Using Silent Installation" on page 92)

NOTE

- Before installing the driver software, verify that the Windows operating system has been upgraded to the latest version with the latest service pack applied.
- Ensure that a network device driver is physically installed before the Marvell 57*xx* and 57*xxx* controllers are used with your Windows operating system. Drivers are located on the installation CD.
- The TCP/IP offload engine (TOE) is not supported in Windows Server 2016 or later. You must have a license key installed on the motherboard (for LOMs). For add-in NICs, the license key is preprogrammed in the hardware.
- QCC GUI is not supported on the Server Core installation option for Microsoft Windows Server.

Using the Installer

In addition to the Marvell device drivers, the installer installs the management applications. The following are installed when running the installer:

- **QLogic Device Drivers** installs the Marvell device drivers.
- **Control Suite** is the QLogic Control Suite (QCS) CLI.
- **QCC** is the QConverge Console GUI.
- **SNMP** installs the SNMP sub agent.

- **NX RPC Remote Agent** installs the RPC remote agent software.
- **iSCSI Crash Dump Driver** installs the driver needed for the iSCSI Crash Dump utility.
- **FCoE Crash Dump Driver** installs the driver needed for the FCoE Crash Dump utility.
- **FastLinQ HBA Device Mgmt Agent** installs the agent for device management.

To install the Marvell 57*xx* and 57*xxx* drivers and management applications:

1. When the **Found New Hardware Wizard** appears, click **Cancel**.
2. From either the driver source media or the location in which you downloaded the software driver package, do the following:

   a. Open the folder for your operating system.

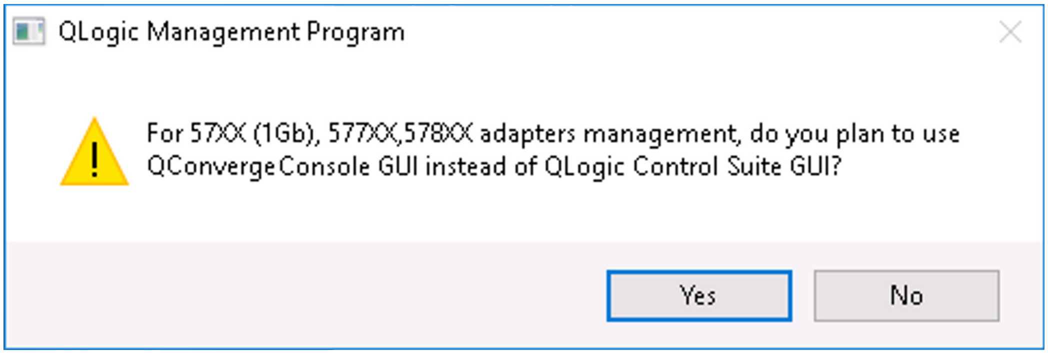
   b. Open the `MUPS` folder, and then extract the folder according to your operating system configuration.

   c. Double-click the **`Setup.exe`** file.

   The InstallShield Wizard for QLogic Drivers and Management Applications opens to the Welcome window.
3. At the InstallShield Wizard prompt (Figure 9-1), select the adapter management utility that you want to use:

   ❑ Click **Yes** to use QConvergeConsole GUI.

   ❑ Click **No** to use QLogic Control Suite.


QLogic Management Program

For 57XX (1Gb), 577XX,578XX adapters management, do you plan to use QConvergeConsole GUI instead of QLogic Control Suite GUI?

Figure 9-1. InstallShield Wizard Prompt for Management Utility

4. At the InstallShield Wizard prompt, “Do you want to skip installing WMI?”, select one of these options:

   ❑ Click **Yes** to defer installation of the Windows Management Instrumentation (WMI) initiative.

- ❑ Click **No** to install WMI.

5. On the InstallShield Welcome window, click **Next** to continue.
6. After you review the license agreement, click **I accept the terms in the license agreement**, and then click **Next** to continue.
7. Select the features you want to install.
8. Click **Install**.
9. Click **Finish** to close the wizard.
10. The installer determines if a system restart is necessary. Follow the on-screen instructions.

To install the Microsoft iSCSI Software Initiator for iSCSI Crash Dump:

If supported and if you will use the Marvell iSCSI Crash Dump utility, it is important to follow the installation sequence:

1. Run the installer.
2. Install Microsoft iSCSI Software Initiator along with the patch (MS KB939875).

NOTE

If you are upgrading the device drivers from the installer, re-enable **iSCSI Crash Dump** from the **Advanced** section of the QCC GUI Configuration page.

Using Silent Installation

NOTE

- All commands are case sensitive.
- For detailed instructions and information about unattended installs, refer to the `silent.txt` file in the `Driver_Management_Apps_Installer` folder.

To perform a silent install from within the installer source folder:

Issue the following command:

```
setup /s /v/qn
```

To perform a silent upgrade from within the installer source folder:

Issue the following command:

```
setup /s /v/qn
```

To perform a silent reinstall of the same installer:

Issue the following command:

```
setup /s /v"/qn REINSTALL=ALL"
```

NOTE

The REINSTALL switch should only be used if the same installer is already installed on the system. If upgrading an earlier version of the installer, use `setup /s /v/qn` as listed in the preceding.

To perform a silent install by feature:

Use the ADDSOURCE to include any of the following features.

Issue the following command according to platform:

IA32 platforms:

```
setup /s /v"/qn ADDSOURCE=Driversi32,BACSi32,BASPi32,SNMPi32,CIMi32"
```

AMD/EM64T platforms:

```
setup /s /v"/qn ADDSOURCE=Driversa64,BACSa64,BASPa64,SNMPa64,CIMa64"
```

The following command-line statement installs only the Marvell drivers according to platform:

IA32 platforms:

```
setup /s /v"/qn ADDSOURCE=Driversi32"
```

AMD64 platforms:

```
setup /s /v"/qn ADDSOURCE=Driversa64"
```

NOTE

The Marvell device drivers are a required feature and are always installed, even if you do not specify `ADDSOURCE`.

To perform a silent install from within a batch file:

To perform a silent install from within a batch file and to wait for the install to complete before continuing with the next command line, issue the following command:

```
start /wait setup /s /w /v/qn
```

Modifying the Driver Software

To modify the driver software:

1. In the Control Panel, double-click **Add or Remove Programs**.
2. Click **QLogic Drivers and Management Applications**, and then click **Change**.
3. Click **Next** to continue.
4. Click **Modify, Add, or Remove** to change program features.

NOTE

This option does not install drivers for new adapters. For information on installing drivers for new adapters, see "Repairing or Reinstalling the Driver Software" on page 94.

5. Click **Next** to continue.
6. Click on an icon to change how a feature is installed.
7. Click **Next**.
8. Click **Install**.
9. Click **Finish** to close the wizard.
10. The installer will determine if a system restart is necessary. Follow the on-screen instructions.

Repairing or Reinstalling the Driver Software

To repair or reinstall the driver software:

1. In Control Panel, double-click **Add or Remove Programs**.
2. Click **QLogic Drivers and Management Applications**, and then click **Change**.
3. Click **Next** to continue.
4. Click **Repair or Reinstall** to repair errors or install drivers for new adapters.
5. Click **Next** to continue.
6. Click **Install**.
7. Click **Finish** to close the wizard.
8. The installer will determine if a system restart is necessary. Follow the on-screen instructions.

Removing the Device Drivers

When removing the device drivers, any management application that is installed is also removed.

To remove the device drivers:

1. In the Control Panel, double-click **Add or Remove Programs**.
2. Click **QLogic Drivers and Management Applications**, and then click **Remove**. Follow the on-screen prompts.
3. Reboot your system to completely remove the drivers. If you fail to reboot your system, you cannot successfully install the drivers.

Viewing or Changing the Properties of the Adapter

To view or change the properties of a Marvell network adapter:

1. In the Control Panel, click **Marvell Control Suite**.
2. Click the **Advanced** section of the Configurations page.

Setting Power Management Options

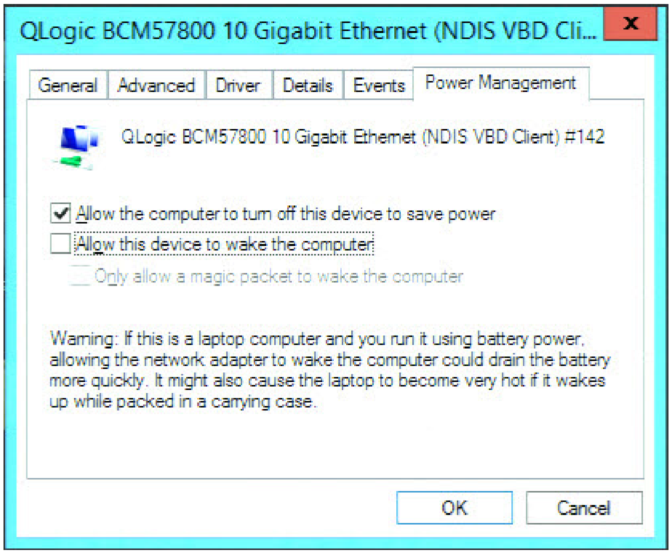
You can set power management options to allow the operating system to turn off the controller to save power or to allow the controller to wake up the computer. If the device is busy doing something (servicing a call, for example) however, the operating system will not shut down the device. The operating system attempts to shut down every possible device only when the computer attempts to go into hibernation.

To have the controller stay on at all times:

On the adapter properties' Power Management page, clear the **Allow the computer to turn off the device to save power** check box, as shown in Figure 9-2.

NOTE

Power management options are not available on blade servers.



QLogic BCM57800 10 Gigabit Ethernet (NDIS VBD Client) #142

☑ Allow the computer to turn off this device to save power

☐ Allow this device to wake the computer

☐ Only allow a magic packet to wake the computer

Warning: If this is a laptop computer and you run it using battery power, allowing the network adapter to wake the computer could drain the battery more quickly. It might also cause the laptop to become very hot if it wakes up while packed in a carrying case.

Figure 9-2. Device Power Management Options

NOTE

- The Power Management page is available only for servers that support power management.
- To enable wake on LAN (WoL) when the computer is on standby, select the **Allow the device to wake the computer** check box.
- If you select the **Only allow a magic packet to wake the computer** check box, the computer can be brought out of standby *only* by a magic packet.

CAUTION

Do not select **Allow the computer to turn off the device to save power** for any adapter that is a member of a team.

Configuring the Communication Protocol to Use with QCC GUI, QCC PowerKit, and QCS CLI

There are two main components of the QCC GUI, QCC PowerKit, and QCS CLI management applications: the RPC agent and the client software. An RPC agent is installed on a server, or managed host, that contains one or more Converged Network Adapters. The RPC agent collects information on the Converged Network Adapters and makes it available for retrieval from a management PC on which the client software is installed. The client software enables viewing information from the RPC agent and configuring the Converged Network Adapters.The management software includes QCC GUI and QCS CLI.

A communication protocol enables communication between the RPC agent and the client software. Depending on the mix of operating systems (Linux, Windows, or both) on the clients and managed hosts in your network, you can choose an appropriate utility.

For installation instructions for these management applications, refer to the following documents:

- *User’s Guide, QLogic Control Suite CLI* (part number BC0054511-00)
- *User’s Guide, PowerShell* (part number BC0054518-00)
- *Installation Guide, QConvergeConsole GUI* (part number SN0051105-00)

To locate these documents, see “Laser Safety Information” on page xxiii.

10 Citrix XenServer Driver Software

This chapter describes how to install the Citrix driver on a XenServer operating system using the driver update disk (DUD).

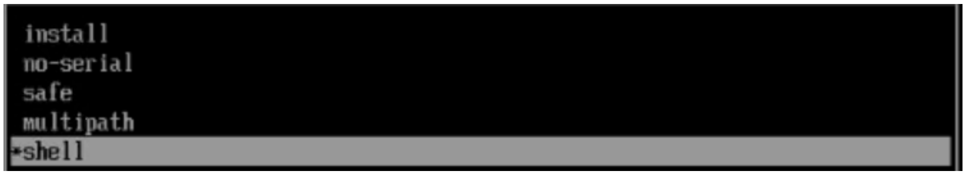
NOTE

The procedures in this section apply only to Citrix XenServer 8.0 and later distributions.

These procedures use both the DUD and the OS installation disk.

To install the Citrix hypervisor driver:

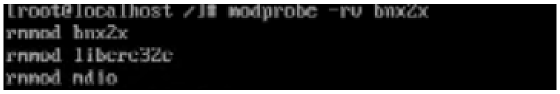
1. Insert the XenServer installation CD and begin the installation in shell mode (see Figure 10-1).



```
install
no-serial
safe
multipath
*shell
```

Figure 10-1. Starting in Shell Mode

2. When the system boots to shell mode, unload (should be upload?) the inbox bnx2x driver (see Figure 10-2).

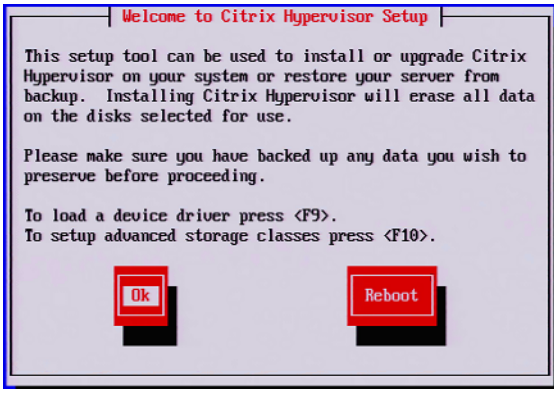


```
[root@localhost /]# modprobe -rv bnx2x
rmmod bnx2x
rmmod libcrc32c
rmmod mdio
```

Figure 10-2. Installing the bnx2x Driver

3. Type **exit**, and then press ENTER, to return to the GUI installer.

4. Insert the DUD CD/ISO. The GUI Welcome screen appears (see Figure 10-3).



Welcome to Citrix Hypervisor Setup

This setup tool can be used to install or upgrade Citrix Hypervisor on your system or restore your server from backup. Installing Citrix Hypervisor will erase all data on the disks selected for use.

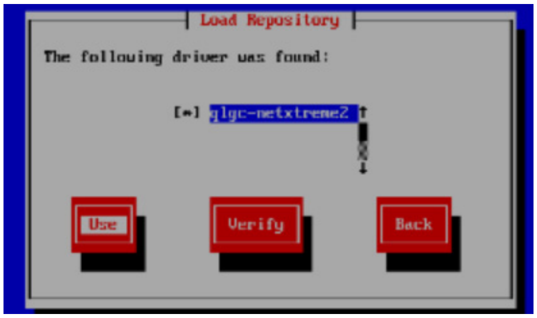
Please make sure you have backed up any data you wish to preserve before proceeding.

To load a device driver press <F9>.
To setup advanced storage classes press <F10>.

Figure 10-3. Loading the Device Driver

Press F9 to load the driver.

The Load Repository window appears (see Figure 10-4.)



Load Repository

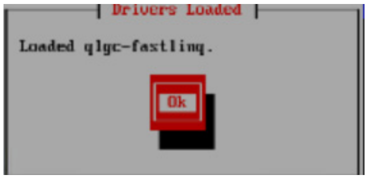
The following driver was found:

[*] qlgc-netxtreme2

Figure 10-4. Locating the Device Driver

5. Click **Use**.

The Drivers Loaded window appears (see Figure 10-5).

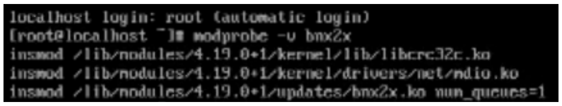


Drivers Loaded

Loaded qlgc-fastlinq.

Figure 10-5. Driver Installed Successfully

6. Press ALT+F2 to return to shell mode, and then load the out-of-box (OOB) driver (see Figure 10-6).



```
localhost login: root (automatic login)
[root@localhost ~]# modprobe -v bnx2x
insmod /lib/modules/4.19.0+1/kernel/lib/libcrc32c.ko
insmod /lib/modules/4.19.0+1/kernel/drivers/net/mdio.ko
insmod /lib/modules/4.19.0+1/updates/bnx2x.ko num_queues=1
```

Figure 10-6. Loading the OOB Driver

7. Press ALT+F1 to return to the GUI installer, and then continue the installation.

   Do **not** remove the driver CD/ISO.

8. When prompted, skip the supplemental package installation.

9. When prompted, reboot the system after removing the OS installer CD and the DUD.

   The hypervisor should boot with the new driver installed.

11 iSCSI Protocol

This chapter provides the following information about the iSCSI protocol:

- iSCSI Boot
- "iSCSI Crash Dump" on page 128
- "iSCSI Offload in Windows Server" on page 128

iSCSI Boot

Marvell 57*xx* and 57*xxx* gigabit Ethernet (GbE) adapters support iSCSI boot to enable network boot of operating systems to diskless systems. iSCSI boot allows a Windows, Linux, or VMware operating system boot from an iSCSI target machine located remotely over a standard IP network.

For both Windows and Linux operating systems, iSCSI boot can be configured to boot with two distinctive paths: non-offload (also known as Microsoft/Open-iSCSI initiator) and offload (Marvell's offload iSCSI driver or Host Bus Adapter). Configure the path with the **HBA Boot Mode** option located on the General Parameters page of the iSCSI Configuration utility. See Table 11-1 on page 105 for more information on all General Parameters page configuration options.

NOTE

If you are using iSCSI boot on 57*xxx* based designs, SR-IOV must be disabled in the system before upgrading from a 7.2.*x* (or earlier) release to release 7.4.*x* or later.

Supported Operating Systems for iSCSI Boot

The Marvell 57*xx* and 57*xxx* gigabit Ethernet adapters support iSCSI boot on the following operating systems:

- Windows Server 2012 and later 32-bit and 64-bit (supports offload and non-offload paths)
- Linux RHEL 6 and later and SLES 11.1 and later (supports offload and non-offload paths)
- VMware ESX in Layer 2 path

In addition, the adapters support iSCSI boot for unspecified path types on supported Windows (see "Microsoft Windows" on page 18), RHEL (see "Linux" on page 18), and Linux ("Linux" on page 18) OSs.

Jumbo frames with iSCSI boot are supported only on Windows OSs, when the adapter is used as either an NDIS or HBA offload device.

iSCSI Boot Setup

The iSCSI boot setup includes:

- Configuring the iSCSI Target
- Configuring iSCSI Boot Parameters
- Preparing the iSCSI Boot Image
- Booting

Configuring the iSCSI Target

Configuring the iSCSI target varies by target vendors. For information on configuring the iSCSI target, refer to the documentation provided by the vendor. The general steps include:

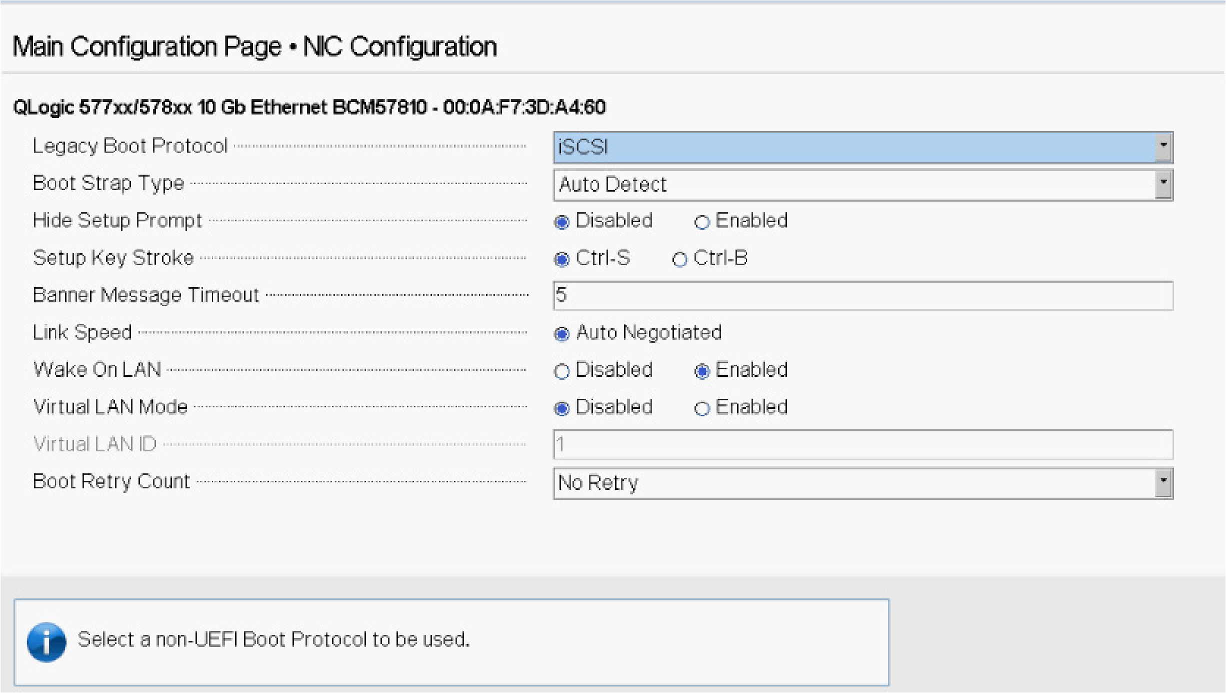
1. Create an iSCSI target.
2. Create a virtual disk.
3. Map the virtual disk to the iSCSI target created in Step 1.
4. Associate an iSCSI initiator with the iSCSI target.
5. Record the iSCSI target name, TCP port number, iSCSI LUN, initiator Internet Qualified Name (IQN), and CHAP authentication details.
6. After configuring the iSCSI target, obtain the following:
   - ❑ Target IQN
   - ❑ Target IP address
   - ❑ Target TCP port number
   - ❑ Target LUN

- ❑ Initiator IQN
- ❑ CHAP ID and secret

Configuring iSCSI Boot Parameters

To configure the iSCSI boot parameters:

1. In the NIC Configuration page, in the **Legacy Boot Protocol** drop-down menu, select **iSCSI** (see Figure 11-1).



Main Configuration Page • NIC Configuration

QLogic 577xx/578xx 10 Gb Ethernet BCM57810 - 00:0A:F7:3D:A4:60

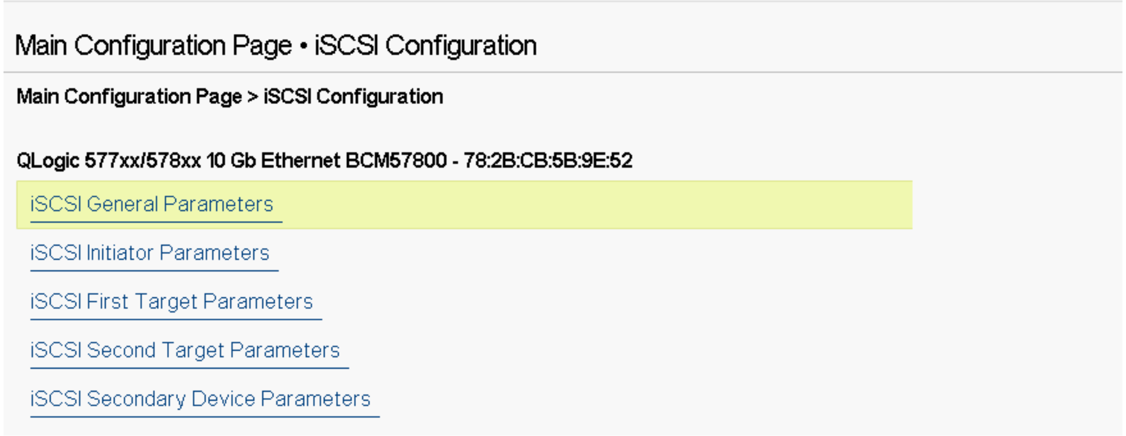
Legacy Boot Protocol	iSCSI
Boot Strap Type	Auto Detect
Hide Setup Prompt	● Disabled ○ Enabled
Setup Key Stroke	● Ctrl-S ○ Ctrl-B
Banner Message Timeout	5
Link Speed	● Auto Negotiated
Wake On LAN	○ Disabled ● Enabled
Virtual LAN Mode	● Disabled ○ Enabled
Virtual LAN ID	1
Boot Retry Count	No Retry

Select a non-UEFI Boot Protocol to be used.

Figure 11-1. Legacy Boot Protocol Selection

As shown in Figure 11-1, UEFI is not supported for the iSCSI protocol for the 57*xx* and 57*xxx* adapters.

2. Configure the iSCSI boot software for either static or dynamic configuration in the CCM, UEFI (see Figure 11-2), QCC GUI, or QCS CLI.



Main Configuration Page • iSCSI Configuration

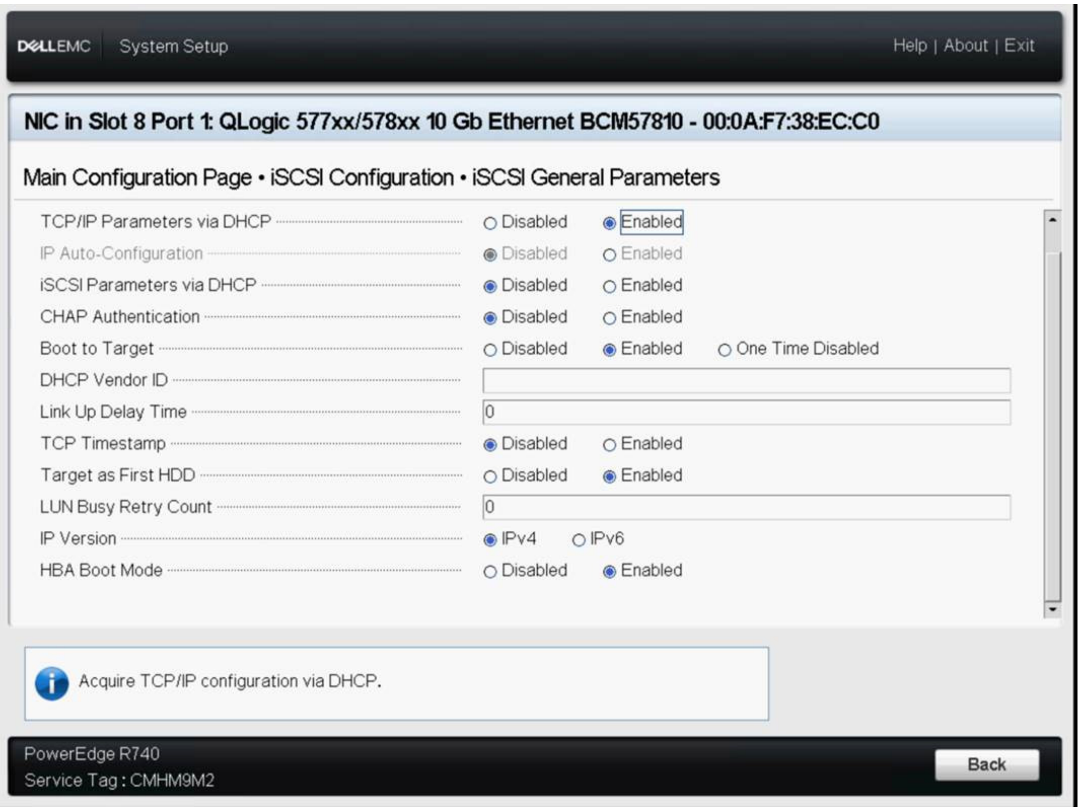
Main Configuration Page > iSCSI Configuration

QLogic 577xx/578xx 10 Gb Ethernet BCM57800 - 78:2B:CB:5B:9E:52

iSCSI General Parameters
iSCSI Initiator Parameters
iSCSI First Target Parameters
iSCSI Second Target Parameters
iSCSI Secondary Device Parameters

Figure 11-2. UEFI, iSCSI Configuration

The configuration options available on the General Parameters window (see Figure 11-3) are listed in Table 11-1.



NIC in Slot 8 Port 1: QLogic 577xx/578xx 10 Gb Ethernet BCM57810 - 00:0A:F7:38:EC:C0

Main Configuration Page • iSCSI Configuration • iSCSI General Parameters

TCP/IP Parameters via DHCP	○ Disabled	● Enabled	
IP Auto-Configuration	● Disabled	○ Enabled	
iSCSI Parameters via DHCP	● Disabled	○ Enabled	
CHAP Authentication	● Disabled	○ Enabled	
Boot to Target	○ Disabled	● Enabled	○ One Time Disabled
DHCP Vendor ID			
Link Up Delay Time	0		
TCP Timestamp	● Disabled	○ Enabled	
Target as First HDD	○ Disabled	● Enabled	
LUN Busy Retry Count	0		
IP Version	● IPv4	○ IPv6	
HBA Boot Mode	○ Disabled	● Enabled	

Acquire TCP/IP configuration via DHCP.

PowerEdge R740
Service Tag : CMHM9M2

Figure 11-3. UEFI, iSCSI Configuration, iSCSI General Parameters

Table 11-1 lists parameters for both IPv4 and IPv6. Parameters specific to either IPv4 or IPv6 are noted.

NOTE

Availability of IPv6 iSCSI boot is platform and device dependent.

Table 11-1. Configuration Options

Option	Description
TCP/IP parameters through DHCP	This option is specific to IPv4. Controls whether the iSCSI boot host software acquires the IP address information using DHCP (Enabled) or use a static IP configuration (Disabled).

Table 11-1. Configuration Options (Continued)

Option	Description
IP Autoconfiguration	This option is specific to IPv6. Controls whether the iSCSI boot host software will configure a stateless link-local address and/or stateful address if DHCPv6 is present and used (Enabled). Router Solicit packets are sent out up to three times with 4 second intervals in between each retry. Or use a static IP configuration (Disabled).
iSCSI Parameters via DHCP	Controls whether the iSCSI boot host software acquires its iSCSI target parameters using DHCP (Enabled) or through a static configuration (Disabled). The static information is entered through the iSCSI Initiator Parameters Configuration window.
CHAP Authentication	Controls whether the iSCSI boot host software uses CHAP authentication when connecting to the iSCSI target. If CHAP Authentication is enabled, the CHAP ID and CHAP Secret are entered through the iSCSI Initiator Parameters Configuration window.
Boot to Target	The first time a connection is made, this option controls whether the designated iSCSI LUN will either be: ■ Not be booted from (Disabled) ■ Always be booted from (Enabled) ■ Not booted After a reboot, this option must be set to Enabled and always booted from (One Time Disabled). This control allows the boot OS to be installed on the connected LUN, which must be done the first time the system is setup. Afterwards, this system can connect and boot from that installed OS image once this control is set to Enabled.
DHCP Vendor ID	Controls how the iSCSI boot host software interprets the Vendor Class ID field used during DHCP. If the Vendor Class ID field in the DHCP Offer packet matches the value in the field, the iSCSI boot host software looks into the DHCP Option 43 fields for the required iSCSI boot extensions. If DHCP is disabled, this value does not need to be set.
Link Up Delay Time	Controls how long the iSCSI boot host software waits, in seconds, after an Ethernet link is established before sending any data over the network. The valid values are 0 to 255. As an example, a user may need to set a value for this option if a network protocol, such as Spanning Tree, is enabled on the switch interface to the client system.
TCP Timestamp	Controls if the TCP Timestamp option is enabled or disabled.
Target as First HDD	Allows specifying that the iSCSI target drive will appear as the first hard drive in the system.

Table 11-1. Configuration Options (Continued)

Option	Description
LUN Busy Retry Count	Controls the quantity of connection retries the iSCSI Boot initiator will attempt if the iSCSI target LUN is busy.
IP Version	This option is specific to IPv6. Toggles between the IPv4 or IPv6 protocol. All IP settings will be lost when switching from one protocol version to another.
HBA Boot Mode	Set to **disable** when the host OS is configured for software initiator mode and to **enable** for HBA (or iSCSI offload) initiator mode. This option is available on 57*xx* and 57*xxx* adapters. (**Note**: This parameter cannot be changed when the adapter is in Multi-Function mode.)

MBA Boot Protocol Configuration

To configure the boot protocol, see Chapter 6 Boot Agent Driver Software.

iSCSI Boot Configuration

- Static iSCSI Boot Configuration
- Dynamic iSCSI Boot Configuration

Static iSCSI Boot Configuration

In a static configuration, you must enter data for the system’s IP address, the system’s initiator IQN, and the target parameters obtained in “Configuring the iSCSI Target” on page 102. For information on configuration options, see Table 11-1 on page 105.

To configure the iSCSI boot parameters using static configuration:

1. On the General Parameters Menu page, set the following:
   - ❑ **TCP/IP parameters via DHCP**: Disabled (for IPv4)
   - ❑ **IP Autoconfiguration**: Disabled (for IPv6, non-offload)
   - ❑ **iSCSI parameters via DHCP**: Disabled
   - ❑ **CHAP Authentication**: Disabled
   - ❑ **Boot from Target**: see NOTE
   - ❑ **DHCP Vendor ID**: QLGC ISAN
   - ❑ **Link Up Delay Time**: 0
   - ❑ **Use TCP Timestamp**: Enabled (for some targets, such as the Dell or EMC AX100i, it is necessary to enable **Use TCP Timestamp**)
   - ❑ **Target as First HDD**: Enabled

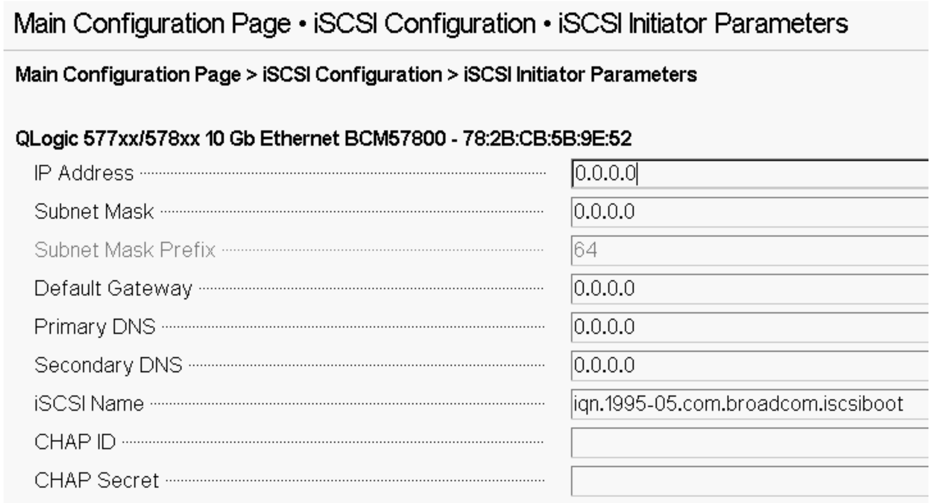
- **LUN Busy Retry Count**: 0
- **IP Version**: IPv6 (for IPv6, non-offload)
- **HBA Boot Mode**: Disabled

NOTE

For initial OS installation to a blank iSCSI target LUN from a CD/DVD-ROM or mounted bootable OS installation image, set **Boot from Target** to **One Time Disabled**. This setting causes the system not to boot from the configured iSCSI target after establishing a successful login and connection. This setting will revert to **Enabled** after the next system reboot. **Enabled** means to connect to an iSCSI target and attempt to boot from it. **Disabled** means to connect to an iSCSI target and not boot from that device, but instead hand off the boot vector to the next bootable device in the boot sequence.

2. Press the ESC key to return to the **Main** menu.
3. On the **Main** menu, select **iSCSI Initiator Parameters**.

   The iSCSI Initiator Parameters window appears (see Figure 11-4).



Main Configuration Page • iSCSI Configuration • iSCSI Initiator Parameters

Main Configuration Page > iSCSI Configuration > iSCSI Initiator Parameters

QLogic 577xx/578xx 10 Gb Ethernet BCM57800 - 78:2B:CB:5B:9E:52

IP Address	0.0.0.0
Subnet Mask	0.0.0.0
Subnet Mask Prefix	64
Default Gateway	0.0.0.0
Primary DNS	0.0.0.0
Secondary DNS	0.0.0.0
iSCSI Name	iqn.1995-05.com.broadcom.iscsiboot
CHAP ID	
CHAP Secret	

Figure 11-4. iSCSI Configuration, iSCSI Initiator Parameters

4. On the iSCSI Initiator Parameters window (Figure 11-4), type values for the following:

   - ❑ IP Address (unspecified IPv4 and IPv6 addresses should be `0.0.0.0` and `::`, respectively)

     NOTE

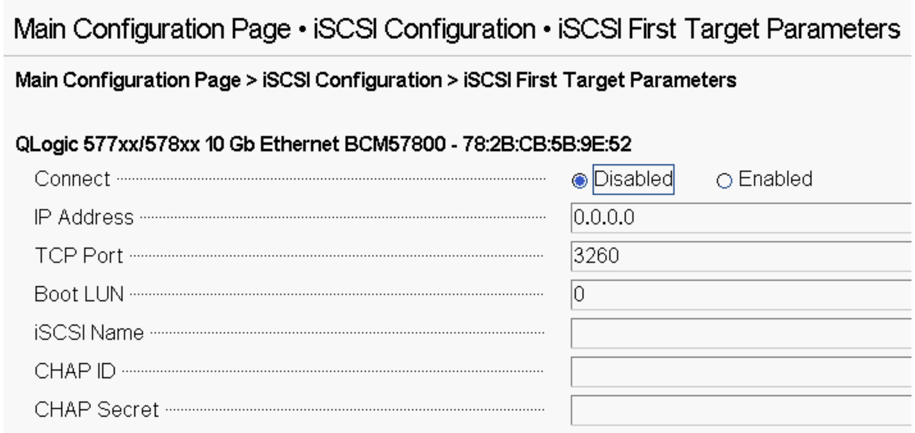
     Carefully enter the IP address. There is no error-checking performed against the IP address to check for duplicates or incorrect segment or network assignment.

   - ❑ Subnet Mask
   - ❑ Subnet Mask Prefix
   - ❑ Default Gateway
   - ❑ Primary DNS
   - ❑ Secondary DNS
   - ❑ iSCSI Name (corresponds to the iSCSI initiator name to be used by the client system)
   - ❑ CHAP ID
   - ❑ CHAP Secret

5. Press the ESC key to return to the **Main** menu.

6. On the **Main** menu, select **iSCSI First Target Parameters**.

   The iSCSI First Target Parameters window appears (see Figure 11-5).



Main Configuration Page • iSCSI Configuration • iSCSI First Target Parameters

Main Configuration Page > iSCSI Configuration > iSCSI First Target Parameters

QLogic 577xx/578xx 10 Gb Ethernet BCM57800 - 78:2B:CB:5B:9E:52

Connect	● Disabled ○ Enabled
IP Address	0.0.0.0
TCP Port	3260
Boot LUN	0
iSCSI Name	
CHAP ID	
CHAP Secret	

Figure 11-5. iSCSI Configuration, iSCSI First Target Parameters

7. On the iSCSI First Target Parameters window (Figure 11-5):
   a. Enable **Connect** to connect to the iSCSI target.
   b. Type values for the following using the values used when configuring the iSCSI target:
      - IP Address
      - TCP Port
      - Boot LUN
      - iSCSI Name
      - CHAP ID
      - CHAP Secret
8. Press ESC to return to the **Main** menu.
9. (Optional) Configure a secondary iSCSI target by repeating these steps in the iSCSI Second Target Parameter window.
10. Press ESC and select **Exit and Save Configuration**.
11. Press the F4 key to save your MBA configuration.

Dynamic iSCSI Boot Configuration

In a dynamic configuration, you only need to specify that the system's IP address and target/initiator information are provided by a DHCP server (see IPv4 and IPv6 configurations in "Configuring the DHCP Server to Support iSCSI Boot" on page 112). For IPv4, with the exception of the initiator iSCSI name, any settings on the Initiator Parameters, 1st Target Parameters, or 2nd Target Parameters windows are ignored and do not need to be cleared. For IPv6, with the exception of the CHAP ID and Secret, any settings on the Initiator Parameters, 1st Target Parameters, or 2nd Target Parameters windows are ignored and do not need to be cleared. For information on configuration options, see Table 11-1 on page 105.

NOTE

When using a DHCP server, the DNS server entries are overwritten by the values provided by the DHCP server. This overwrite occurs even if the locally provided values are valid and the DHCP server provides no DNS server information. When the DHCP server provides no DNS server information, both the primary and secondary DNS server values are set to `0.0.0.0`. When the Windows OS takes over, the Microsoft iSCSI initiator retrieves the iSCSI Initiator parameters and configures the appropriate registries statically. It will overwrite whatever is configured. Because the DHCP daemon runs in the Windows environment as a user process, all TCP/IP parameters must be statically configured before the stack comes up in the iSCSI Boot environment.

If DHCP Option 17 is used, the target information is provided by the DHCP server, and the initiator iSCSI name is retrieved from the value programmed on the Initiator Parameters window. If no value was selected, the controller defaults to the following name:

```
iqn.1995-05.com.qlogic.<11.22.33.44.55.66>.iscsiboot
```

Where the string `11.22.33.44.55.66` corresponds to the controller's MAC address.

If DHCP option 43 (IPv4 only) is used, any settings on the Initiator Parameters, 1st Target Parameters, or 2nd Target Parameters windows are ignored and do not need to be cleared.

To configure the iSCSI boot parameters using dynamic configuration:

1. On the General Parameters Menu window, set the following:
   - ❑ **TCP/IP parameters via DHCP**: Enabled. (For IPv4.)
   - ❑ **IP Autoconfiguration**: Enabled (For IPv6, non-offload)
   - ❑ **iSCSI parameters via DHCP**: Enabled
   - ❑ **CHAP Authentication**: Disabled
   - ❑ **Boot from Target**: see NOTE
   - ❑ **DHCP Vendor ID**: QLGC ISAN
   - ❑ **Link Up Delay Time**: 0
   - ❑ **Use TCP Timestamp**: Enabled (for some targets such as the Dell or EMC AX100i, it is necessary to enable **Use TCP Timestamp**)
   - ❑ **Target as First HDD**: Disabled
   - ❑ **LUN Busy Retry Count**: 0
   - ❑ **IP Version**: IPv6. (For IPv6, non-offload)
   - ❑ **HBA Boot Mode**: Disabled. (**Note**: This parameter cannot be changed when the adapter is in Multi-Function mode.)
2. Press the ESC key to return to the **Main** menu.

NOTE

Information on the Initiator Parameters, and 1st Target Parameters windows are ignored and do not need to be cleared.

3. Select **Exit and Save Configurations**.

Enabling CHAP Authentication

Ensure that CHAP authentication is enabled on the target and initiator.

To enable CHAP authentication:

1. On the iSCSI General Parameters window, set **CHAP Authentication** to **Enabled**.
2. On the iSCSI Initiator Parameters window, type values for the following:
   - ❑ CHAP ID (up to 128 bytes)
   - ❑ CHAP Secret (if authentication is required, and must be a minimum of 12 characters; the maximum length is 16 characters)
3. Press the ESC key to return to the **Main** menu.
4. On the iSCSI First Target Parameters window, type values for the following using the values used when configuring the iSCSI target:
   - ❑ CHAP ID (optional if two-way CHAP)
   - ❑ CHAP Secret (optional if two-way CHAP, and must be a minimum of 12 characters; the maximum length is 16 characters.)
5. Press ESC to return to the **Main** menu.
6. (optional) Add CHAP to the iSCSI Second Target Parameters menu.
7. Press ESC and select **Exit and Save Configuration**.

Configuring the DHCP Server to Support iSCSI Boot

The DHCP server is an optional component and it is only necessary if you will be doing a dynamic iSCSI Boot configuration setup (see “Dynamic iSCSI Boot Configuration” on page 110).

Configuring the DHCP server to support iSCSI boot is different for IPv4 and IPv6.

- DHCP iSCSI Boot Configuration for IPv4
- DHCP iSCSI Boot Configuration for IPv6

DHCP iSCSI Boot Configuration for IPv4

The DHCP protocol includes a quantity of options that provide configuration information to the DHCP client. For iSCSI boot, Marvell adapters support the following DHCP configurations:

- DHCP Option 17, Root Path
- DHCP Option 43, Vendor-Specific Information

DHCP Option 17, Root Path

Option 17 is used to pass the iSCSI target information to the iSCSI client.

The format of the root path as defined in IETC RFC 4173 is:

```
"iscsi:"<servername>":"<protocol>":"<port>":"<LUN>":"<targetname>"
```

Table 11-2 lists the parameters and definitions.

Table 11-2. DHCP Option 17 Parameter Definition

Parameter	Definition
`"iscsi:"`	A literal string
`<servername>`	The IP address or FQDN of the iSCSI target
`":"`	Separator
`<protocol>`	IP protocol used to access the iSCSI target. Currently, only TCP is supported so the protocol is 6.
`<port>`	Port number associated with the protocol. The standard port number for iSCSI is 3260.
`<LUN>`	LUN to use on the iSCSI target. The value of the LUN must be represented in hexadecimal format. Configure a LUN with an ID of 64 as 40 within the option 17 parameter on the DHCP server.
`<targetname>`	Target name in either IQN or extended unique identifier (EUI) format (refer to RFC 3720 for details on both IQN and EUI formats). An example IQN name is `iqn.1995-05.com.Marvell:iscsi-target`.

DHCP Option 43, Vendor-Specific Information

DHCP option 43 (vendor-specific information) provides more configuration options to the iSCSI client than DHCP option 17. In this configuration, three additional suboptions are provided that assign the initiator IQN to the iSCSI boot client along with two iSCSI target IQNs that can be used for booting. The format for the iSCSI target IQN is the same as that of DHCP option 17, while the iSCSI initiator IQN is simply the initiator's IQN.

NOTE

DHCP Option 43 is supported on IPv4 only.

Table 11-3 lists the suboption.

Table 11-3. DHCP Option 43 Suboption Definition

Suboption	Definition
201	First iSCSI target information in the standard root path format `"iscsi:"<servername>":"<protocol>":"<port>":"<LUN>":"<targetname>"`

Using DHCP option 43 requires more configuration than DHCP option 17, but it provides a richer environment and provides more configuration options. Marvell recommends that customers use DHCP option 43 when performing dynamic iSCSI boot configuration.

Configuring the DHCP Server

Configure the DHCP server to support option 17 or option 43.

NOTE

If using Option 43, you also need to configure Option 60. The value of Option 60 should match the **DHCP Vendor ID** value. The **DHCP Vendor ID** value is QLGC ISAN, as shown in the **General Parameters** section of the **iSCSI Boot Configuration** menu.

DHCP iSCSI Boot Configuration for IPv6

The DHCPv6 server can provide a quantity of options, including stateless or stateful IP configuration, as well s information to the DHCPv6 client. For iSCSI boot, Marvell adapters support the following DHCP configurations:

- DHCPv6 Option 16, Vendor Class Option
- DHCPv6 Option 17, Vendor-Specific Information

NOTE

The DHCPv6 standard Root Path option is not yet available. Marvell suggests using Option 16 or Option 17 for dynamic iSCSI Boot IPv6 support.

DHCPv6 Option 16, Vendor Class Option

DHCPv6 Option 16 (vendor class option) must be present and must contain a string that matches your configured **DHCP Vendor ID** parameter. The **DHCP Vendor ID** value is **`QLGC ISAN`**, as shown in **General Parameters** of the iSCSI **Boot Configuration** menu.

The content of Option 16 should be `<2-byte length> <DHCP Vendor ID>`.

DHCPv6 Option 17, Vendor-Specific Information

DHCPv6 Option 17 (vendor-specific information) provides more configuration options to the iSCSI client. In this configuration, three additional suboptions are provided that assign the initiator IQN to the iSCSI boot client along with two iSCSI target IQNs that can be used for booting.

Table 11-4 lists the suboption.

Table 11-4. DHCP Option 17 Suboption Definition

Suboption	Definition
201	First iSCSI target information in the standard root path format `"iscsi:"[<servername>]":"<protocol>":"<port>":"<LUN>":"<targetname>"`

NOTE

In Table 11-4, the brackets [ ] are required for the IPv6 addresses.

The content of option 17 should be `<2-byte Option Number 201|202|203> <2-byte length> <data>`.

Configuring the DHCP Server

Configure the DHCP server to support Option 16 and Option 17.

NOTE

The format of DHCPv6 Option 16 and Option 17 are fully defined in RFC 3315.

Preparing the iSCSI Boot Image

- Windows Server 2016/2019/Azure Stack HCI iSCSI Boot Setup
- Linux iSCSI Boot Setup
- SUSE 11.1 Remote DVD Installation Workaround
- Removing Inbox Drivers from Windows OS Image
- Injecting (Slipstreaming) Marvell Drivers into Windows Image Files

Windows Server 2016/2019/Azure Stack HCI iSCSI Boot Setup

Windows Server 2016/2019/Azure Stack HCI support booting as well as installing in either the offload or non-offload paths. Marvell requires the use of a “slipstream” DVD with the latest Marvell drivers injected (see “Injecting (Slipstreaming) Marvell Drivers into Windows Image Files” on page 122). Also refer to the Microsoft knowledge base topic KB974072 at support.microsoft.com.

NOTE

The Microsoft procedure injects only the EVBD and NDIS drivers. Marvell recommends that all drivers (EVBD, VBD, BXND, OIS, FCoE, and NDIS) be injected.

To prepare the image for installation and booting in either the offload or non-offload path:

1. Remove any local hard drives on the system to be booted (the “remote system”).
2. Load the latest Marvell MBA and iSCSI boot images into the NVRAM of the adapter.
3. Configure the BIOS on the remote system to have the Marvell MBA as the first bootable device and the CDROM as the second device.
4. Configure the iSCSI target to allow a connection from the remote device. Ensure that the target has sufficient disk space to hold the new O/S installation.
5. Boot up the remote system. When the Preboot Execution Environment (PXE) banner appears, press the CTRL+S keys to enter the PXE menu.
6. At the PXE menu, set **Boot Protocol** to **iSCSI**.
7. Enter the iSCSI target parameters.
8. Set **HBA Boot Mode** to **Enabled** or **Disabled**. (Note: This parameter cannot be changed when the adapter is in Multi-Function mode.)
9. Save the settings and reboot the system.

   The remote system should connect to the iSCSI target and then boot from the DVDROM device.
10. Boot from DVD and begin installation.
11. Answer all the installation questions appropriately (specify the operating system you want to install, accept the license terms, and so on).

    When the **Where do you want to install Windows?** window appears, the target drive should be visible. The target drive is connected through the iSCSI boot protocol, located in the remote iSCSI target.

12. Select **Next** to proceed with Windows Server installation.

    A few minutes after the Windows Server DVD installation process starts, a system reboot occurs. After the reboot, the Windows Server installation routine should resume and complete the installation.

13. Following another system restart, verify that the remote system is able to boot to the desktop.

14. After Windows Server boots to the OS, Marvell recommends running the driver installer to complete the Marvell drivers and application installation.

Linux iSCSI Boot Setup

Linux iSCSI boot is supported on Red Hat Enterprise Linux 5.5 and later and SUSE Linux Enterprise Server 11 (SLES 11) SP1 and later in both the offload and non-offload paths.

To set up Linux iSCSI boot:

1. For driver update, obtain the latest QLogic Linux driver CD.

2. Configure the iSCSI Boot Parameters for DVD direct install to target by disabling the boot-from-target option on the network adapter.

3. Configure to install through the non-offload path by setting **HBA Boot Mode** to **Disabled** in the NVRAM Configuration. (Note: This parameter cannot be changed when the adapter is in Multi-Function mode.).

4. Change the boot order as follows:

    a. Boot from the network adapter.

    b. Boot from the CD or DVD driver.

5. Reboot the system.

    The system will connect to iSCSI target, and then will boot from the CD or DVD drive.

6. For SUSE 11.*x*, choose **installation** and type with **iscsi=1 netsetup=1** at the boot option. If driver update is required, choose **YES** for the F6 driver option.

7. At the `networking device` prompt, choose the required network adapter port, and then click **OK**.

8. At the `configure TCP/IP` prompt, configure the way the system acquire IP address, and then click **OK**.

9. If static IP was chosen, you must enter IP information for iSCSI initiator.

10. (RHEL) Choose to “skip” media testing.

11. Continue installation as needed. A drive will be available at this point. After file copying is done, remove the CD or DVD and reboot the system.

12. When the system reboots, enable “boot from target” in iSCSI Boot Parameters and continue with installation until it is done.

At this stage, the initial installation phase is complete.

To create a new customized initrd for any new components update:

1. Update the iSCSI initiator if needed. You must first remove the existing initiator using `rpm -e`.
2. Make sure all runlevels of network service are on:

   ```
   chkconfig network on
   ```

3. Make sure 2, 3, and 5 runlevels of iSCSI service are on:

   ```
   chkconfig -level 235 iscsi on
   ```

4. For Red Hat 6.0, make sure Network Manager service is stopped and disabled.
5. (Optional) Install iscsiuio (not required for SUSE 10).
6. (Optional) Install the linux-nx2 package.
7. Install the bibt package.
8. Remove `ifcfg-eth*`.
9. Reboot.
10. For SUSE 11.1, follow the remote DVD installation workaround shown in the next section.
11. After the system reboots, log in, change to the `/opt/bcm/bibt` folder, and run the `iscsi_setup.sh` script to create the offload and the non-offload initrd image.
12. Copy the initrd image or images, offload and non-offload, to the `/boot` folder.
13. Change the grub menu to point to the new initrd image.
14. To enable CHAP, you need to modify iscsid.conf (Red Hat only).
15. Reboot.
16. (Optional) Change CHAP parameters.

17. Continue booting into the iSCSI boot image and select one of the images you created (non-offload or offload). Your choice must correspond with your choice in the **iSCSI Boot parameters** section. If **HBA Boot Mode** was enabled in the **iSCSI Boot Parameters** section, you must boot the offload image.

NOTE

Marvell supports Host Bus Adapter (offload) starting in SLES 11 SP1 and later.

18. For IPv6, you can now change the IP address for both the initiator and the target to the IPv6 address that you want in the NVRAM configuration.

NOTE

In SLES 15 (all SPs), perform the following steps so that the OS detects the iSCSI LUN:

1. Add the following line to the grub boot option:

   ```
   iomem=relaxed
   ```

2. Press F6, and then select **YES** to select the driver update disk.

SUSE 11.1 Remote DVD Installation Workaround

1. Create a new file called `boot.open-iscsi` with the content shown in Step 2.
2. Copy the file you just created to `/etc/init.d/` folder and overwrite the existing one.

Contents of the new `boot.open-iscsi` file:

```
#!/bin/bash
#
# /etc/init.d/iscsi
#
### BEGIN INIT INFO
# Provides:          iscsiboot
# Required-Start:
# Should-Start:      boot.multipath
# Required-Stop:
# Should-Stop:       $null
# Default-Start:     B
# Default-Stop:
# Short-Description: iSCSI initiator daemon root-fs support
```

```
# Description:       Starts the iSCSI initiator daemon if the
#                    root-filesystem is on an iSCSI device
#
### END INIT INFO

ISCSIADM=/sbin/iscsiadm
ISCSIUIO=/sbin/iscsiuio
CONFIG_FILE=/etc/iscsid.conf
DAEMON=/sbin/iscsid
ARGS="-c $CONFIG_FILE"

# Source LSB init functions
. /etc/rc.status

#
# This service is run right after booting. So all targets activated
# during mkinitrd run should not be removed when the open-iscsi
# service is stopped.
#
iscsi_load_iscsiuio()
{
    TRANSPORT=`$ISCSIADM -m session 2> /dev/null | grep "bnx2i"`
    if [ "$TRANSPORT" ] ; then
    echo -n "Launch iscsiuio "
        startproc $ISCSIUIO
    fi
}

iscsi_mark_root_nodes()
{
    $ISCSIADM -m session 2> /dev/null | while read t num i target ;
do
    ip=${i%%:*}
    STARTUP=`$ISCSIADM -m node -p $ip -T $target 2> /dev/null |
grep "node.conn\[0\].startup" | cut -d' ' -f3`
    if [ "$STARTUP" -a "$STARTUP" != "onboot" ] ; then
    $ISCSIADM -m node -p $ip -T $target -o update -n
node.conn[0].startup -v onboot
    fi
    done
}
```

```
# Reset status of this service
rc_reset

# We only need to start this for root on iSCSI
if ! grep -q iscsi_tcp /proc/modules ; then
    if ! grep -q bnx2i /proc/modules ; then
        rc_failed 6
        rc_exit
    fi
fi

case "$1" in
    start)
    echo -n "Starting iSCSI initiator for the root device: "
    iscsi_load_iscsiuio
    startproc $DAEMON $ARGS
    rc_status -v
    iscsi_mark_root_nodes
    ;;
    stop|restart|reload)
    rc_failed 0
    ;;
    status)
    echo -n "Checking for iSCSI initiator service: "
    if checkproc $DAEMON ; then
    rc_status -v
    else
    rc_failed 3
    rc_status -v
    fi
    ;;
    *)
    echo "Usage: $0 {start|stop|status|restart|reload}"
    exit 1
    ;;
esac
rc_exit
```

Removing Inbox Drivers from Windows OS Image

1. Create a temporary folder, such as `D:\temp`.
2. Create the following two subfolders in the temporary folder:
   - ❑ `Win2008R2Copy`
   - ❑ `Win2008R2Mod`
3. Copy all the contents from the DVD installation media into the `Win2008R2Copy` folder.
4. Open the Windows Automated Installation Kit (AIK) command prompt in elevated mode from All program, and then issue the following command:

   ```
   attrib -r D:\Temp\Win2008R2Copy\sources\boot.wim
   ```

5. Issue the following command to mount the `boot.wim` image:

   ```
   dism /Mount-WIM
   /WimFile:D:\Temp\Win2008R2Copy\sources\boot.wim /index:1 /
   MountDir:D:\Temp\Win2008R2Mod
   ```

6. The `boot.wim` image was mounted in the `Win2008R2Mod` folder. In the subfolders of the `Win2008R2Mod` folder, locate and delete all instances of the following files:
   - ❑ `netevbda.inf`
   - ❑ `netevbda.pnf`
   - ❑ `evbda.sys`
   - ❑ `netbxnda.inf`
   - ❑ `netbxnda.pnf`
   - ❑ `bxnd60a.sys`
   - ❑ `bxvbda.sys`
   - ❑ `netbvbda.inf`
   - ❑ `netbvbda.pnf`

   To easily find all the instances of the files to be deleted, issue the following command:

   ```
   dir /s D:\Temp\Win2008R2Mod\filename
   ```

7. To unmount the `Boot.wim` image, issue the following command:

   ```
   dism /unmount-wim /Mountdir:D:\Temp\Win2008R2Mod /commit
   ```

8. Repeat steps 5 to 7, but set the `index = 2` for the command in Step 5.

   In this example, `index 2` is specified for the standard edition. For other editions, change the index accordingly.

Injecting (Slipstreaming) Marvell Drivers into Windows Image Files

See these instructions in the FCoE topic.

To inject Marvell drivers into the Windows image files, you must obtain the driver installation packages for the applicable Windows Server version.

Place these driver packages to a working directory. For example, copy all driver packages and files applicable to your Windows Server version to example folder location in Step 3:

- `C:\Temp\drivers`

Finally, inject these drivers into the Windows Image (WIM) files and install the applicable Windows Server version from the updated images.

To inject Marvell drivers into Windows image files:

1. For Windows Server 2016/2019/Azure Stack HCI, install the Windows Assessment and Deployment Kit (ADK).
2. Issue the following commands to create a temporary folder and set it as the current folder for all later steps:

   ```
   md C:\Temp
   cd /d C:\Temp
   ```

3. Issue the following commands to create two subfolders, in `C:\temp`:

   ```
   md src
   md mnt
   md drivers
   ```

4. Issue the following command to copy the original DVD into the `src` subdirectory.

   ```
   xcopy N:\ .\src /e /c /i /f /h /k /y /q
   ```

   Note that in this example, the installation DVD is in the `N:` drive.

5. Open a Deployment and Imaging Tools command prompt in elevated (Administrator) mode.Then, set `c:\Temp` as the current folder.

   Note that you will use this command prompt window in all subsequent steps.

6. Issue the following commands:

   ```
   attrib -r .\src\sources\boot.wim
   attrib -r .\src\sources\install.wim
   ```

7. Issue the following command to mount the `boot.wim` image:

   ```
   dism /mount-wim /wimfile:.\src\sources\boot.wim /index:2
   /mountdir:.\mnt
   ```

   Note: you must always use “2” for the index value.

8. Issue the following commands to add the following drivers to the currently mounted image:

```
dism /image:.\mnt /add-driver /driver:C:\Temp\drivers /Recurse /ForceUnsigned
```

9. Issue the following command to unmount the `boot.wim` image:

   ```
   dism /unmount-wim /mountdir:.\mnt /commit
   ```

10. Issue the following command to determine the index of the SKU that you want in the `install.wim` image:

    ```
    dism /get-wiminfo /wimfile:.\src\sources\install.wim
    ```

11. Issue the following command to mount the `install.wim` image:

    ```
    dism /mount-wim /wimfile:.\src\sources\install.wim /index:X
    /mountdir:.\mnt
    ```

    Note: *X* is a placeholder for the index value that you obtained in the previous step.

12. To add these drivers to the currently mounted image, issue the following commands:

```
dism /image:.\mnt /add-driver /driver:C:\Temp\drivers /Recurse /ForceUnsigned
```

13. To unmount the `install.wim` image, issue the following command:

    ```
    dism /unmount-wim /mountdir:.\mnt /commit
    ```

14. Prepare for ISO creation by copying boot files to `C:\temp`:

    ```
    copy "<AIK or ADK path>\..\etfsboot.com" C:\Temp
    copy "<AIK or ADK path>\..\efisys.bin" C:\Temp
    ```

15. Issue the following command to create an `.iso` file:

```
oscdimg -m -o -u2 -udfver102 -lslipstream -bootdata:2#p0,e,b"c:\Temp\
etfsboot.com"#pEF,e,b"C:\Temp\efisys.bin" c:\temp\src c:\temp\Win20xxMOD.iso
```

    Note: the `xx` in the filenames is a placeholder for the Windows Server OS version.

16. Using a DVD-burning application, burn the `.iso` file you created to a DVD.
17. Use the DVD that you created in the previous step to install the applicable Windows Server version.

Booting

After that the system has been prepared for an iSCSI boot and the operating system is present on the iSCSI target, the last step is to perform the actual boot. The system will boot to Windows or Linux over the network and operate as if it were a local disk drive.

1. Reboot the server.
2. Press the CTRL+S keys.
3. To boot through an offload path, set the **HBA Boot Mode** to **Enabled**. To boot through a non-offload path, set the **HBA Boot Mode** to **Disabled**. (This parameter cannot be changed when the adapter is in multi-function mode.)

If CHAP authentication is needed, enable CHAP authentication after determining that booting is successful (see “Enabling CHAP Authentication” on page 112).

Other iSCSI Boot Considerations

Consider these additional factors when configuring a system for iSCSI boot.

Changing the Speed and Duplex Settings in Windows Environments

Changing the Speed & Duplex settings on the boot port using Windows Device Manager when performing iSCSI boot through the offload path is not supported. Booting through the NDIS path is supported. The Speed & Duplex settings can be changed using the QCS management utility for iSCSI boot through the offload and NDIS paths.

Locally Administered Address

A user-defined MAC address assigned through the **Locally Administered Address** property of the **Advanced** section on the applicable utility Configurations page is not supported on iSCSI boot-enabled devices.

Virtual LANs

Virtual LAN (VLAN) tagging is not supported for iSCSI boot with the Microsoft iSCSI software initiator.

“DD” Method of Creating an iSCSI Boot Image

If installation directly to a remote iSCSI target is not an option, use the “DD” method as an alternate way to create such an image. This method requires you to install the image directly to a local hard drive, and then create an iSCSI boot image for the subsequent boot.

To create an iSCSI boot image with “DD”:

1. Install Linux OS on your local hard drive and ensure that the Open-iSCSI initiator is up to date.
2. Ensure that all run levels of network service are on.
3. Ensure that the 2, 3, and 5 runlevels of iSCSI service are on.
4. Update iscsiuio. You can get the iscsiuio package from the QLogic CD. This step is not needed for SUSE 10.
5. Install the linux-nx2 package on your Linux system. You can get this package from the QLogic CD.
6. Install the bibt package on your Linux system. You can get this package from QLogic CD.
7. Delete all `ifcfg-eth*` files.
8. Configure one port of the network adapter to connect to iSCSI target (for instructions, see “Configuring the iSCSI Target” on page 102).
9. Connect to the iSCSI target.
10. Issue the `DD` command to copy from the local hard drive to iSCSI target.
11. When `DD` is done, issue the `sync` command two times, log out, and then log in to iSCSI target again.
12. Issue the `fsck` command on all partitions created on the iSCSI target.
13. Change to the `/OPT/bcm/bibt` folder and run the `iscsi_setup.sh` script to create the initrd images. Option 0 creates a non-offload image and option 1 creates an offload image. The `iscsi_script.sh` script creates the non-offload image only on SUSE 10 because offload is not supported on SUSE 10.
14. Mount the `/boot` partition on the iSCSI target.
15. Copy the initrd images you created in Step 13 from your local hard drive to the partition mounted in Step 14.
16. On the partition mounted in Step 14, edit the grub menu to point to the new initrd images.
17. Unmount the `/boot` partition on the iSCSI target.
18. (Red Hat Only) To enable CHAP, you need to modify the CHAP section of the `iscsid.conf` file on the iSCSI target. As needed, edit the `iscsid.conf` file with one-way or two-way CHAP information.
19. Shut down the system and disconnect the local hard drive.

    You are now ready to iSCSI boot the iSCSI target.

20. (Optional) Configure iSCSI boot parameters, including CHAP parameters (see “Configuring the iSCSI Target” on page 102).
21. Continue booting into the iSCSI boot image and choose one of the images you created (non-offload or offload). Your choice should correspond with your choice in the **iSCSI Boot Parameters** section. If **HBA Boot Mode** was enabled in the **iSCSI Boot Parameters** section, you must boot the offload image.

Troubleshooting iSCSI Boot

The following troubleshooting tips are useful for iSCSI boot.

Problem: The Marvell iSCSI Crash Dump utility will not work properly to capture a memory dump when the link speed for iSCSI boot is configured for 10Mbps or 100Mbps.
Solution: The iSCSI Crash Dump utility is supported when the link speed for iSCSI boot is configured for 1Gbps or 10Gbps. 10Mbps and 100Mbps are not supported.

Problem: When switching iSCSI boot from the Microsoft standard path to Marvell iSCSI offload, the booting fails to complete.
Solution: Prior to switching the iSCSI boot path, install or upgrade the Marvell Virtual Bus Device (VBD) driver to and OIS driver to the latest versions.

Problem: The iSCSI configuration utility will not run.
Solution: Ensure that the iSCSI Boot firmware is installed in the NVRAM.

Problem: A system blue screen occurs when installing the Marvell drivers through Windows Plug-and-Play (PnP).
Solution: Install the drivers through the Setup installer.

Problem: For static IP configuration when switching from Layer 2 iSCSI boot to Marvell iSCSI Host Bus Adapter, you receive an IP address conflict.
Solution: Change the IP address of the network property in the OS.

Problem: After configuring the iSCSI boot LUN to 255, a system blue screen appears when performing iSCSI boot.
Solution: Although Marvell’s iSCSI solution supports a LUN range from 0 to 255, the Microsoft iSCSI software initiator does not support a LUN of 255. Configure a LUN value from 0 to 254.

Problem: NDIS miniports with Code 31 yellow-bang after Layer 2 iSCSI boot install.
Solution: Run the latest version of the driver installer.

Problem: Unable to update inbox driver if a non-inbox hardware ID present.
Solution: Create a custom slipstream DVD image with supported drivers present on the install media.

Problem: iSCSI-Offload boot from SAN fails to boot after installation.
Solution: Follow the instructions in "Linux" on page 289.

Problem: Installing Windows onto an iSCSI target through iSCSI boot fails when connecting to a 1Gbps switch port.
Solution: This failure is a limitation relating to adapters that use SFP+ as the physical connection. SFP+ defaults to 10Gbps operation and does not support autonegotiation.

iSCSI Crash Dump

If you use the Marvell iSCSI Crash Dump utility, you must install the iSCSI Crash Dump driver. For more information, see "Using the Installer" on page 90.

iSCSI Offload in Windows Server

iSCSI traffic may be segregated offload is a technology that offloads iSCSI protocol processing overhead from host processors to the iSCSI Host Bus Adapter to increase network performance and throughput while helping to optimize server processor utilization.

This section covers Marvell's iSCSI offload feature for the 57*xx* and 57*xxx* family of network adapters on Windows Server systems. For Linux iSCSI offload, see "Linux iSCSI Offload" on page 58.

Configuring iSCSI Offload

With the proper iSCSI offload licensing, you can configure your iSCSI-capable 57*xx* and 57*xxx* network adapter to offload iSCSI processing from the host processor. The following process enables your system to take advantage of Marvell's iSCSI offload feature.

- Installing Marvell Drivers and Management Applications
- Installing the Microsoft iSCSI Initiator
- Configuring Marvell iSCSI Using QCC
- Configuring Microsoft Initiator to Use the Marvell iSCSI Offload

Installing Marvell Drivers and Management Applications

Install the Windows drivers and management applications.

Installing the Microsoft iSCSI Initiator

For Windows Server 2012 and later, the iSCSI initiator is included inbox. To download the iSCSI initiator from Microsoft (if not already installed), locate the direct link for your system here:

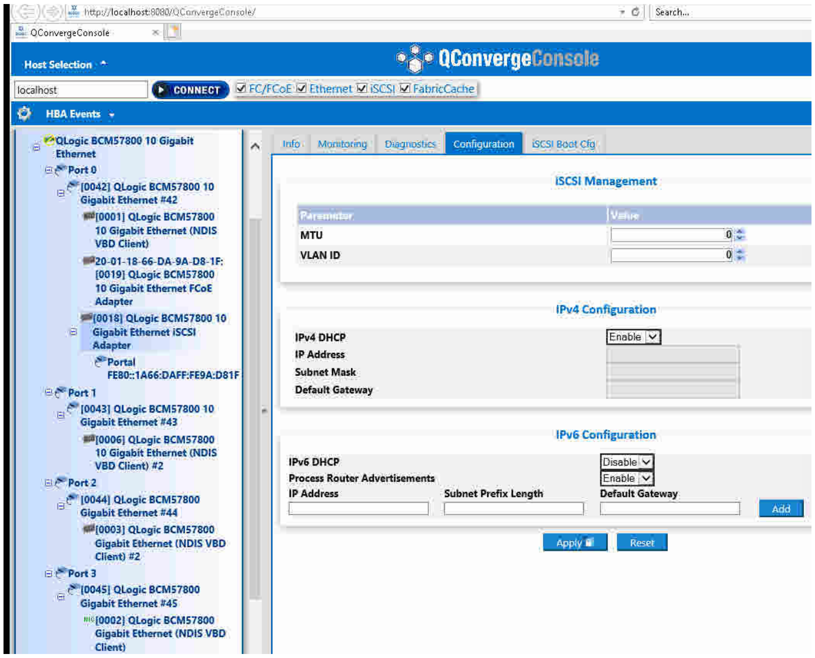
http://www.microsoft.com/en-us/download/details.aspx?displaylang=en&id=18986

Configuring Marvell iSCSI Using QCC

Use the QConvergeConsole (QCC) GUI to manage all of Marvell's network adapters and advanced features. For more information, see the QCC GUI online help.

1. Open QCC GUI.
2. Select the Marvell 57*xx* and 57*xxx* C-NIC iSCSI adapter. If the C-NIC iSCSI adapter instance is not present in the QCC GUI tree view, select the VBD device (the item between the PORT and the Ethernet/NDIS or iSCSI-Offload or FCoE-Offload items in the tree view) and enable iSCSI offload by selecting **iSCSI Offload Engine** from the **Resource Config** tab (see Figure 14-30 on page 239).
3. In the iSCSI-Offload item (in the tree view), select the **Configuration** tab.

On this page, you can change the iSCSI-Offload MTU size, the iSCSI-Offload VLAN ID, the IPv4/IPv6 DHCP setting, the IPv4/IPv6 Static Address/Subnet Mask/Default Gateway settings, and the IPv6 Process Router Advertisements setting (see Figure 11-6).


http://localhost:8080/QConvergeConsole/
Search...
QConvergeConsole
Host Selection
QConvergeConsole
localhost
CONNECT
☑ FC/FCoE ☑ Ethernet ☑ iSCSI ☑ FabricCache
HBA Events
QLogic BCM57800 10 Gigabit Ethernet
Port 0
[0042] QLogic BCM57800 10 Gigabit Ethernet #42
[0001] QLogic BCM57800 10 Gigabit Ethernet (NDIS VBD Client)
20-01-18-66-DA-9A-D8-1F: [0019] QLogic BCM57800 10 Gigabit Ethernet FCoE Adapter
[0018] QLogic BCM57800 10 Gigabit Ethernet iSCSI Adapter
Portal FE80::1A66:DAFF:FE9A:D81F
Port 1
[0043] QLogic BCM57800 10 Gigabit Ethernet #43
[0006] QLogic BCM57800 10 Gigabit Ethernet (NDIS VBD Client) #2
Port 2
[0044] QLogic BCM57800 Gigabit Ethernet #44
[0003] QLogic BCM57800 Gigabit Ethernet (NDIS VBD Client) #2
Port 3
[0045] QLogic BCM57800 Gigabit Ethernet #45
[0002] QLogic BCM57800 Gigabit Ethernet (NDIS VBD Client)
Info
Monitoring
Diagnostics
Configuration
iSCSI Boot Cfg
Add
Apply
Reset

iSCSI Management

Parameter	Value
MTU	0
VLAN ID	0

IPv4 Configuration

IPv4 DHCP: Enable
IP Address
Subnet Mask
Default Gateway

IPv6 Configuration

IPv6 DHCP: Disable
Process Router Advertisements: Enable
IP Address
Subnet Prefix Length
Default Gateway

Figure 11-6. Configuring iSCSI Using QCC

4. DHCP is the default for IP address assignment, but you can change it to a static IP address assignment, if this is the preferred method of IP address assignment.

NOTE

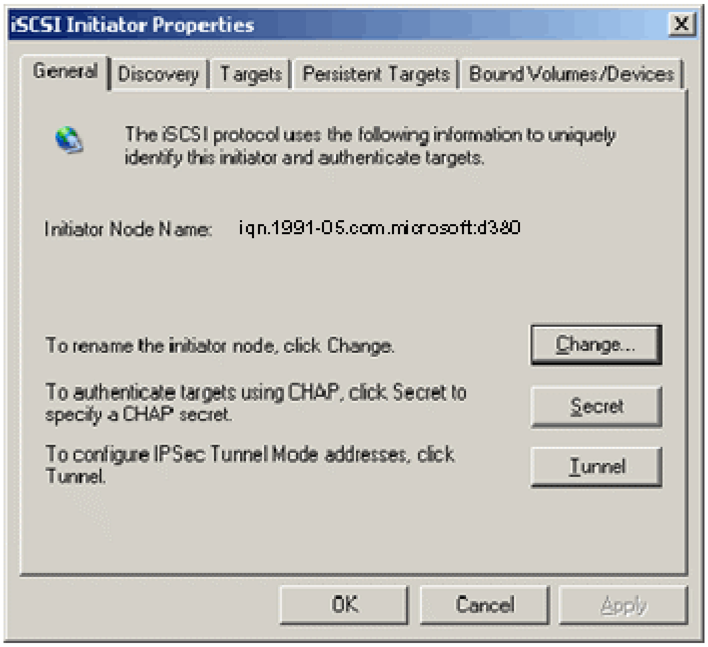
The iSCSI-Offload IP address assignment method cannot be changed if the adapter port was used for iSCSI-Offload remote boot.

5. Click **Apply** and close QCC GUI.

Configuring Microsoft Initiator to Use the Marvell iSCSI Offload

After you have configured the IP address for the iSCSI adapter, you must use Microsoft Initiator to configure and add a connection to the iSCSI target using a Marvell iSCSI adapter. See Microsoft's user guide for more details on the Microsoft Initiator.

1. Open Microsoft Initiator.
2. Configure the initiator IQN name according to your setup. On the iSCSI Initiator Properties, General page (see Figure 11-7), click **Change**.



The iSCSI protocol uses the following information to uniquely identify this initiator and authenticate targets.

Initiator Node Name: iqn.1991-05.com.microsoft:d380

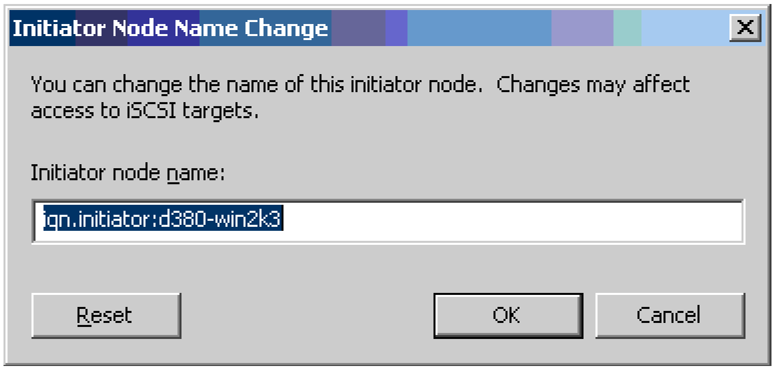
To rename the initiator node, click Change.

To authenticate targets using CHAP, click Secret to specify a CHAP secret.

To configure IPSec Tunnel Mode addresses, click Tunnel.

Figure 11-7. iSCSI Initiator Properties: General Page

3. In the Initiator Node Name Change dialog box (see Figure 11-8), type the initiator IQN name, and then click **OK**.



You can change the name of this initiator node. Changes may affect access to iSCSI targets.

Initiator node name: iqn.initiator:d380-win2k3

Figure 11-8. Changing the Initiator Node Name

4. On the iSCSI Initiator Properties (Figure 11-9), click the **Discovery** tab, and then under **Target Portals**, click **Add**.

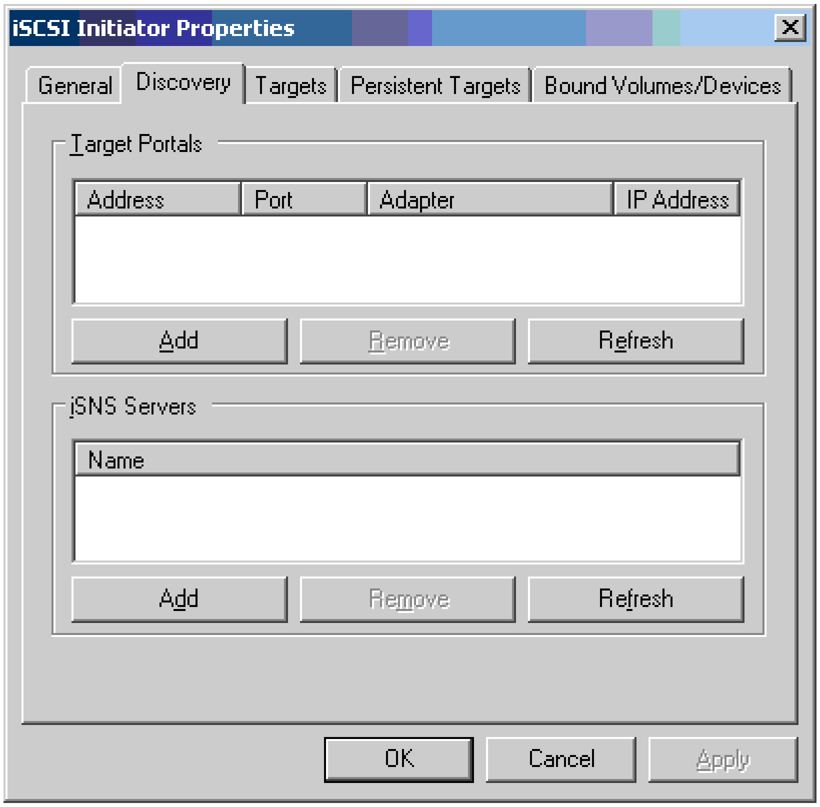
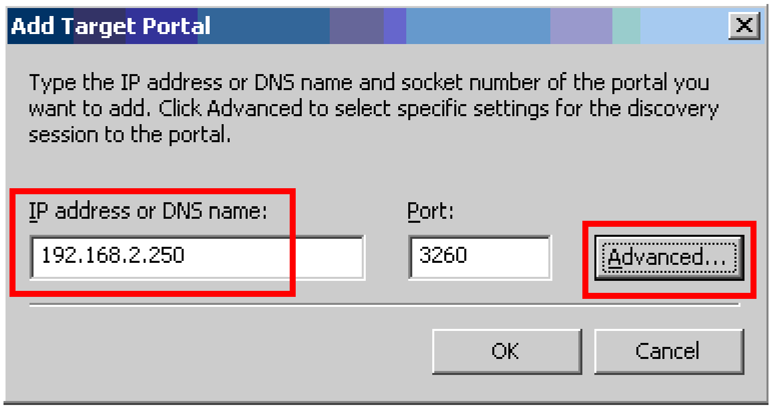


Figure 11-9. iSCSI Initiator Properties: Discovery Page

5. On the Add Target Portal dialog box (Figure 11-10), type the IP address of the target, and then click **Advanced**.



Type the IP address or DNS name and socket number of the portal you want to add. Click Advanced to select specific settings for the discovery session to the portal.

IP address or DNS name: 192.168.2.250

Port: 3260

Figure 11-10. Add Target Portal Dialog Box

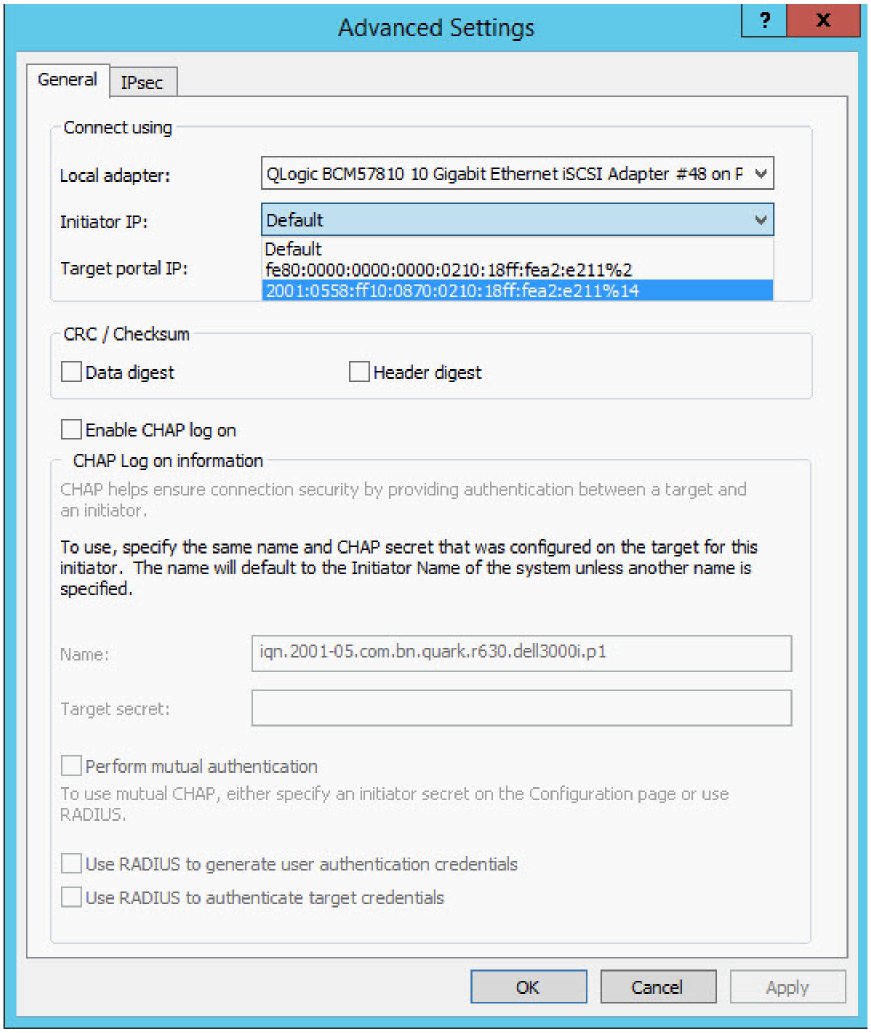
6. On the Advanced Settings dialog box, complete the General page as follows:

   a. For the **Local adapter**, select the Marvell 57*xx* and 57*xxx* C-NIC iSCSI adapter.

   b. For the **Source IP**, select the IP address for the adapter.

   c. To close the Advanced Settings dialog box and save your changes, click **OK**.

Figure 11-11 shows an example.



Advanced Settings

General | IPsec

Connect using

Local adapter: QLogic BCM57810 10 Gigabit Ethernet iSCSI Adapter #48 on F

Initiator IP: Default
- Default
- fe80:0000:0000:0000:0210:18ff:fea2:e211%2
- 2001:0558:ff10:0870:0210:18ff:fea2:e211%14

Target portal IP:

CRC / Checksum

□ Data digest □ Header digest

□ Enable CHAP log on

CHAP Log on information

CHAP helps ensure connection security by providing authentication between a target and an initiator.

To use, specify the same name and CHAP secret that was configured on the target for this initiator. The name will default to the Initiator Name of the system unless another name is specified.

Name: iqn.2001-05.com.bn.quark.r630.dell3000i.p1

Target secret:

□ Perform mutual authentication

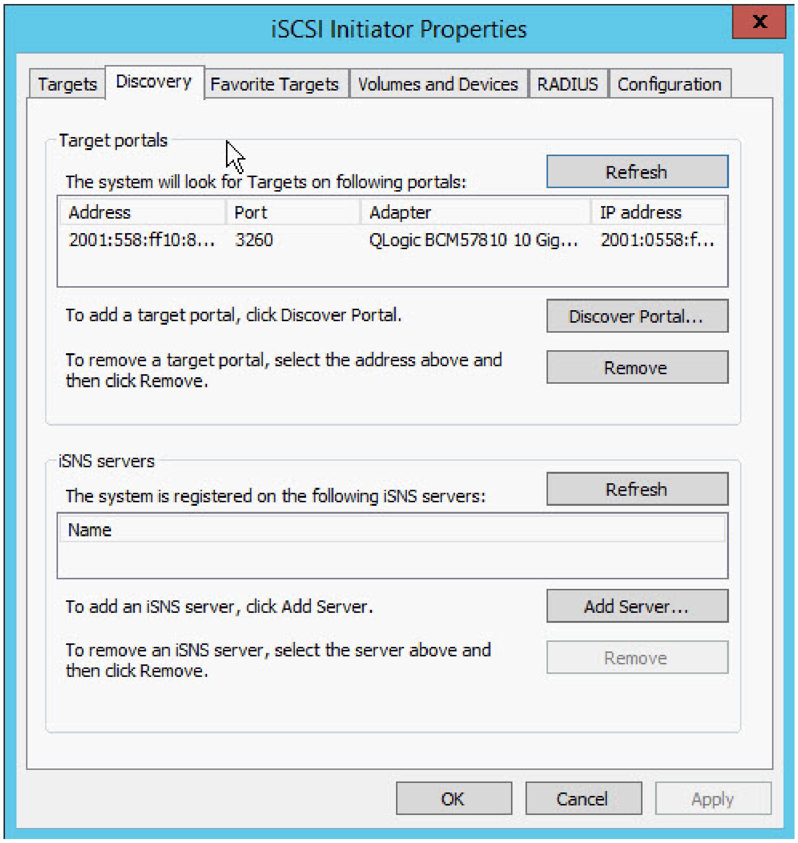
To use mutual CHAP, either specify an initiator secret on the Configuration page or use RADIUS.

□ Use RADIUS to generate user authentication credentials

□ Use RADIUS to authenticate target credentials

Figure 11-11. Advanced Settings: General Page

7. On the iSCSI Initiator Properties, click the **Discovery** tab, and then on the Discovery page, click **OK** to add the target portal. Figure 11-12 shows an example.



The system will look for Targets on following portals:

Address	Port	Adapter	IP address
2001:558:ff10:8...	3260	QLogic BCM57810 10 Gig...	2001:0558:f...

To add a target portal, click Discover Portal.

To remove a target portal, select the address above and then click Remove.

The system is registered on the following iSNS servers:

Name

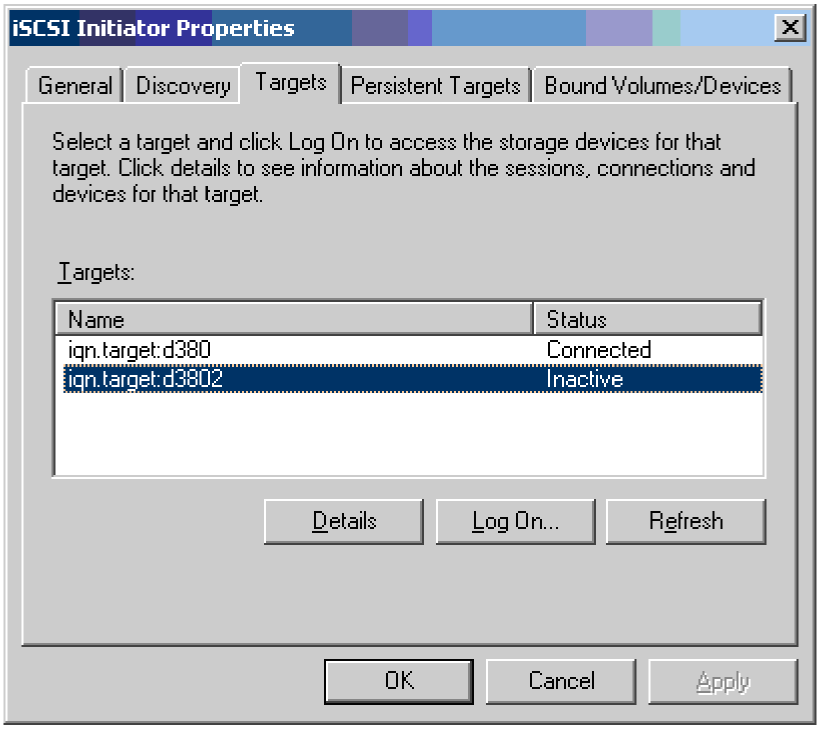
To add an iSNS server, click Add Server.

To remove an iSNS server, select the server above and then click Remove.

Figure 11-12. iSCSI Initiator Properties: Discovery Page

8. On the iSCSI Initiator Properties, click the **Targets** tab.

9. On the Targets page, select the target, and then click **Log On** to log into your iSCSI target using the Marvell iSCSI adapter. Figure 11-13 shows an example.



iSCSI Initiator Properties

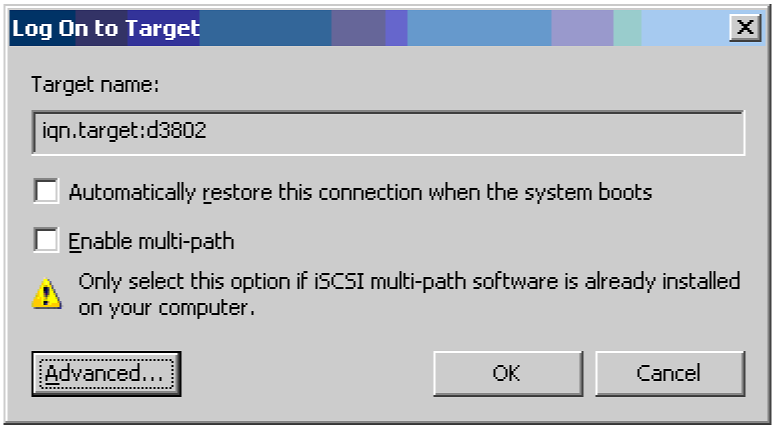
Select a target and click Log On to access the storage devices for that target. Click details to see information about the sessions, connections and devices for that target.

Targets:

Name	Status
iqn.target:d380	Connected
iqn.target:d3802	Inactive

Figure 11-13. iSCSI Initiator Properties: Targets Page

10. On the Log On To Target dialog box (Figure 11-14), click **Advanced**.



Log On to Target

Target name:

iqn.target:d3802

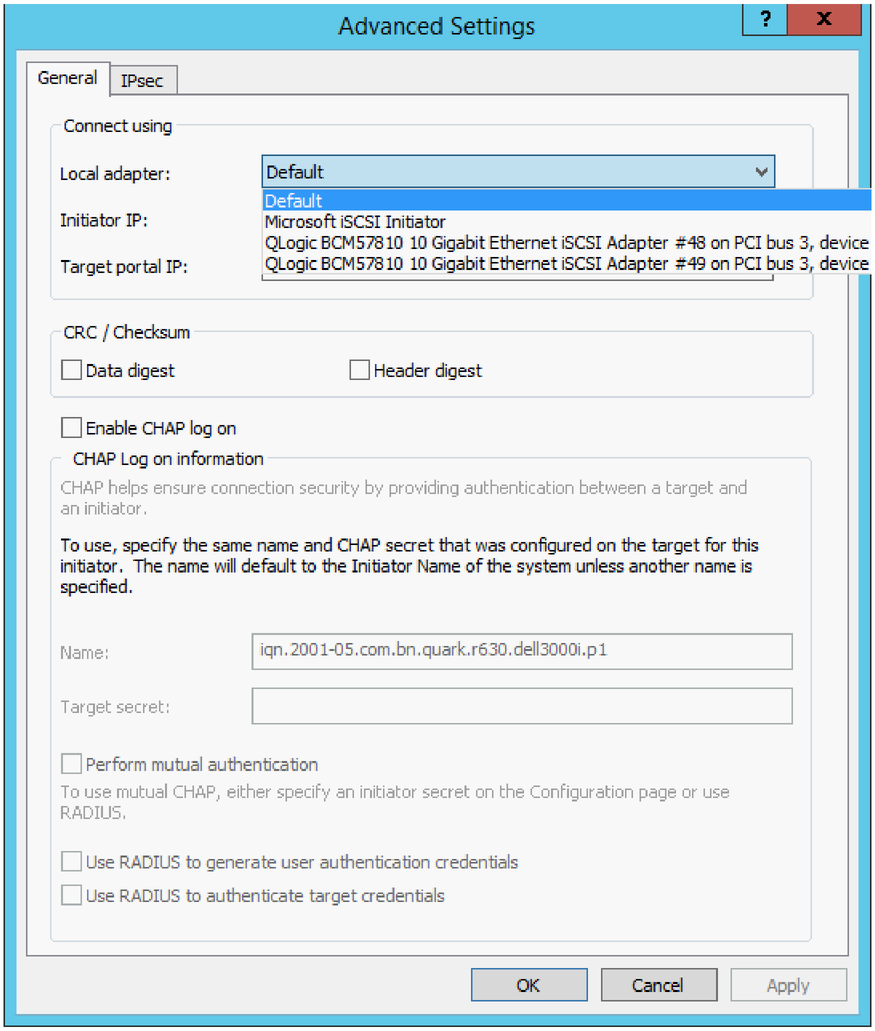
□ Automatically restore this connection when the system boots

□ Enable multi-path

⚠ Only select this option if iSCSI multi-path software is already installed on your computer.

Figure 11-14. Log On to Target

11. On the Advanced Settings dialog box, General page, select the Marvell 57*xx* and 57*xxx* C-NIC iSCSI adapters as the **Local adapter**, and then click **OK**. Figure 11-15 shows an example.



Advanced Settings

General | IPsec

Connect using

Local adapter: Default
- Default
- Microsoft iSCSI Initiator
- QLogic BCM57810 10 Gigabit Ethernet iSCSI Adapter #48 on PCI bus 3, device
- QLogic BCM57810 10 Gigabit Ethernet iSCSI Adapter #49 on PCI bus 3, device

Initiator IP:

Target portal IP:

CRC / Checksum

☐ Data digest ☐ Header digest

☐ Enable CHAP log on

CHAP Log on information

CHAP helps ensure connection security by providing authentication between a target and an initiator.

To use, specify the same name and CHAP secret that was configured on the target for this initiator. The name will default to the Initiator Name of the system unless another name is specified.

Name: iqn.2001-05.com.bn.quark.r630.dell3000i.p1

Target secret:

☐ Perform mutual authentication

To use mutual CHAP, either specify an initiator secret on the Configuration page or use RADIUS.

☐ Use RADIUS to generate user authentication credentials

☐ Use RADIUS to authenticate target credentials

Figure 11-15. Advanced Settings: General Page, Local Adapter

12. Click **OK** to close the Microsoft Initiator.

13. To format your iSCSI partition, use Disk Manager.

NOTE

- Teaming does not support iSCSI adapters.
- Teaming does not support NDIS adapters that are in the boot path.
- Teaming supports NDIS adapters that are not in the iSCSI boot path, but only for the SLB or switch-independent team type.

iSCSI Offload FAQs

Question: How do I assign an IP address for iSCSI offload?
Answer: Use the Configurations page in the applicable management utility.

Question: What tools should be used to create the connection to the target?
Answer: Use Microsoft iSCSI Software Initiator (version 2.08 or later).

Question: How do I know that the connection is offloaded?
Answer: Use Microsoft iSCSI Software Initiator. From a command line, type `iscsicli sessionlist`. From **Initiator Name**, an iSCSI offloaded connection will display an entry beginning with "B06BDRV…" (for the 57*xx*) or "EBDRV…" (for the 57*xxx*). A non-offloaded connection displays an entry beginning with "Root...".

Question: What configurations should be avoided?
Answer: The IP address should not be the same as the LAN.

Question: Why does the install fail when attempting to complete an iSCSI offload install using Windows Server OS for 57*xx* and 57*xxx* adapters?
Answer: There is a conflict with the internal inbox driver.

Event Log Messages for Offload iSCSI (OIS) Driver

Table 11-5 lists the offload iSCSI driver event log messages.

Table 11-5. Offload iSCSI (OIS) Driver Event Log Messages

Message Number	Severity	Message
1	Error	Initiator failed to connect to the target. Target IP address and TCP Port number are specific in dump data.
2	Error	The initiator could not allocate resources for an iSCSI session.
3	Error	Maximum command sequence number is not serially greater than expected command sequence number in login response. Dump data contains Expected Command Sequence number followed by Maximum Command Sequence number.

Table 11-5. Offload iSCSI (OIS) Driver Event Log Messages (Continued)

Message Number	Severity	Message
4	Error	MaxBurstLength is not serially greater than FirstBurstLength. Dump data contains FirstBurstLength followed by MaxBurstLength.
5	Error	Failed to setup initiator portal. Error status is specified in the dump data.
6	Error	The initiator could not allocate resources for an iSCSI connection.
7	Error	The initiator could not send an iSCSI PDU. Error status is specified in the dump data.
8	Error	Target or discovery service did not respond in time for an iSCSI request sent by the initiator. iSCSI Function code is specified in the dump data. For details about iSCSI Function code, refer to the iSCSI user's guide.
9	Error	Target did not respond in time for a SCSI request. The CDB is specified in the dump data.
10	Error	Login request failed. The login response packet is specified in the dump data.
11	Error	Target returned an invalid login response packet. The login response packet is specified in the dump data.
12	Error	Target provided invalid data for login redirect. Dump data contains the data returned by the target.
13	Error	Target offered an unknown AuthMethod. Dump data contains the data returned by the target.
14	Error	Target offered an unknown digest algorithm for CHAP. Dump data contains the data returned by the target.
15	Error	CHAP challenge specified by the target contains invalid characters. Dump data contains the specified challenge.
16	Error	An invalid key was received during CHAP negotiation. The key=value pair is specified in the dump data.
17	Error	CHAP Response specified by the target did not match the expected one. Dump data contains the CHAP response.
18	Error	Header Digest is required by the initiator, but target did not offer it.
19	Error	Data Digest is required by the initiator, but target did not offer it.
20	Error	Connection to the target was lost. The initiator will attempt to retry the connection.
21	Error	Data Segment Length specified in the header exceeds `MaxRecvDataSegmentLength` declared by the target.

Table 11-5. Offload iSCSI (OIS) Driver Event Log Messages (Continued)

Message Number	Severity	Message
22	Error	Header digest error was detected for the specified PDU. Dump data contains the header and digest.
23	Error	Target sent an invalid iSCSI PDU. Dump data contains the entire iSCSI header.
24	Error	Target sent an iSCSI PDU with an invalid opcode. Dump data contains the entire iSCSI header.
25	Error	Data digest error was detected. Dump data contains the calculated checksum followed by the specified checksum.
26	Error	Target trying to send more data than requested by the initiator.
27	Error	Initiator could not find a match for the initiator task tag in the received PDU. Dump data contains the entire iSCSI header.
28	Error	Initiator received an invalid R2T packet. Dump data contains the entire iSCSI header.
29	Error	Target rejected an iSCSI PDU sent by the initiator. Dump data contains the rejected PDU.
30	Error	Initiator could not allocate a work item for processing a request.
31	Error	Initiator could not allocate resource for processing a request.
32	Information	Initiator received an asynchronous logout message. The Target name is specified in the dump data.
33	Error	Challenge size specified by the target exceeds the maximum specified in iSCSI specification.
34	Information	A connection to the target was lost, but Initiator successfully reconnected to the target. Dump data contains the target name.
35	Error	Target CHAP secret is smaller than the minimum size (12 bytes) required by the specification.
36	Error	Initiator CHAP secret is smaller than the minimum size (12 bytes) required by the specification. Dump data contains the specified CHAP secret.
37	Error	FIPS service could not be initialized. Persistent logons will not be processed.
38	Error	Initiator requires CHAP for logon authentication, but target did not offer CHAP.
39	Error	Initiator sent a task management command to reset the target. The target name is specified in the dump data.

Table 11-5. Offload iSCSI (OIS) Driver Event Log Messages (Continued)

Message Number	Severity	Message
40	Error	Target requires logon authentication through CHAP, but Initiator is not configured to perform CHAP.
41	Error	Target did not send AuthMethod key during security negotiation phase.
42	Error	Target sent an invalid status sequence number for a connection. Dump data contains Expected Status Sequence number followed by the specified status sequence number.
43	Error	Target failed to respond in time for a login request.
44	Error	Target failed to respond in time for a logout request.
45	Error	Target failed to respond in time for a login request. This login request was for adding a new connection to a session.
46	Error	Target failed to respond in time for a SendTargets command.
47	Error	Target failed to respond in time for a SCSI command sent through a WMI request.
48	Error	Target failed to respond in time to a NOP request.
49	Error	Target failed to respond in time to a Task Management request.
50	Error	Target failed to respond in time to a Text Command sent to renegotiate iSCSI parameters.
51	Error	Target failed to respond in time to a logout request sent in response to an asynchronous message from the target.
52	Error	Initiator Service failed to respond in time to a request to configure IPSec resources for an iSCSI connection.
53	Error	Initiator Service failed to respond in time to a request to release IPSec resources allocated for an iSCSI connection.
54	Error	Initiator Service failed to respond in time to a request to encrypt or decrypt data.
55	Error	Initiator failed to allocate resources to send data to target.
56	Error	Initiator could not map an user virtual address to kernel virtual address resulting in I/O failure.
57	Error	Initiator could not allocate required resources for processing a request resulting in I/O failure.
58	Error	Initiator could not allocate a tag for processing a request resulting in I/O failure.

Table 11-5. Offload iSCSI (OIS) Driver Event Log Messages (Continued)

Message Number	Severity	Message
59	Error	Target dropped the connection before the initiator could transition to Full Feature Phase.
60	Error	Target sent data in SCSI Response PDU instead of Data_IN PDU. Only Sense Data can be sent in SCSI Response.
61	Error	Target set DataPduInOrder to NO when initiator requested YES. Login will be failed.
62	Error	Target set DataSequenceInOrder to NO when initiator requested YES. Login will be failed.
63	Error	Cannot reset the target or LUN. Will attempt session recovery.
64	Information	Attempt to bootstrap Windows using iSCSI NIC Boot (iBF).
65	Error	Booting from iSCSI, but could not set any NIC in Paging Path.
66	Error	Attempt to disable the Nagle Algorithm for iSCSI connection failed.
67	Information	If Digest support selected for iSCSI Session, will use Processor support for Digest computation.
68	Error	After receiving an async logout from the target, attempt to relogin the session failed. Error status is specified in the dump data.
69	Error	Attempt to recover an unexpected terminated session failed. Error status is specified in the dump data.
70	Error	Error occurred when processing iSCSI logon request. The request was not retried. Error status is specified in the dump data.
71	Information	Initiator did not start a session recovery upon receiving the request. Dump data contains the error status.
72	Error	Unexpected target portal IP types. Dump data contains the expected IP type.

12 Marvell Teaming Services

This chapter describes teaming for adapters in Windows Server systems (excluding Windows Server 2016 and later). For more information on a similar technologies on other operating systems (for example, Linux Channel Bonding), refer to your operating system documentation.

Microsoft recommends using their in-OS NIC teaming service instead of any adapter vendor-proprietary NIC teaming driver on Windows Server 2012 and later. Marvell's NIC teaming driver is not supported on Windows Server 2016 and later.

Executive Summary

Marvell teaming services are summarized in the following sections:

- Glossary
- Teaming Concepts
- Software Components
- Hardware Requirements
- Teaming Support by Processor
- Configuring Teaming
- Supported Features by Team Type
- Selecting a Team Type

This section describes the technology and implementation considerations when working with the network teaming services offered by the Marvell software shipped with Dell's servers and storage products. The goal of Marvell teaming services is to provide fault tolerance and link aggregation across a team of two or more adapters. The information in this document is provided to assist IT professionals during the deployment and troubleshooting of system applications that require network fault tolerance and load balancing.

Glossary

Table 12-1 defines terminology used in teaming.

Table 12-1. Glossary

Term	Definition
ARP	address resolution protocol
CLI	command line interface
DNS	domain name service
G-ARP	gratuitous address resolution protocol
GUI	graphical user interface
HSRP	hot standby router protocol
ICMP	Internet control message protocol
IGMP	Internet group management protocol
IPv6	Version 6 of the IP
iSCSI	Internet small computer systems interface
Layer 2	Network traffic that is not offloaded, and where hardware only performs Layer 2 operations on the traffic. Layer 3 (IP) and Layer 4 (TCP) protocols are processed in software.
Layer 4	Network traffic that is heavily offloaded to the hardware, where much of the Layer 3 (IP) and Layer 4 (TCP) processing is done in the hardware to improve performance.
LACP	link aggregation control protocol
link aggregation (802.3ad)	Switch-dependent load balancing and failover type of team with LACP in which the intermediate driver manages outgoing traffic and the switch manages incoming traffic
LOM	LAN on motherboard
NDIS	Network Driver Interface Specification

Table 12-1. Glossary (Continued)

Term	Definition
PXE	pre-execution environment
QCC	QConvergeConsole
QCS	QLogic Control Suite
RAID	redundant array of inexpensive disks
TCP	transmission control protocol
UDP	user datagram protocol
WINS	Windows Internet Name Service

Teaming Concepts

The concept of grouping multiple physical devices to provide fault tolerance and load balancing is not new. It has been around for years. Storage devices use RAID technology to group individual hard drives. Switch ports can be grouped together using technologies such as Cisco Gigabit EtherChannel, IEEE 802.3ad Link Aggregation, Bay Networks Multilink Trunking, and Extreme Network Load Sharing. Network interfaces on Dell servers can be grouped together into a team of physical ports called a virtual adapter.

This section provides the following information about teaming concepts:

- Network Addressing
- Teaming and Network Addresses
- Description of Teaming Types

Network Addressing

To understand how teaming works, it is important to understand how node communications work in an Ethernet network. This document is based on the assumption that the reader is familiar with the basics of IP and Ethernet network communications.

The following information provides a high-level overview of the concepts of network addressing used in an Ethernet network. Every Ethernet network interface in a host platform, such as a computer system, requires a globally unique Layer 2 address and at least one globally unique Layer 3 address. Layer 2 is the data link layer, and Layer 3 is the network layer as defined in the OSI model. The Layer 2 address is assigned to the hardware and is often referred to as the MAC address or physical address. This address is pre-programmed at the factory and stored in NVRAM on a network interface card or on the system motherboard for an embedded LAN interface. The Layer 3 addresses are referred to as the protocol or logical address assigned to the software stack. IP and IPX are examples of Layer 3 protocols. In addition, Layer 4 (Transport Layer) uses port numbers for each network upper level protocol such as Telnet or FTP. These port numbers are used to differentiate traffic flows across applications. Layer 4 protocols such as TCP or UDP are most commonly used in today's networks. The combination of the IP address and the TCP port number is called a socket.

Ethernet devices communicate with other Ethernet devices using the MAC address, not the IP address. However, most applications work with a host name that is translated to an IP address by a naming service such as Windows Internet Name Service (WINS) and DNS. Therefore, a method of identifying the MAC address assigned to the IP address is required. The address resolution protocol for an IP network provides this mechanism. For IPX, the MAC address is part of the network address and ARP is not required. ARP is implemented using an ARP Request and ARP Reply frame. ARP Requests are typically sent to a broadcast address while the ARP Reply is typically sent as unicast traffic. A unicast address corresponds to a single MAC address or a single IP address. A broadcast address is sent to all devices on a network.

Teaming and Network Addresses

A team of adapters function as a single virtual network interface and does not appear any different to other network devices than a non-teamed adapter. A virtual network adapter advertises a single Layer 2 and one or more Layer 3 addresses. When the teaming driver initializes, it selects one MAC address from one of the physical adapters that make up the team to be the Team MAC address. This address is typically taken from the first adapter that gets initialized by the driver. When the system hosting the team receives an ARP request, it selects one MAC address from among the physical adapters in the team to use as the source MAC address in the ARP Reply. In Windows operating systems, the `IPCONFIG /all` command shows the IP and MAC address of the virtual adapter and not the individual physical adapters. The protocol IP address is assigned to the virtual network interface and not to the individual physical adapters.

For switch-independent teaming modes, all physical adapters that make up a virtual adapter must use the unique MAC address assigned to them when transmitting data. That is, the frames that are sent by each of the physical adapters in the team must use a unique MAC address to be IEEE compliant. It is important to note that ARP cache entries are not learned from received frames, but only from ARP requests and ARP replies.

Description of Teaming Types

Teaming types described in this section include:

- Smart Load Balancing and Failover
- Generic Trunking
- Link Aggregation (IEEE 802.3ad LACP)
- SLB (Auto-Fallback Disable)

The three methods for classifying the supported teaming types are based on:

- Whether the switch port configuration must also match the adapter teaming type.
- The functionality of the team: whether it supports load balancing and failover, or just failover.
- Whether or not the link aggregation control protocol (LACP) is used.

Table 12-2 shows a summary of the teaming types and their classification.

Table 12-2. Available Teaming Types

Teaming Type	Switch-Dependent [a]	LACP Support Required on the Switch	Load Balancing	Failover
Smart Load Balancing and Failover (with two to eight load balance team members)	—	—	✓	✓
SLB (Auto-Fallback Disable)	—	—	✓	✓
Link Aggregation (802.3ad)	?	?	✓	✓
Generic Trunking (FEC/GEC)/802.3ad-Draft Static	?	—	✓	✓

[a] Switch must support specific type of team.

Smart Load Balancing and Failover

The Smart Load Balancing and Failover type of team provides both load balancing and failover when configured for load balancing, and only failover when configured for fault tolerance. This type of team works with any Ethernet switch and requires no trunking configuration on the switch. The team advertises multiple MAC addresses and one or more IP addresses (when using secondary IP addresses). The team MAC address is selected from the list of load balance members. When the system receives an ARP request, the software-networking stack will always send an ARP Reply with the team MAC address. To begin the load balancing process, the teaming driver will modify this ARP Reply by changing the source MAC address to match one of the physical adapters.

Smart Load Balancing enables both transmit and receive load balancing based on the Layer 3 and Layer 4 IP address and TCP/UDP port number. In other words, the load balancing is not done at a byte or frame level but on a TCP/UDP session basis. This methodology is required to maintain in-order delivery of frames that belong to the same socket conversation. Load balancing is supported on two to eight ports. These ports can include any combination of add-in adapters and LAN on motherboard (LOM) devices.

Transmit load balancing is achieved by creating a hashing table using the source and destination IP addresses and TCP/UDP port numbers.The same combination of source and destination IP addresses and TCP/UDP port numbers generally yield the same hash index and therefore point to the same port in the team. When a port is selected to carry all the frames of a specific socket, the unique MAC address of the physical adapter is included in the frame, and not the team MAC address. This inclusion is required to comply with the IEEE 802.3 standard. If two adapters transmit using the same MAC address, a duplicate MAC address situation would occur that the switch could not handle.

NOTE

IPv6 addressed traffic is load balanced by SLB because ARP is not a feature of IPv6.

Receive load balancing is achieved through an intermediate driver by sending gratuitous ARPs on a client-by-client basis using the unicast address of each client as the destination address of the ARP request (also known as a directed ARP). This practice is considered client load balancing and not traffic load balancing. When the intermediate driver detects a significant load imbalance between the physical adapters in an SLB team, it generates G-ARPs in an effort to redistribute incoming frames. It is important to understand that receive load balancing is a function of the quantity of clients that are connecting to the system through the team interface.

SLB receive load balancing attempts to load balance incoming traffic for client machines across physical ports in the team. It uses a modified gratuitous ARP to advertise a different MAC address for the team IP address in the sender physical and protocol address. The G-ARP is unicast with the MAC and IP Address of a client machine in the target physical and protocol address, respectively. This action causes the target client to update its ARP cache with a new MAC address map to the team IP address. G-ARPs are not broadcast because this would cause all clients to send their traffic to the same port. As a result, the benefits achieved through client load balancing would be eliminated, and could cause out-of-order frame delivery. This receive load-balancing scheme works as long as all clients and the teamed system are on the same subnet or broadcast domain.

When the clients and the system are on different subnets, and incoming traffic has to traverse a router, the received traffic destined for the system is not load balanced. The physical adapter that the intermediate driver has selected to carry the IP flow carries all of the traffic. When the router sends a frame to the team IP address, it broadcasts an ARP request (if not in the ARP cache). The server software stack generates an ARP reply with the team MAC address, but the intermediate driver modifies the ARP reply and sends it over a specific physical adapter, establishing the flow for that session.

The reason is that ARP is not a routable protocol. It does not have an IP header and therefore, is not sent to the router or default gateway. ARP is only a local subnet protocol. In addition, because the G-ARP is not a broadcast packet, the router will not process it and will not update its own ARP cache.

The only way that the router would process an ARP that is intended for another network device is if it has Proxy ARP enabled and the host has no default gateway. This situation is very rare and not recommended for most applications.

Transmit traffic through a router is load balanced because transmit load balancing is based on the source and destination IP address and TCP/UDP port number. Because routers do not alter the source and destination IP address, the load balancing algorithm works as intended.

Configuring routers for hot standby routing protocol (HSRP) does not allow for receive load balancing to occur in the adapter team. In general, HSRP allows for two routers to act as one router, advertising a virtual IP and virtual MAC address. One physical router is the active interface while the other is standby. Although HSRP can also load share nodes (using different default gateways on the host nodes) across multiple routers in HSRP groups, it always points to the primary MAC address of the team.

Generic Trunking

Generic Trunking is a switch-assisted teaming mode and requires configuring ports at both ends of the link: server interfaces and switch ports. This port configuration is often referred to as Cisco Fast EtherChannel or Gigabit EtherChannel. In addition, generic trunking supports similar implementations by other switch OEMs such as Extreme Networks Load Sharing and Bay Networks or IEEE 802.3ad Link Aggregation static mode. In this mode, the team advertises one MAC Address and one IP Address when the protocol stack responds to ARP Requests. In addition, each physical adapter in the team uses the same team MAC address when transmitting frames. Use of the address is possible because the switch at the other end of the link is aware of the teaming mode and will handle the use of a single MAC address by every port in the team. The forwarding table in the switch will reflect the trunk as a single virtual port.

In this teaming mode, the intermediate driver controls load balancing and failover for outgoing traffic only, while incoming traffic is controlled by the switch firmware and hardware. Most switches implement an XOR hashing of the source and destination MAC address.

NOTE

Generic trunking is not supported on iSCSI offload adapters.

Link Aggregation (IEEE 802.3ad LACP)

Link aggregation is similar to generic trunking except that it uses the link aggregation control protocol (LACP) to negotiate the ports that will make up the team. LACP must be enabled at both ends of the link for the team to be operational. If LACP is not available at both ends of the link, 802.3ad provides a manual aggregation that only requires both ends of the link to be in a link up state. Because manual aggregation provides for the activation of a member link without performing the LACP message exchanges, it should not be considered as reliable and robust as an LACP negotiated link. LACP automatically determines which member links can be aggregated and then aggregates them. It provides for the controlled addition and removal of physical links for the link aggregation so that no frames are lost or duplicated. The removal of aggregate link members is provided by the marker protocol that can be optionally enabled for Link Aggregation Control Protocol (LACP) enabled aggregate links.

The Link Aggregation group advertises a single MAC address for all the ports in the trunk. The MAC address of the Aggregator can be the MAC addresses of one of the MACs that make up the group. LACP and marker protocols use a multicast destination address.

The Link Aggregation control function determines which links may be aggregated and then binds the ports to an Aggregator function in the system and monitors conditions to determine if a change in the aggregation group is required. Link aggregation combines the individual capacity of multiple links to form a high performance virtual link. The failure or replacement of a link in an LACP trunk will not cause loss of connectivity. The traffic will simply be failed over to the remaining links in the trunk.

SLB (Auto-Fallback Disable)

This type of team is identical to the Smart Load Balance and Failover type of team, with the following exception—when the standby member is active, if a primary member comes back on line, the team continues using the standby member rather than switching back to the primary member. This type of team is supported only for situations in which the network cable is disconnected and reconnected to the network adapter. It is not supported for situations in which the adapter is removed/installed through Device Manager or Hot-Plug PCI.

If any primary adapter assigned to a team is disabled, the team functions as a Smart Load Balancing and Failover type of team in which auto-fallback occurs.

Software Components

Teaming is implemented through an NDIS intermediate driver in the Windows operating system environment. This software component works with the miniport driver, the NDIS layer, and the protocol stack to enable the teaming architecture (see Figure 12-2 on page 159). The miniport driver controls the host LAN controller directly to enable functions such as sends, receives, and interrupt processing. The intermediate driver fits between the miniport driver and the protocol layer multiplexing several miniport driver instances, and creating a virtual adapter that looks like a single adapter to the NDIS layer. NDIS provides a set of library functions to enable the communications between either miniport drivers or intermediate drivers and the protocol stack. The protocol stack implements IP, IPX, and ARP. A protocol address such as an IP address is assigned to each miniport device instance, but when an Intermediate driver is installed, the protocol address is assigned to the virtual team adapter and not to the individual miniport devices that make up the team.

The Marvell-supplied teaming support is provided by three individual software components that work together and are supported as a package. When one component is upgraded, all the other components must be upgraded to the supported versions.

Table 12-3 describes the four software components and their associated files for supported operating systems.

Table 12-3. Marvell Teaming Software Component

Software Component	Marvell Name	Network Adapter or Operating System	System Architecture	Windows File Name
—	Virtual Bus Driver (VBD)	57*xx*	32-bit	bxvbdx.sys
		57*xx*	64-bit	bxvbda.sys
		5771*x*, 578*xx*	32-bit	evbdx.sys
		5771*x*, 578*xx*	64-bit	evbda.sys
Miniport Driver	QLogic Base Driver	Windows Server 2012, 2012 R2	64-bit	bxnd60a.sys
Configuration User Interface	QCS CLI	Windows Server 2012, 2012 R2	—	qcscli.exe

Hardware Requirements

Hardware requirements for teaming include the following:

- Repeater Hub
- Switching Hub
- Router

The various teaming modes described in this document place specific restrictions on the networking equipment used to connect clients to teamed systems. Each type of network interconnect technology has an effect on teaming as described in the following sections.

Repeater Hub

A Repeater Hub allows a network administrator to extend an Ethernet network beyond the limits of an individual segment. The repeater regenerates the input signal received on one port onto all other connected ports, forming a single collision domain. This domain means that when a station attached to a repeater sends an Ethernet frame to another station, every station within the same collision domain will also receive that message. If two stations begin transmitting at the same time, a collision occurs, and each transmitting station must retransmit its data after waiting a random amount of time.

The use of a repeater requires that each station participating within the collision domain operate in half-duplex mode. Although half-duplex mode is supported for Gigabit Ethernet (GbE) adapters in the IEEE 802.3 specification, half-duplex mode is not supported by the majority of GbE adapter manufacturers. Therefore, half-duplex mode is not considered here.

Teaming across hubs is supported for troubleshooting purposes (such as connecting a network analyzer) for SLB teams only.

Switching Hub

Unlike a repeater hub, a switching hub (or more simply a switch) allows an Ethernet network to be broken into multiple collision domains. The switch is responsible for forwarding Ethernet packets between hosts based solely on Ethernet MAC addresses. A physical network adapter that is attached to a switch may operate in half-duplex or full-duplex mode.

To support Generic Trunking and 802.3ad Link Aggregation, a switch must specifically support such functionality. If the switch does not support these protocols, it may still be used for Smart Load Balancing.

NOTE

All modes of network teaming are supported across switches when operating as a stackable switch.

Router

A router is designed to route network traffic based on Layer 3 or higher protocols, although it often also works as a Layer 2 device with switching capabilities. The teaming of ports connected directly to a router is not supported.

Teaming Support by Processor

All team types are supported by the IA-32 and EM64T processors.

Configuring Teaming

The QConvergeConsole (QCC) GUI and QLogic Control Suite (QCS) CLI utility are used to configure teaming in the supported operating system environments.

These utilities run on 32-bit and 64-bit Windows family of operating systems. Use these utilities to configure VLANs and load balancing and fault tolerance teaming. In addition, they display the MAC address, driver version, and status information for each network adapter. These utilities also include several diagnostics tools such as hardware diagnostics, cable testing, and a network topology test.

Supported Features by Team Type

Table 12-4 provides a feature comparison across the team types supported by Dell. Use this table to determine the best type of team for your application. The teaming software supports up to eight ports in a single team and up to 16 teams in a single system. These teams can be any combination of the supported teaming types, but each team must be on a separate network or subnet.

Table 12-4. Comparison of Team Types

Type of Team	Fault Tolerance	Load Balancing	Switch-Dependent Static Trunking	Switch-Independent Dynamic Link Aggregation (IEEE 802.3ad)
Function	SLB with Standby [a]	SLB	Generic Trunking	Link Aggregation
Quantity of ports per team (same broadcast domain)	2–16	2–16	2–16	2–16
Quantity of teams	16	16	16	16
Adapter fault tolerance	Yes	Yes	Yes	Yes
Switch link fault tolerance (same broadcast domain)	Yes	Yes	Switch-dependent	Switch-dependent
TX load balancing	No	Yes	Yes	Yes
RX load balancing	No	Yes	Yes (performed by the switch)	Yes (performed by the switch)
Requires compatible switch	No	No	Yes	Yes
Heartbeats to check connectivity	No	No	No	No
Mixed media (adapters with different media)	Yes	Yes	Yes (switch-dependent)	Yes
Mixed speeds (adapters that do not support a common speed, but can operate at different speeds)	Yes	Yes	No	No

Table 12-4. Comparison of Team Types (Continued)

Type of Team	Fault Tolerance	Load Balancing	Switch-Dependent Static Trunking	Switch-Independent Dynamic Link Aggregation (IEEE 802.3ad)
Function	**SLB with Standby [a]**	**SLB**	**Generic Trunking**	**Link Aggregation**
Mixed speeds (adapters that support a common speed, but can operate at different speeds)	Yes	Yes	No (must be the same speed)	Yes
Load balances TCP/IP	No	Yes	Yes	Yes
Mixed vendor teaming	Yes [b]	Yes [b]	Yes [b]	Yes [b]
Load balances non-IP	No	Yes (IPX outbound traffic only)	Yes	Yes
Same MAC address for all team members	No	No	Yes	Yes
Same IP address for all team members	Yes	Yes	Yes	Yes
Load balancing by IP address	No	Yes	Yes	Yes
Load balancing by MAC address	No	Yes (used for no-IP/IPX)	Yes	Yes

[a] SLB with one primary and one standby member.

[b] Requires at least one Marvell adapter in the team.

Selecting a Team Type

The following flow chart provides the decision flow when planning for Layer 2 teaming. The primary rationale for teaming is the need for additional network bandwidth and fault tolerance. Teaming offers link aggregation and fault tolerance to meet both of these requirements. Preference teaming should be selected in the following order:

- First choice: Link Aggregation

- Second choice: Generic Trunking
- Third choice: SLB, when using unmanaged switches or switches that do not support the first two choices. If switch fault tolerance is a requirement, either SLB or in-OS switch independent NIC teaming is the only choice.

Figure 12-1 shows a flow chart for determining the team type.

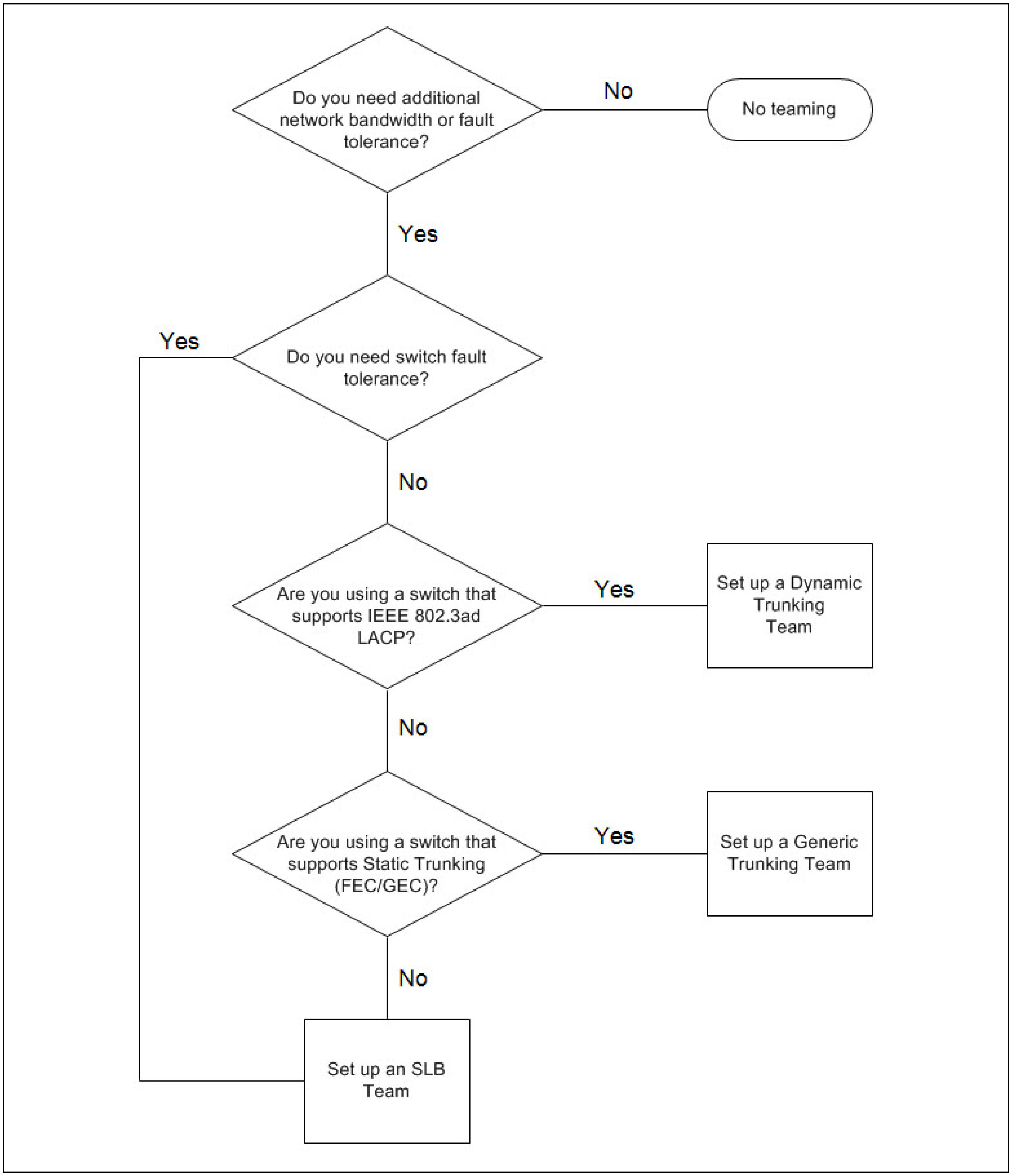

Do you need additional network bandwidth or fault tolerance?
No
No teaming
Yes
Yes
Do you need switch fault tolerance?
No
Are you using a switch that supports IEEE 802.3ad LACP?
Yes
Set up a Dynamic Trunking Team
No
Are you using a switch that supports Static Trunking (FEC/GEC)?
Yes
Set up a Generic Trunking Team
No
Set up an SLB Team

Figure 12-1. Process for Selecting a Team Type

Teaming Mechanisms

This section provides the following information about teaming mechanisms:

- Architecture
- Types of Teams
- Attributes of the Features Associated with Each Type of Team
- Speeds Supported for Each Type of Team

Architecture

The NDIS intermediate driver (see Figure 12-2) operates below protocol stacks such as TCP/IP and IPX and appears as a virtual adapter. This virtual adapter inherits the MAC Address of the first port initialized in the team. A Layer 3 address must also be configured for the virtual adapter. The primary function of the driver is to balance inbound (for SLB) and outbound traffic (for all teaming modes) among the physical adapters installed on the system selected for teaming. The inbound and outbound algorithms are independent and orthogonal to each other. The outbound traffic for a specific session can be assigned to a specific port while its corresponding inbound traffic can be assigned to a different port.

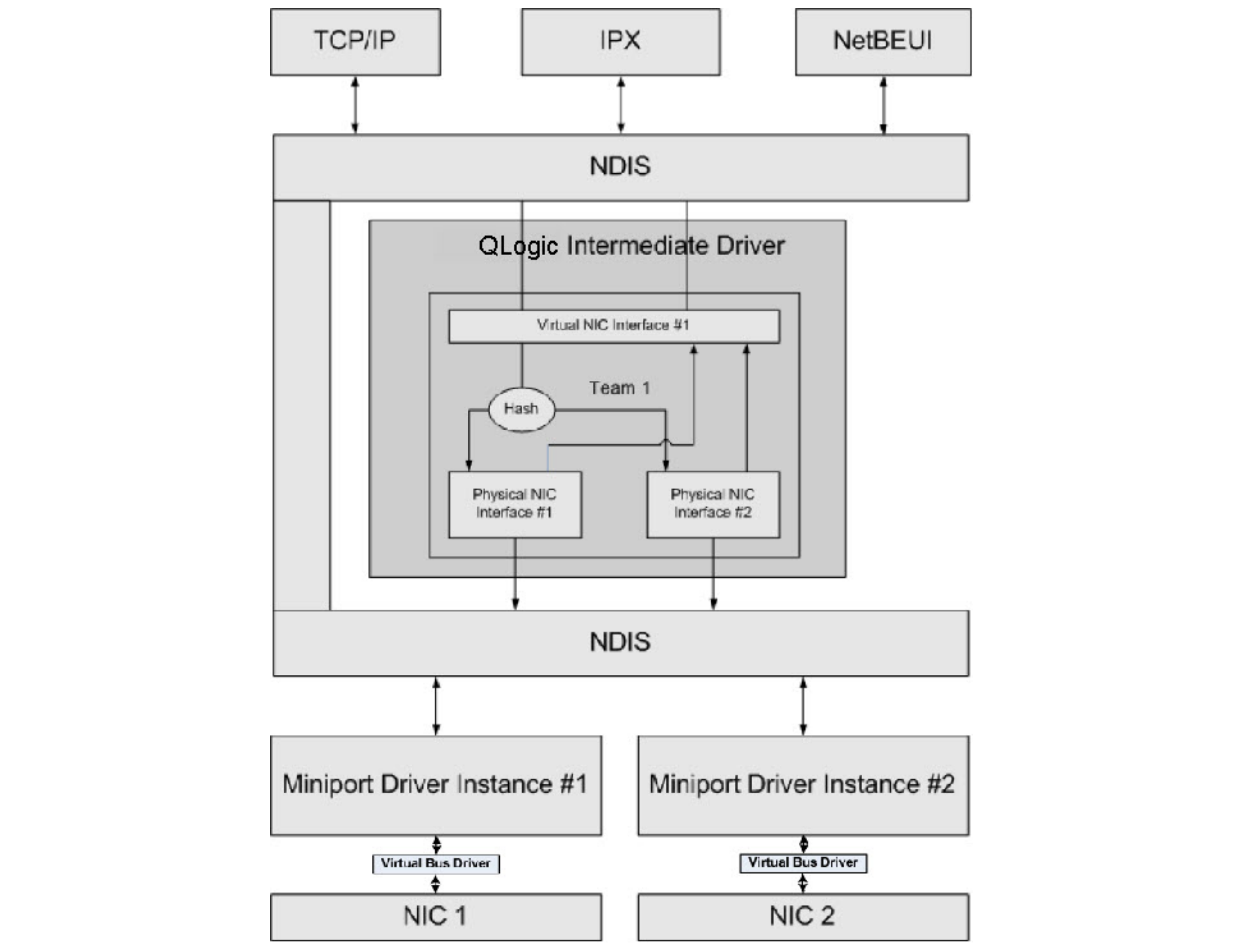

TCP/IP
IPX
NetBEUI
NDIS
QLogic Intermediate Driver
Virtual NIC Interface #1
Team 1
Hash
Physical NIC Interface #1
Physical NIC Interface #2
NDIS
Miniport Driver Instance #1
Miniport Driver Instance #2
Virtual Bus Driver
Virtual Bus Driver
NIC 1
NIC 2

Figure 12-2. Intermediate Driver

Outbound Traffic Flow

The Marvell intermediate driver manages the outbound traffic flow for all teaming modes. For outbound traffic, every packet is first classified into a flow, and then distributed to the selected physical adapter for transmission. The flow classification involves an efficient hash computation over known protocol fields. The resulting hash value is used to index into an Outbound Flow Hash Table. The selected Outbound Flow Hash Entry contains the index of the selected physical adapter responsible for transmitting this flow. The source MAC address of the packets will then be modified to the MAC address of the selected physical adapter. The modified packet is then passed to the selected physical adapter for transmission.

The outbound TCP and UDP packets are classified using Layer 3 and Layer 4 header information. This scheme improves the load distributions for popular Internet protocol services using well-known ports such as HTTP and FTP. Therefore, QLASP performs load balancing on a TCP session basis and not on a packet-by-packet basis.

In the Outbound Flow Hash Entries, statistics counters are also updated after classification. The load-balancing engine uses these counters to periodically distribute the flows across teamed ports. The outbound code path has been designed to achieve best possible concurrency where multiple concurrent accesses to the Outbound Flow Hash Table are allowed.

For protocols other than TCP/IP, the first physical adapter will always be selected for outbound packets. The exception is Address Resolution Protocol (ARP), which is handled differently to achieve inbound load balancing.

Inbound Traffic Flow (SLB Only)

The Marvell intermediate driver manages the inbound traffic flow for the SLB teaming mode. Unlike outbound load balancing, inbound load balancing can only be applied to IP addresses that are located in the same subnet as the load-balancing server. Inbound load balancing exploits a unique characteristic of address resolution protocol (RFC0826), in which each IP host uses its own ARP cache to encapsulate the IP datagram into an Ethernet frame. QLASP carefully manipulates the ARP response to direct each IP host to send the inbound IP packet to the physical adapter that you want. Therefore, inbound load balancing is a plan-ahead scheme based on statistical history of the inbound flows. New connections from a client to the server will always occur over the primary physical adapter (because the ARP Reply generated by the operating system protocol stack will always associate the logical IP address with the MAC address of the primary physical adapter).

Like the outbound case, there is an Inbound Flow Head Hash Table. Each entry inside this table has a singly linked list and each link (Inbound Flow Entries) represents an IP host located in the same subnet.

When an inbound IP Datagram arrives, the appropriate Inbound Flow Head Entry is located by hashing the source IP address of the IP Datagram. Two statistics counters stored in the selected entry are also updated. These counters are used in the same fashion as the outbound counters by the load-balancing engine periodically to reassign the flows to the physical adapter.

On the inbound code path, the Inbound Flow Head Hash Table is also designed to allow concurrent access. The link lists of Inbound Flow Entries are only referenced in the event of processing ARP packets and the periodic load balancing. There is no per packet reference to the Inbound Flow Entries. Even though the link lists are not bounded; the overhead in processing each non-ARP packet is always a constant. The processing of ARP packets, both inbound and outbound, however, depends on the quantity of links inside the corresponding link list.

On the inbound processing path, filtering is also employed to prevent broadcast packets from looping back through the system from other physical adapters.

Protocol Support

ARP and IP/TCP/UDP flows are load balanced. If the packet is an IP protocol only, such as ICMP or IGMP, all data flowing to a specific IP address will go out through the same physical adapter. If the packet uses TCP or UDP for the Layer 4 protocol, the port number is added to the hashing algorithm, so that two separate Layer 4 flows can go out through two separate physical adapters to the same IP address.

For example, assume the client has an IP address of 10.0.0.1. All IGMP and ICMP traffic will go out the same physical adapter because only the IP address is used for the hash. The flow would look something like this:

```
IGMP ------> PhysAdapter1 ------> 10.0.0.1
ICMP ------> PhysAdapter1 ------> 10.0.0.1
```

If the server also sends an TCP and UDP flow to the same 10.0.0.1 address, they can be on the same physical adapter as IGMP and ICMP, or on completely different physical adapters from ICMP and IGMP. The stream may look like this:

```
IGMP ------> PhysAdapter1 ------> 10.0.0.1
ICMP ------> PhysAdapter1 ------> 10.0.0.1
TCP  ------> PhysAdapter1 ------> 10.0.0.1
UDP  ------> PhysAdatper1 ------> 10.0.0.1
```

Or the streams may look like this:

```
IGMP ------> PhysAdapter1 ------> 10.0.0.1
ICMP ------> PhysAdapter1 ------> 10.0.0.1
TCP  ------> PhysAdapter2 ------> 10.0.0.1
UDP  ------> PhysAdatper3 ------> 10.0.0.1
```

The actual assignment between adapters may change over time, but any protocol that is not TCP/UDP based goes over the same physical adapter because only the IP address is used in the hash.

Performance

Modern network interface cards provide many hardware features that reduce CPU utilization by offloading specific CPU intensive operations (see “Teaming and Other Advanced Networking Properties” on page 168). In contrast, the QLASP intermediate driver is a purely software function that must examine every packet received from the protocol stacks and react to its contents before sending it out through a specific physical interface. Though the QLASP driver can process each outgoing packet in near constant time, some applications that may already be CPU bound may suffer if operated over a teamed interface. Such an application may be better suited to take advantage of the failover capabilities of the intermediate driver rather than the load balancing features, or it may operate more efficiently over a single physical adapter that provides a specific hardware feature such as Large Send Offload.

Types of Teams

Team types include switch-independent, switch dependent, and LiveLink.

Switch-Independent

The Marvell Smart Load Balancing type of team allows two to eight physical adapters to operate as a single virtual adapter. The greatest benefit of the SLB type of team is that it operates on any IEEE compliant switch and requires no special configuration.

Smart Load Balancing and Failover

SLB provides for switch-independent, bidirectional, fault-tolerant teaming and load balancing. Switch independence implies that there is no specific support for this function required in the switch, allowing SLB to be compatible with all switches. Under SLB, all adapters in the team have separate MAC addresses. The load-balancing algorithm operates on Layer 3 addresses of the source and destination nodes, which enables SLB to load balance both incoming and outgoing traffic.

The QLASP intermediate driver continually monitors the physical ports in a team for link loss. In the event of link loss on any port, traffic is automatically diverted to other ports in the team. The SLB teaming mode supports switch fault tolerance by allowing teaming across different switches- provided the switches are on the same physical network or broadcast domain.

Network Communications

Key attributes of SLB include:

- Failover mechanism—Link loss detection.
- Load Balancing Algorithm—Inbound and outbound traffic are balanced through a Marvell proprietary mechanism based on Layer 4 flows.
- Outbound Load Balancing using MAC Address—No
- Outbound Load Balancing using IP Address—Yes
- Multivendor Teaming—Supported (must include at least one Marvell Ethernet adapter as a team member).

Applications

The SLB algorithm is most appropriate in home and small business environments where cost is a concern or with commodity switching equipment. SLB teaming works with unmanaged Layer 2 switches and is a cost-effective way of getting redundancy and link aggregation at the server. Smart Load Balancing also supports teaming physical adapters with differing link capabilities. In addition, SLB is recommended when switch fault tolerance with teaming is required.

Configuration Recommendations

SLB supports connecting the teamed ports to hubs and switches if they are on the same broadcast domain. It does not support connecting to a router or Layer 3 switches because the ports must be on the same subnet.

Switch-Dependent

Generic Static Trunking

This mode supports a variety of environments where the adapter link partners are statically configured to support a proprietary trunking mechanism. This mode could be used to support the Lucent Open Trunk, the Cisco Fast EtherChannel (FEC), and the Cisco Gigabit EtherChannel (GEC). In the static mode, as in generic link aggregation, the switch administrator needs to assign the ports to the team, as there is no exchange of the link aggregation control protocol (LACP) frame.

With this mode, all adapters in the team are configured to receive packets for the same MAC address. Trunking operates on Layer 2 addresses and supports load balancing and failover for both inbound and outbound traffic.

The attached switch must support the appropriate trunking scheme for this mode of operation. Both the QLASP and the switch continually monitor their ports for link loss. In the event of link loss on any port, traffic is automatically diverted to other ports in the team.

Network Communications

The following are the key attributes of Generic Static Trunking:

- Failover mechanism—Link loss detection
- Load Balancing Algorithm—Outbound traffic is balanced through Marvell proprietary mechanism-based Layer 4 flows. Inbound traffic is balanced according to a switch specific mechanism.
- Outbound Load Balancing using MAC Address—No
- Outbound Load Balancing using IP Address—Yes
- Multivendor teaming—Supported (Must include at least one Marvell Ethernet adapter as a team member)

Applications

Generic trunking works with switches that support Cisco Fast EtherChannel, Cisco Gigabit EtherChannel, Extreme Networks Load Sharing and Bay Networks or IEEE 802.3ad Link Aggregation static mode. Because load balancing is implemented on Layer 2 addresses, all higher protocols such as IP, IPX, and NetBEUI are supported. Therefore, this is the recommended teaming mode when the switch supports generic trunking modes over SLB.

Configuration Recommendations

Static trunking supports connecting the teamed ports to switches if they are on the same broadcast domain and support generic trunking. It does not support connecting to a router or Layer 3 switches because the ports must be on the same subnet.

Dynamic Trunking (IEEE 802.3ad Link Aggregation)

This mode supports link aggregation through static and dynamic configuration through the link aggregation control protocol (LACP). With this mode, all adapters in the team are configured to receive packets for the same MAC address. The MAC address of the first adapter in the team is used and cannot be substituted for a different MAC address. The QLASP driver determines the load-balancing scheme for outbound packets, using Layer 4 protocols previously discussed, whereas the team’s link partner determines the load-balancing scheme for inbound packets. Because the load balancing is implemented on Layer 2, all higher protocols such as IP, IPX, and NetBEUI are supported. The attached switch must support the 802.3ad Link Aggregation standard for this mode of operation. The switch manages the inbound traffic to the adapter while the QLASP manages the outbound traffic. Both the QLASP and the switch continually monitor their ports for link loss. In the event of link loss on any port, traffic is automatically diverted to other ports in the team.

Network Communications

The following are the key attributes of dynamic trunking:

- Failover mechanism—Link loss detection
- Load Balancing Algorithm—Outbound traffic is balanced through a Marvell proprietary mechanism based on Layer 4 flows. Inbound traffic is balanced according to a switch specific mechanism.
- Outbound Load Balancing using MAC Address—No
- Outbound Load Balancing using IP Address—Yes
- Multivendor teaming—Supported (Must include at least one Marvell Ethernet adapter as a team member)

Applications

Dynamic trunking works with switches that support IEEE 802.3ad Link Aggregation dynamic mode using LACP. Inbound load balancing is switch dependent. In general, the switch traffic is load balanced based on Layer 2 addresses. In this case, all network protocols such as IP, IPX, and NetBEUI are load balanced. Therefore, this is the recommended teaming mode when the switch supports LACP, except when switch fault tolerance is required. SLB is the only teaming mode that supports switch fault tolerance.

Configuration Recommendations

Dynamic trunking supports connecting the teamed ports to switches as long as they are on the same broadcast domain and supports IEEE 802.3ad LACP trunking. It does not support connecting to a router or Layer 3 switches because the ports must be on the same subnet.

LiveLink

LiveLink is a feature of QLASP that is available for the Smart Load Balancing (SLB) and SLB (Auto-Fallback Disable) types of teaming. The purpose of LiveLink is to detect link loss beyond the switch and to route traffic only through team members that have a live link. This function is accomplished though the teaming software. The teaming software periodically probes (issues a link packet from each team member) one or more specified target network devices. The probe target(s) responds when it receives the link packet. If a team member does not detect the response within a specified amount of time, this indicates that the link has been lost, and the teaming software discontinues passing traffic through that team member. Later, if that team member begins to detect a response from a probe target, this indicates that the link has been restored, and the teaming software automatically resumes passing traffic through that team member. LiveLink works only with TCP/IP.

LiveLink functionality is supported in both 32-bit and 64-bit Windows operating systems. For similar functionality in Linux operating systems, see the Channel Bonding information in your Red Hat documentation.

Attributes of the Features Associated with Each Type of Team

The attributes of the features associated with each type of team are summarized in Table 12-5.

Table 12-5. Teaming Attributes

Feature	Attribute
Smart Load Balancing	
User interface	QCS CLI or QCC GUI
Quantity of teams	Maximum 16
Quantity of adapters per team	Maximum 16
Hot replace	Yes
Hot add	Yes
Hot remove	Yes
Link speed support	Different speeds
Frame protocol	IP
Incoming packet management	QLASP
Outgoing packet management	QLASP
LiveLink support	Yes
Failover event	Loss of link
Failover time	<500ms
Fallback time	1.5s (approximate) [a]
MAC address	Different
Multivendor teaming	Yes
Generic Trunking	
User interface	QCS CLI or QCC GUI
Quantity of teams	Maximum 16
Quantity of adapters per team	Maximum 16
Hot replace	Yes

Table 12-5. Teaming Attributes (Continued)

Feature	Attribute
Hot add	Yes
Hot remove	Yes
Link speed support	Different speeds [b]
Frame protocol	All
Incoming packet management	Switch
Outgoing packet management	QLASP
Failover event	Loss of link only
Failover time	< 500ms
Fallback time	1.5s (approximate) [a]
MAC address	Same for all adapters
Multivendor teaming	Yes
Dynamic Trunking	
User interface	QCS CLI or QCC GUI
Quantity of teams	Maximum 16
Quantity of adapters per team	Maximum 16
Hot replace	Yes
Hot add	Yes
Hot remove	Yes
Link speed support	Different speeds
Frame protocol	All
Incoming packet management	Switch
Outgoing packet management	QLASP
Failover event	Loss of link only
Failover time	< 500ms
Fallback time	1.5s (approximate) [a]
MAC address	Same for all adapters
Multivendor teaming	Yes

[a] Make sure that Port Fast or Edge Port is enabled.

[b] Some switches require matching link speeds to correctly negotiate between trunk connections.

Speeds Supported for Each Type of Team

The various link speeds that are supported for each type of team are listed in Table 12-6. Mixed speed refers to the capability of teaming adapters that are running at different link speeds.

Table 12-6. Link Speeds in Teaming

Type of Team	Link Speed	Traffic Direction	Speed Support
SLB	10/100/1000/10000	Incoming and outgoing	Mixed speed
FEC	100	Incoming and outgoing	Same speed
GEC	1000	Incoming and outgoing	Same speed
IEEE 802.3ad	10/100/1000/10000	Incoming and outgoing	Mixed speed

Teaming and Other Advanced Networking Properties

This section covers the following teaming and advanced networking properties:

- Checksum Offload
- IEEE 802.1p QoS Tagging
- Large Send Offload
- Jumbo Frames
- IEEE 802.1Q VLANs
- Wake on LAN
- Preboot Execution Environment

Before creating a team, adding or removing team members, or changing advanced settings of a team member, make sure each team member has been configured similarly. Settings to check include VLANs and QoS Packet Tagging, Jumbo Frames, and the various offloads. Advanced adapter properties and teaming support are listed in Table 12-7.

Table 12-7. Advanced Adapter Properties and Teaming Support

Adapter Properties	Supported by Teaming Virtual Adapter
Checksum Offload	Yes
IEEE 802.1p QoS Tagging	No
Large Send Offload	Yes [a]
Jumbo Frames	Yes [b]
IEEE 802.1Q VLANs	Yes [c]
Wake on LAN	No [d]
Preboot Execution environment (PXE)	Yes [e]

[a] All adapters on the team must support this feature. Some adapters may not support this feature if ASF/IPMI is also enabled.

[b] Must be supported by all adapters in the team.

[c] Only for Marvell adapters.

[d] See Wake on LAN.

[e] As a PXE sever only, not as a client.

A team does not necessarily inherit adapter properties; rather various properties depend on the specific capability. For instance, an example would be flow control, which is a physical adapter property and has nothing to do with QLASP, and will be enabled on a specific adapter if the miniport driver for that adapter has flow control enabled.

NOTE

All adapters on the team must support the property listed in Table 12-7 in order for the team to support the property.

Checksum Offload

Checksum Offload is a property of the Marvell network adapters that allows the TCP/IP/UDP checksums for send and receive traffic to be calculated by the adapter hardware rather than by the host CPU. In high-traffic situations, this can allow a system to handle more connections more efficiently than if the host CPU were forced to calculate the checksums. This property is inherently a hardware property and would not benefit from a software-only implementation. An adapter that supports Checksum Offload advertises this capability to the operating system so that the checksum does not need to be calculated in the protocol stack. Checksum Offload is only supported for IPv4 at this time.

IEEE 802.1p QoS Tagging

The IEEE 802.1p standard includes a 3-bit field (supporting a maximum of 8 priority levels), which allows for traffic prioritization.

Large Send Offload

Large Send Offload (LSO) is a feature provided by Marvell network adapters that prevents an upper level protocol such as TCP from breaking a large data packet into a series of smaller packets with headers appended to them. The protocol stack need only generate a single header for a data packet as large as 64 KB, and the adapter hardware breaks the data buffer into appropriately-sized Ethernet frames with the correctly sequenced header (based on the single header originally provided).

Jumbo Frames

The use of jumbo frames was originally proposed by Alteon Networks, Inc. in 1998 and increased the maximum size of an Ethernet frame to a maximum size of 9600 bytes. Though never formally adopted by the IEEE 802.3 Working Group, support for jumbo frames has been implemented in Marvell 57*xx* and 57*xxx* adapters. The QLASP intermediate driver supports jumbo frames, provided that all of the physical adapters in the team also support jumbo frames and the same size is set on all adapters in the team.

IEEE 802.1Q VLANs

In 1998, the IEEE approved the 802.3ac standard, which defines frame format extensions to support Virtual Bridged Local Area Network tagging on Ethernet networks as specified in the IEEE 802.1Q specification. The VLAN protocol permits insertion of a tag into an Ethernet frame to identify the VLAN to which a frame belongs. If present, the 4-byte VLAN tag is inserted into the Ethernet frame between the source MAC address and the length/type field. The first 2-bytes of the VLAN tag consist of the IEEE 802.1Q tag type, whereas the second 2 bytes include a user priority field and the VLAN identifier (VID). Virtual LANs (VLANs) allow the user to split the physical LAN into logical subparts. Each defined VLAN behaves as its own separate network, with its traffic and broadcasts isolated from the others, thus increasing bandwidth efficiency within each logical group. VLANs also enable the administrator to enforce appropriate security and quality of service (QoS) policies. The QLASP supports the creation of 64 VLANs per team or adapter: 63 tagged and 1 untagged. The operating system and system resources, however, limit the actual quantity of VLANs. VLAN support is provided according to IEEE 802.1Q and is supported in a teaming environment as well as on a single adapter. Note that VLANs are supported only with homogeneous teaming and not in a multivendor teaming environment. The QLASP intermediate driver supports VLAN tagging. One or more VLANs may be bound to a single instance of the intermediate driver.

Wake on LAN

Wake on LAN (WoL) is a feature that allows a system to be awakened from a sleep state by the arrival of a specific packet over the Ethernet interface. Because a Virtual Adapter is implemented as a software only device, it lacks the hardware features to implement Wake on LAN and cannot be enabled to wake the system from a sleeping state through the virtual adapter. The physical adapters, however, support this property, even when the adapter is part of a team.

NOTE

WoL is only supported on one physical port (Port 1) for the following adapters:

- 957810A1006DC (N20KJ)
- 957810A1006DLPC (Y40PH)

Preboot Execution Environment

The preboot execution environment (PXE) allows a system to boot from an operating system image over the network. By definition, PXE is invoked before an operating system is loaded, so there is no opportunity for the driver to load and enable a team. As a result, teaming is not supported as a PXE client, though a physical adapter that participates in a team when the operating system is loaded may be used as a PXE client. Whereas a teamed adapter cannot be used as a PXE client, it can be used for a PXE server, which provides operating system images to PXE clients using a combination of dynamic host control protocol (DHCP) and the trivial file transfer protocol (TFTP). Both of these protocols operate over IP and are supported by all teaming modes.

General Network Considerations

General network considerations include:

- Teaming with Microsoft Virtual Server 2005
- Teaming Across Switches
- Spanning Tree Algorithm
- Layer 3 Routing and Switching
- Teaming with Hubs (for Troubleshooting Purposes Only)
- Teaming with Microsoft Network Load Balancing

Teaming with Microsoft Virtual Server 2005

The only supported QLASP team configuration when using Microsoft Virtual Server 2005 is with a Smart Load Balancing team-type consisting of a single primary Marvell adapter and a standby Marvell adapter. Make sure to unbind or deselect “Virtual Machine Network Services” from each team member prior to creating a team and prior to creating virtual networks with Microsoft Virtual Server. Additionally, create a virtual network in this software and subsequently bound to the virtual adapter created by a team. Directly binding a guest operating system to a team virtual adapter may not render the results that you want.

NOTE

Microsoft recommends using their in-OS NIC teaming service instead of any adapter vendor-proprietary NIC teaming driver on Windows Server 2012 and later. Marvell's NIC teaming driver is not supported on Windows Server 2016 and later.

Teaming Across Switches

SLB teaming can be configured across switches. The switches, however, must be connected together. Generic Trunking and Link Aggregation do not work across switches because each of these implementations requires that all physical adapters in a team share the same Ethernet MAC address. It is important to note that SLB can only detect the loss of link between the ports in the team and their immediate link partner. SLB has no way of reacting to other hardware failures in the switches and cannot detect loss of link on other ports.

Switch-Link Fault Tolerance

The figures in this section describe the operation of an SLB team in a switch fault tolerant configuration. Marvell shows the mapping of the ping request and ping replies in an SLB team with two active members. All servers (Blue, Gray, and Red) have a continuous ping to each other. These scenarios describe the behavior of teaming across the two switches and the importance of the interconnect link.

- Figure 12-3 is a setup without the interconnect cable in place between the two switches.
- Figure 12-4 has the interconnect cable in place.
- Figure 12-5 is an example of a failover event with the Interconnect cable in place.

The figures show the secondary team member sending the ICMP echo requests (yellow arrows) while the primary team member receives the respective ICMP echo replies (blue arrows). This send-receive illustrates a key characteristic of the teaming software. The load balancing algorithms do not synchronize how frames are load balanced when sent or received. Frames for a specific conversation can go out and be received on different interfaces in the team, which is true for all types of teaming supported by Marvell. Therefore, an interconnect link must be provided between the switches that connect to ports in the same team.

In the configuration without the interconnect, an ICMP Request from Blue to Gray goes out port 82:83 destined for Gray port 5E:CA, but the Top Switch has no way to send it there because it cannot go along the 5E:C9 port on Gray. A similar scenario occurs when Gray attempts to ping Blue. An ICMP Request goes out on 5E:C9 destined for Blue 82:82, but cannot get there. Top Switch does not have an entry for 82:82 in its CAM table because there is no interconnect between the two switches. Pings, however, flow between Red and Blue and between Red and Gray.

Furthermore, a failover event would cause additional loss of connectivity. Consider a cable disconnect on the Top Switch port 4. In this case, Gray would send the ICMP Request to Red 49:C9, but because the Bottom Switch has no entry for 49:C9 in its CAM Table, the frame is flooded to all its ports but cannot find a way to get to 49:C9.

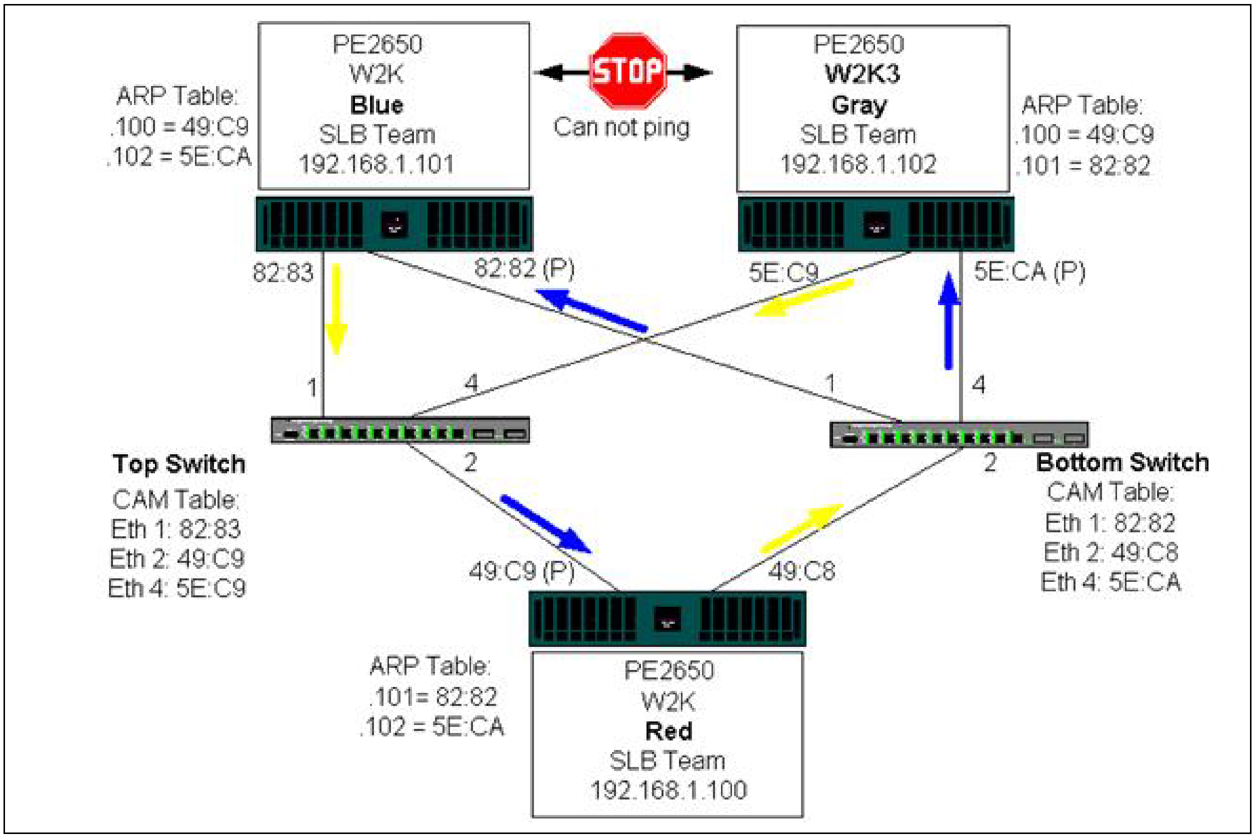


Figure 12-3. Teaming Across Switches Without an Inter-Switch Link

The addition of a link between the switches allows traffic from and to Blue and Gray to reach each other without any problems. Note the additional entries in the CAM table for both switches. The link interconnect is critical for the proper operation of the team. As a result, Marvell highly advises that you have a link aggregation trunk to interconnect the two switches to ensure high availability for the connection.

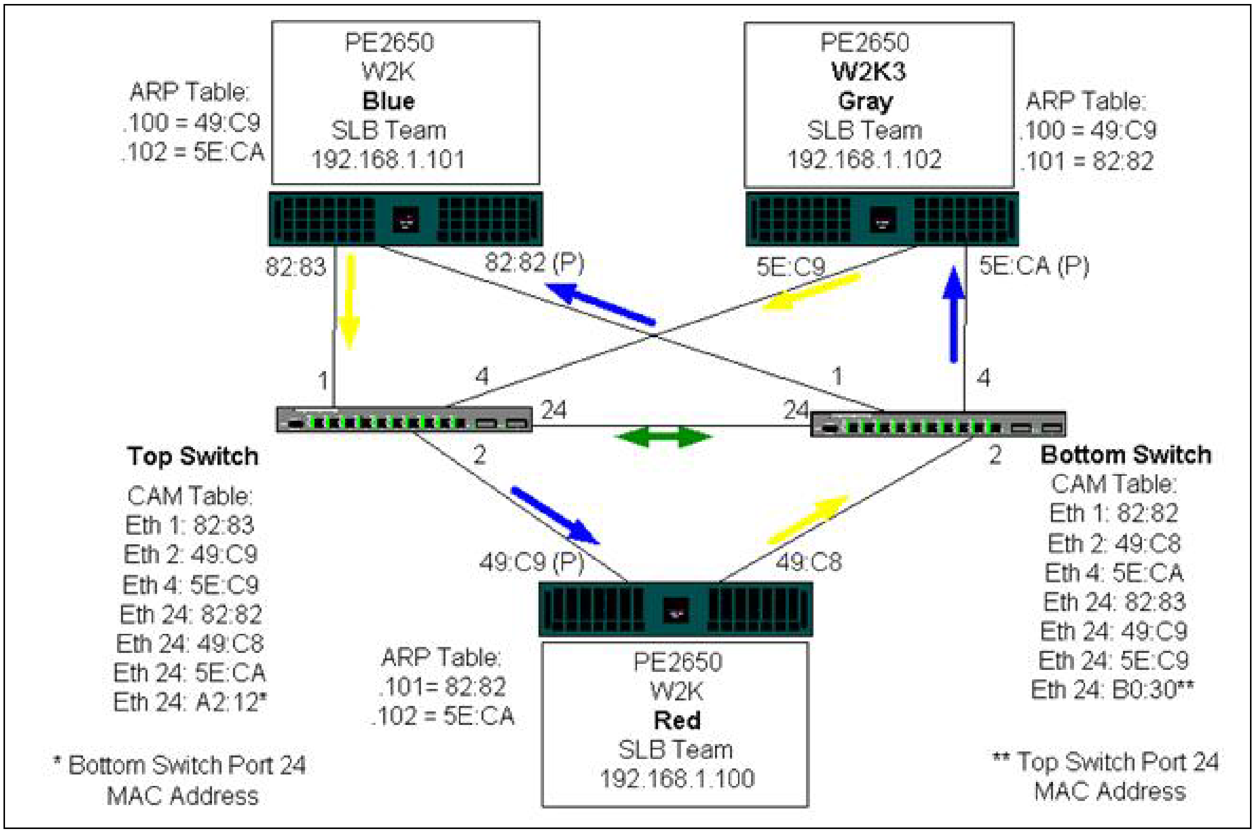


Figure 12-4. Teaming Across Switches with Interconnect

Figure 12-5 represents a failover event in which the cable is unplugged on the Top Switch port 4. This event is a successful failover with all stations pinging each other without loss of connectivity.

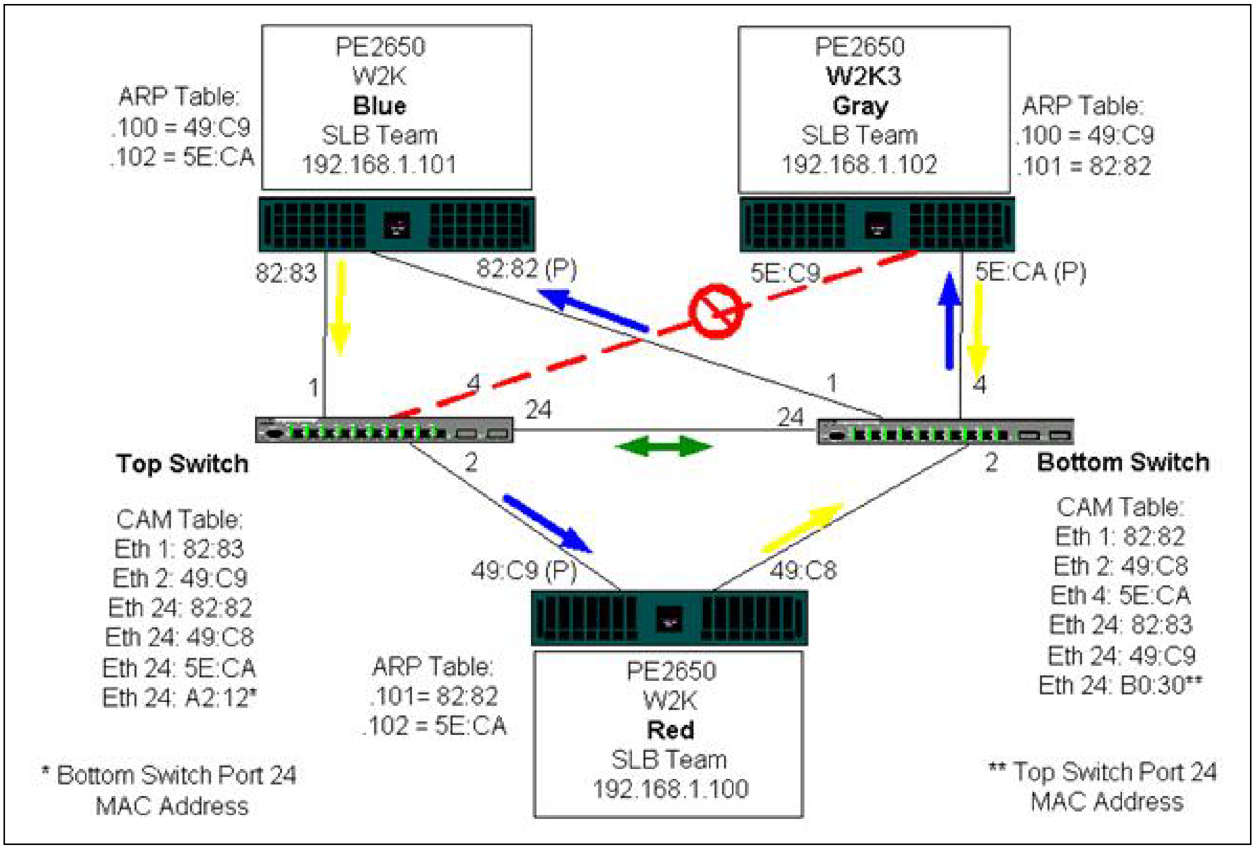


Figure 12-5. Failover Event

Spanning Tree Algorithm

In Ethernet networks, only one active path may exist between any two bridges or switches. Multiple active paths between switches can cause loops in the network. When loops occur, some switches recognize stations on both sides of the switch. This situation causes the forwarding algorithm to malfunction allowing duplicate frames to be forwarded. Spanning tree algorithms provide path redundancy by defining a tree that spans all of the switches in an extended network and then forces specific redundant data paths into a standby (blocked) state. At regular intervals, the switches in the network send and receive spanning tree packets that they use to identify the path. If one network segment becomes unreachable, or if spanning tree costs change, the spanning tree algorithm reconfigures the spanning tree topology and re-establishes the link by activating the standby path. Spanning tree operation is transparent to end stations, which do not detect whether they are connected to a single LAN segment or a switched LAN of multiple segments.

Spanning tree protocol (STP) is a Layer 2 protocol designed to run on bridges and switches. The specification for STP is defined in IEEE 802.1d. The main purpose of STP is to ensure that you do not run into a loop situation when you have redundant paths in your network. STP detects and disables network loops and provides backup links between switches or bridges. It allows the device to interact with other STP compliant devices in your network to ensure that only one path exists between any two stations on the network.

After a stable network topology has been established, all bridges listen for hello BPDUs (bridge protocol data units) transmitted from the root bridge. If a bridge does not get a hello BPDU after a predefined interval (Max Age), the bridge assumes that the link to the root bridge is down. This bridge then initiates negotiations with other bridges to reconfigure the network to re-establish a valid network topology. The process to create a new topology can take up to 50 seconds. During this time, end-to-end communications are interrupted.

Marvell does not recommend the use of spanning tree for ports that are connected to end stations, because by definition, an end station does not create a loop within an Ethernet segment. Additionally, when a teamed adapter is connected to a port with spanning tree enabled, users may experience unexpected connectivity problems. For example, consider a teamed adapter that has a lost link on one of its physical adapters. If the physical adapter were to be reconnected (also known as fallback), the intermediate driver would detect that the link has been reestablished and would begin to pass traffic through the port. Traffic would be lost if the port was temporarily blocked by the (STP).

This section provides details of:

- Topology Change Notice (TCN)
- Port Fast and Edge Port

Topology Change Notice (TCN)

A bridge or switch creates a forwarding table of MAC addresses and port numbers by learning the source MAC address that received on a specific port. The table is used to forward frames to a specific port rather than flooding the frame to all ports. The typical maximum aging time of entries in the table is 5 minutes. Only when a host has been silent for 5 minutes would its entry be removed from the table. It is sometimes beneficial to reduce the aging time. One example is when a forwarding link goes to blocking and a different link goes from blocking to forwarding. This change could take up to 50 seconds. At the end of the STP re-calculation a new path would be available for communications between end stations. However, because the forwarding table would still have entries based on the old topology, communications may not be reestablished until after 5 minutes when the affected ports entries are removed from the table. Traffic would then be flooded to all ports and re-learned. In this case it is beneficial to reduce the aging time. This reduction is the purpose of a topology change notice (TCN) BPDU. The TCN is sent from the affected bridge or switch to the root bridge/switch. As soon as a bridge/switch detects a topology change (a link going down or a port going to forwarding) it sends a TCN to the root bridge through its root port. The root bridge then advertises a BPDU with a topology change to the entire network.The advertisement causes every bridge to reduce the MAC table aging time to 15 seconds for a specified amount of time. The time reduction allows the switch to re-learn the MAC addresses as soon as STP re-converges.

Topology change notice BPDUs are sent when a port that was forwarding changes to blocking or transitions to forwarding. A TCN BPDU does not initiate an STP recalculation. It only affects the aging time of the forwarding table entries in the switch.It will not change the topology of the network or create loops. End nodes such as servers or clients trigger a topology change when they power off and then power back on.

Port Fast and Edge Port

To reduce the effect of TCNs on the network (for example, increasing flooding on switch ports), end nodes that are powered on and off frequently should use the Port Fast or Edge Port setting on the switch port to which they are attached. Port Fast or Edge Port is a command that is applied to specific ports and has the following effects:

- Ports coming from link down to link up will be put in the forwarding STP mode, instead of going from listening to learning and then to forwarding. STP is still running on these ports.
- The switch does not generate a topology change notice when the port is going up or down.

Layer 3 Routing and Switching

The switch that the teamed ports are connected to must not be a Layer 3 switch or router. The ports in the team must be in the same network.

Teaming with Hubs (for Troubleshooting Purposes Only)

SLB teaming can be used with 10 and 100 hubs, but Marvell recommends using it only for troubleshooting purposes, such as connecting a network analyzer in the event that switch port mirroring is not an option.

Hub teaming information includes:

- Hub Usage in Teaming Network Configurations
- SLB Teams
- SLB Team Connected to a Single Hub
- Generic and Dynamic Trunking (FEC/GEC/IEEE 802.3ad)

Hub Usage in Teaming Network Configurations

Although the use of hubs in network topologies is functional in some situations, it is important to consider the throughput ramifications when doing so. Network hubs have a maximum of 100Mbps half-duplex link speed, which severely degrades performance in either a gigabit or 100Mbps switched-network configuration. Hub bandwidth is shared among all connected devices. As a result, when more devices are connected to the hub, the bandwidth available to any single device connected to the hub is reduced in direct proportion to the quantity of devices connected to the hub.

Marvell does not recommend that you connect team members to hubs; only switches should be used to connect to teamed ports. An SLB team, however, can be connected directly to a hub for troubleshooting purposes. Other team types can result in a loss of connectivity if specific failures occur and should not be used with hubs.

SLB Teams

SLB teams are the only teaming type not dependent on switch configuration. The server intermediate driver handles the load balancing and fault tolerance mechanisms with no assistance from the switch. These elements of SLB make it the only team type that maintains failover and fallback characteristics when team ports are connected directly to a hub.

SLB Team Connected to a Single Hub

SLB teams configured as shown in Figure 12-6 maintain their fault tolerance properties. Either server connection could potentially fail, and network functionality is maintained. Clients could be connected directly to the hub, and fault tolerance would still be maintained; server performance, however, would be degraded.

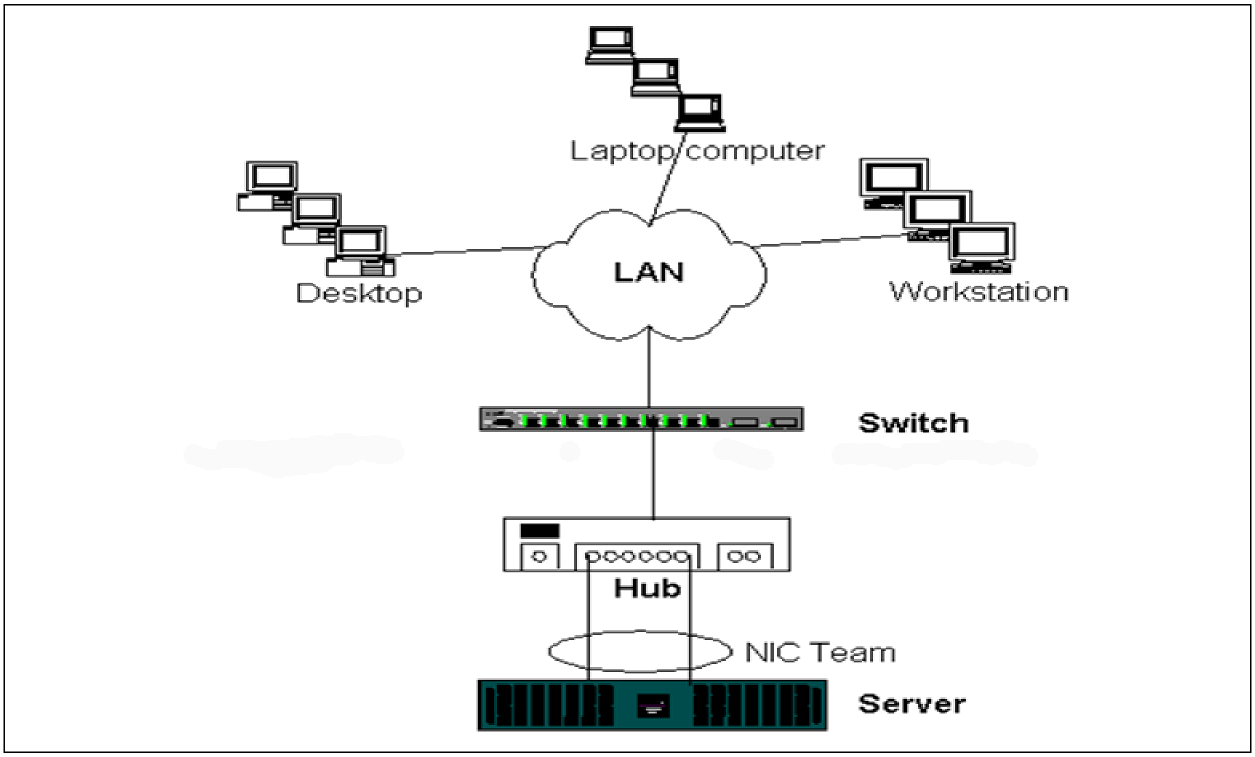


Figure 12-6. Team Connected to a Single Hub

Generic and Dynamic Trunking (FEC/GEC/IEEE 802.3ad)

FEC, GEC, and IEEE 802.3ad teams cannot be connected to any hub configuration. These team types must be connected to a switch that has also been configured for this team type.

Teaming with Microsoft Network Load Balancing

Teaming *does not* work in Microsoft's Network Load Balancing unicast mode, only in multicast mode. Due to the mechanism used by the Network Load Balancing service, the recommended teaming configuration in this environment is Failover (SLB with a standby NIC) as load balancing is managed by Network Load Balancing.

Application Considerations

Application considerations covered:

- Teaming and Clustering
- Teaming and Network Backup

Teaming and Clustering

Teaming and clustering information includes:

- Microsoft Cluster Software
- High-Performance Computing Cluster
- Oracle

Microsoft Cluster Software

Dell Server cluster solutions integrate Microsoft Cluster Services (MSCS) with PowerVault™ SCSI or Dell and EMC Fibre Channel-based storage, Dell servers, storage adapters, storage switches and network adapters to provide high-availability (HA) solutions. HA clustering supports all adapters qualified on a supported Dell server.

In each cluster node, Marvell strongly recommends that customers install at least two network adapters (on-board adapters are acceptable). These interfaces serve two purposes:

- One adapter is used exclusively for intra-cluster *heartbeat* communications. This adapter is referred to as the *private adapter* and usually resides on a separate private subnetwork.
- The other adapter is used for client communications and is referred to as the *public adapter*.

Multiple adapters may be used for each of these purposes: private, intracluster communications and public, external client communications. All Marvell teaming modes are supported with Microsoft Cluster Software for the public adapter only. Private network adapter teaming is not supported. Microsoft indicates that the use of teaming on the private interconnect of a server cluster is not supported because of delays that could possibly occur in the transmission and receipt of heartbeat packets between the nodes. For best results, when you want redundancy for the private interconnect, disable teaming and use the available ports to form a second private interconnect. This interconnect achieves the same end result and provides dual, robust communication paths for the nodes to communicate over.

For teaming in a clustered environment, Marvell recommends that customers use the same brand of adapters.

Figure 12-7 shows a two-node Fibre-Channel cluster with three network interfaces per cluster node: one private and two public. On each node, the two public adapters are teamed, and the private adapter is not. Teaming is supported across the same switch or across two switches. Figure 12-8 on page 184 shows the same two-node Fibre-Channel cluster in this configuration.

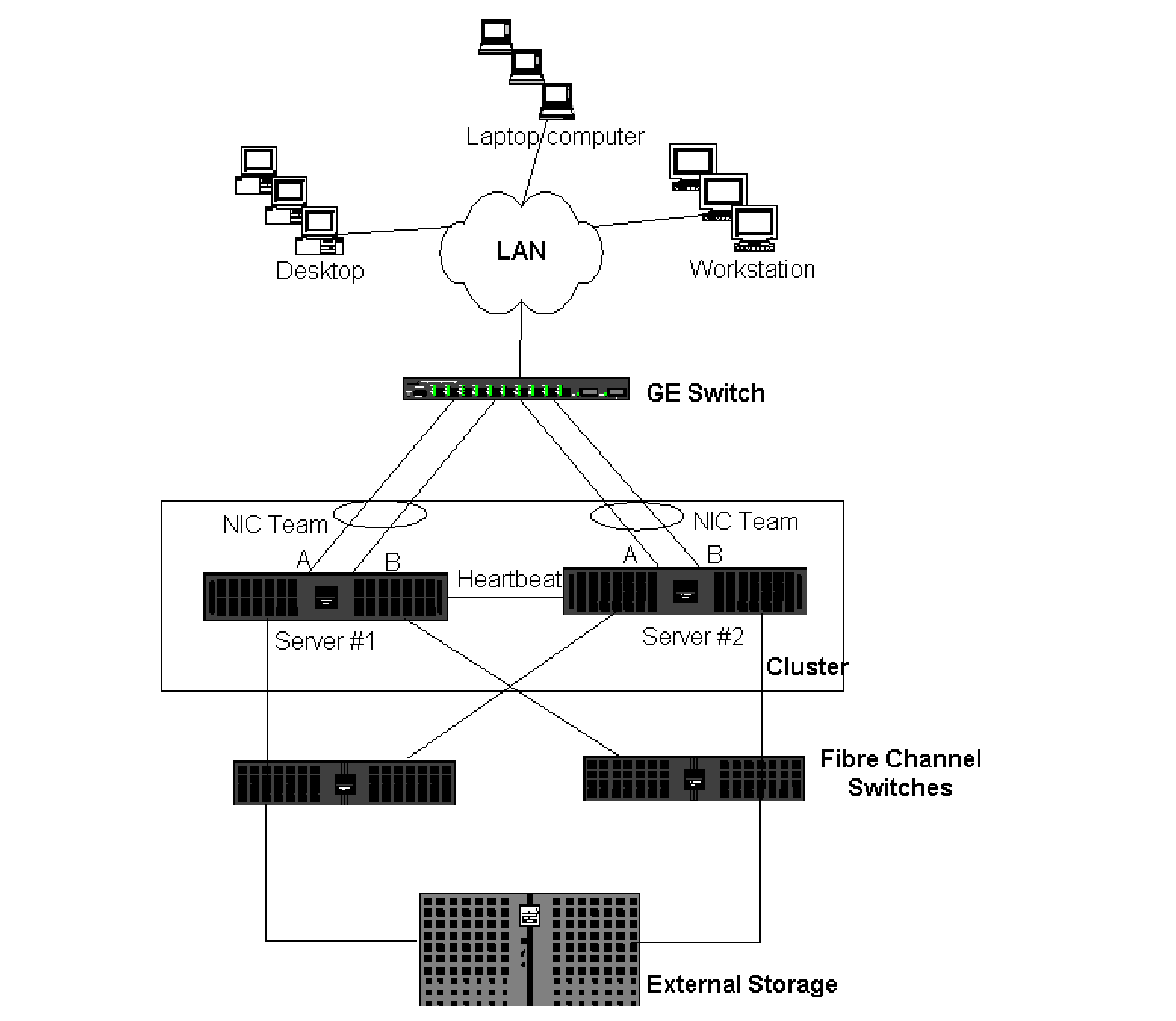


Figure 12-7. Clustering with Teaming Across One Switch

NOTE

Microsoft Network Load Balancing is not supported with Microsoft Cluster Software.

High-Performance Computing Cluster

Gigabit Ethernet is typically used for the following purposes in high-performance computing cluster (HPCC) applications:

- **Inter-process communications (IPC)**: For applications that do not require low-latency, high-bandwidth interconnects (such as Myrinet™ or InfiniBand®), Gigabit Ethernet can be used for communication between the compute nodes.
- **I/O**: Ethernet can be used for file sharing and serving the data to the compute nodes using an NFS server or using parallel file systems such as PVFS.
- **Management and administration**: Ethernet is used for out-of-band (Dell Embedded Remote Access [ERA]) and in-band (Dell OpenManage™ Server Administrator [OMSA]) management of the cluster nodes. It can also be used for job scheduling and monitoring.

In Dell's current HPCC offerings, only one of the on-board adapters is used. If Myrinet or InfiniBand is present, this adapter serves I/O and administration purposes; otherwise, it is also responsible for IPC. In case of an adapter failure, the administrator can use the Felix[1] package to easily configure the second (standby) adapter. Adapter teaming on the host side is neither tested nor supported in HPCC.

Advanced Features

PXE is used extensively for the deployment of the cluster (installation and recovery of compute nodes). Teaming is typically not used on the host side and it is not a part of the Marvell standard offering. Link aggregation is commonly used between switches, especially for large configurations. Jumbo frames, although not a part of the Marvell standard offering, may provide performance improvement for some applications due to reduced CPU overhead.

[1] The 32-bit HPCC configurations from Dell come with the Felix 3.1 Deployment solution stack. Felix is a collaborative effort between MPI Software Technologies Inc. (MSTI) and Dell.

Oracle

In the Marvell Oracle® solution stacks, Marvell supports adapter teaming in both the private network (interconnect between Real Application Cluster [RAC] nodes) and public network with clients or the application layer above the database layer, as shown in Figure 12-8.

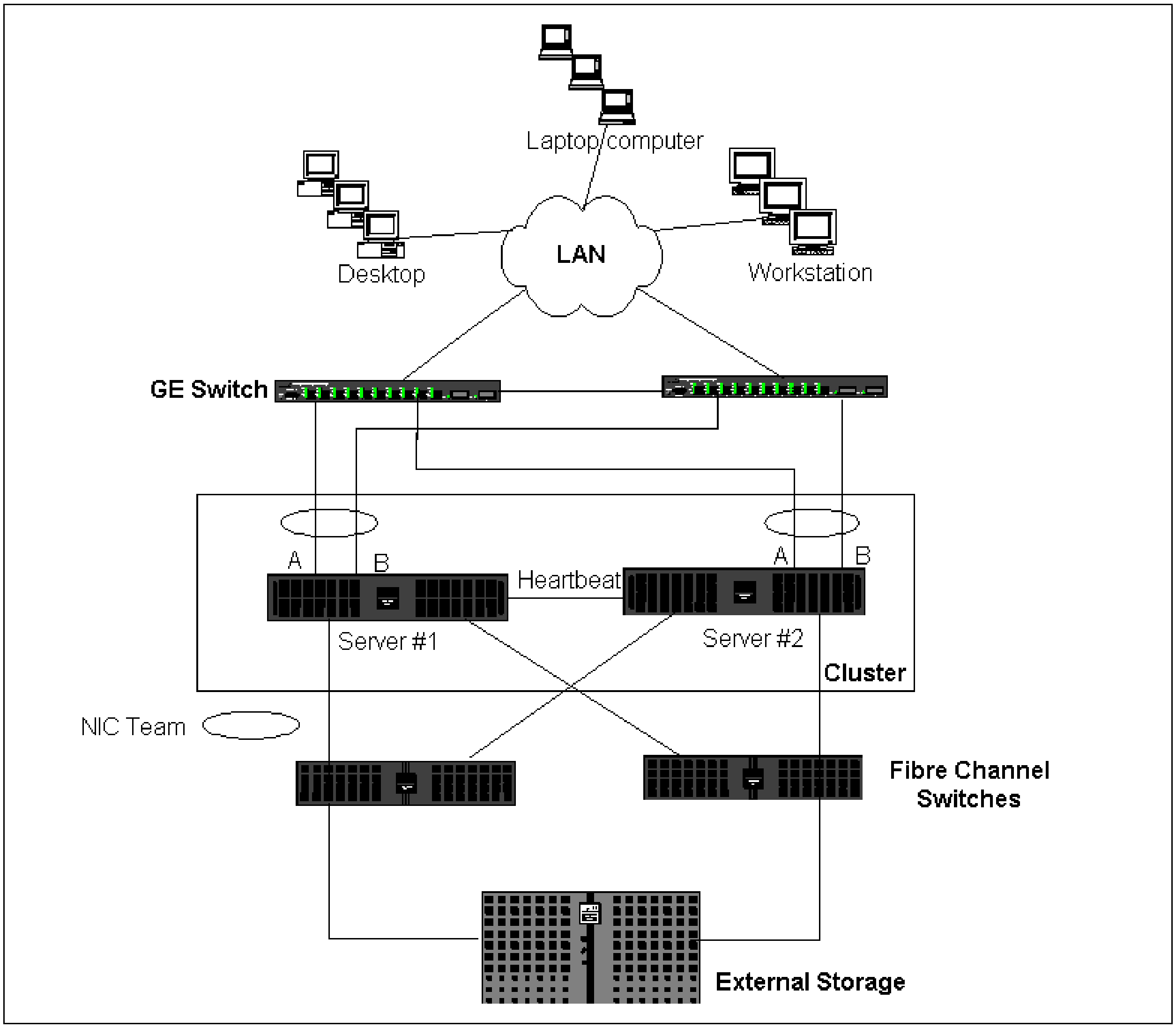


Figure 12-8. Clustering with Teaming Across Two Switches

Teaming and Network Backup

When you perform network backups in a nonteamed environment, overall throughput on a backup server adapter can be easily impacted due to excessive traffic and adapter overloading. Depending on the quantity of backup servers, data streams, and tape drive speed, backup traffic can easily consume a high percentage of the network link bandwidth, thus impacting production data and tape backup performance. Network backups usually consist of a dedicated backup server running with tape backup software such as NetBackup™ or Backup Exec™. Attached to the backup server is either a direct SCSI tape backup unit or a tape library connected through a fiber channel storage area network (SAN). Systems that are backed up over the network are typically called *clients* or *remote servers* and usually have a tape backup software agent installed. Figure 12-9 shows a typical 1Gbps nonteamed network environment with tape backup implementation.

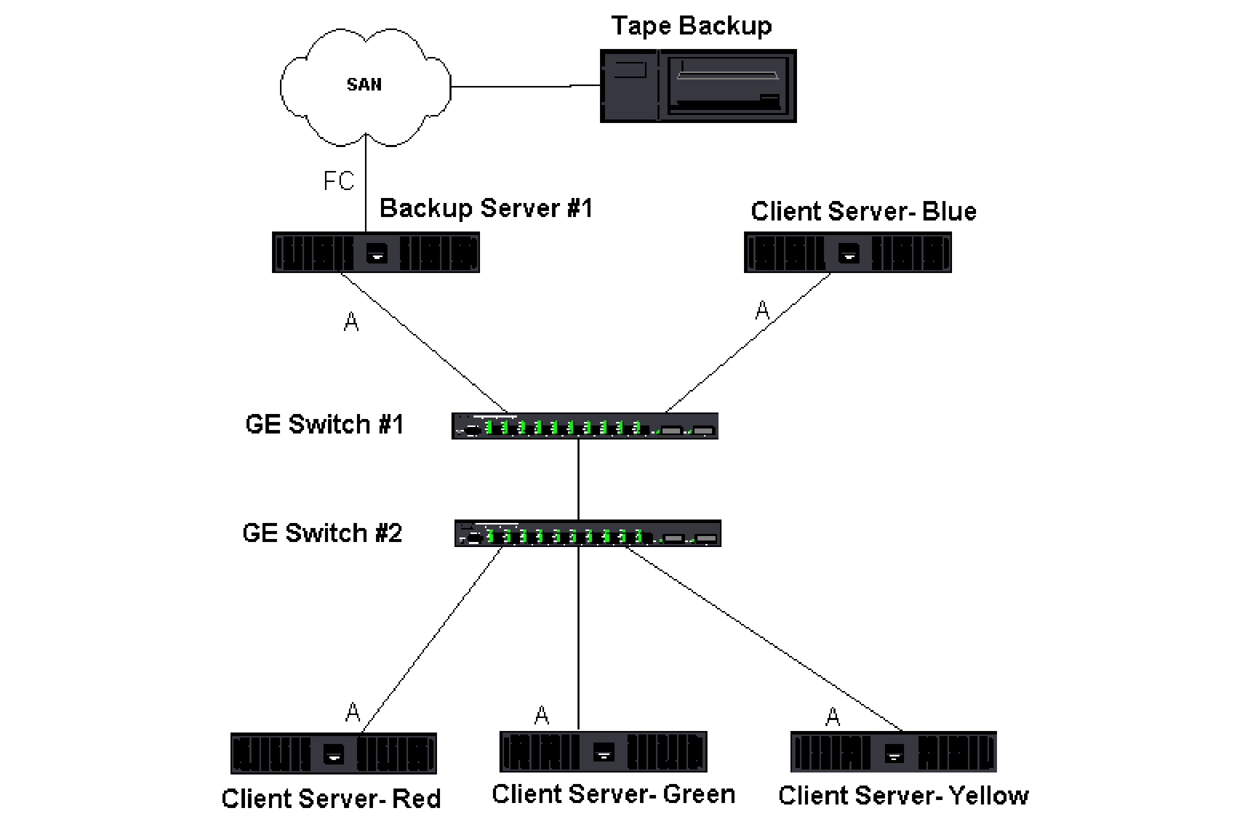


Figure 12-9. Network Backup Without Teaming

Because there are four client servers, the backup server can simultaneously stream four backup jobs (one per client) to a multidrive autoloader. Because of the single link between the switch and the backup server; however, a four-stream backup can easily saturate the adapter and link. If the adapter on the backup server operates at 1Gbps (125MBps), and each client is able to stream data at 20MBps during tape backup, the throughput between the backup server and switch will be at 80MBps (20MBps × 4), which is equivalent to 64 percent of the network bandwidth. Although this is well within the network bandwidth range, the 64 percent constitutes a high percentage, especially if other applications share the same link.

Teaming and network backup information includes:

- Load Balancing and Failover
- Fault Tolerance

Load Balancing and Failover

As the quantity of backup streams increases, the overall throughput increases. Each data stream, however, may not be able to maintain the same performance as a single backup stream of 25MBps. In other words, even though a backup server can stream data from a single client at 25MBps, it is not expected that four simultaneously-running backup jobs will stream at 100MBps (25MBps × 4 streams). Although overall throughput increases as the quantity of backup streams increases, each backup stream can be impacted by tape software or network stack limitations.

For a tape backup server to reliably use adapter performance and network bandwidth when backing up clients, a network infrastructure must implement teaming such as load balancing and fault tolerance. Data centers will incorporate redundant switches, link aggregation, and trunking as part of their fault tolerant solution. Although teaming device drivers will manipulate the way data flows through teamed interfaces and failover paths, this is transparent to tape backup applications and does not interrupt any tape backup process when backing up remote systems over the network. Figure 12-10 on page 189 shows a network topology that demonstrates tape backup in a Marvell teamed environment and how smart load balancing can *load balance* tape backup data across teamed adapters.

The client-server can use four paths to send data to the backup server, but only one of these paths is designated during data transfer. One possible path that Client-Server Red can use to send data to the backup server is shown in the following example.

Example Path: Client-Server Red sends data through Adapter A, Switch 1, Backup Server Adapter A.

The designated path is determined by two factors:

- Client-Server ARP cache points to the backup server MAC address. This address is determined by the Marvell intermediate driver inbound load balancing algorithm.
- The physical adapter interface on Client-Server Red transmits the data. The Marvell intermediate driver outbound load-balancing algorithm determines the data (see “Outbound Traffic Flow” on page 160 and “Inbound Traffic Flow (SLB Only)” on page 160.

The teamed interface on the backup server transmits a gratuitous address resolution protocol (G-ARP) to Client-Server Red, which in turn causes the client-server ARP cache to get updated with the Backup Server MAC address. The load balancing mechanism within the teamed interface determines the MAC address embedded in the G-ARP. The selected MAC address is essentially the destination for data transfer from the client server.

On Client-Server Red, the SLB teaming algorithm will determine which of the two adapter interfaces is used to transmit data. In this example, data from Client Server Red is received on the backup server Adapter A interface. To demonstrate the SLB mechanisms when additional load is placed on the teamed interface, consider the scenario when the backup server initiates a second backup operation: one to Client-Server Red, and one to Client-Server Blue. The route that Client-Server Blue uses to send data to the backup server is dependent on its ARP cache, which points to the backup server MAC address. Because Adapter A of the backup server is already under load from its backup operation with Client-Sever Red, the Backup Server invokes its SLB algorithm to *inform* Client-Server Blue (through an G-ARP) to update its ARP cache to reflect the backup server Adapter B MAC address. When Client-Server Blue needs to transmit data, it uses either one of its adapter interfaces, which is determined by its own SLB algorithm. What is important is that data from Client-Server Blue is received by the Backup Server Adapter B interface, and not by its Adapter A interface. This action is important because, with both backup streams running simultaneously, the backup server must *load balance* data streams from different clients. With both backup streams running, each adapter interface on the backup server is processing an equal load, thus load-balancing data across both adapter interfaces.

The same algorithm applies if a third and fourth backup operation is initiated from the backup server. The teamed interface on the backup server transmits a unicast G-ARP to backup clients to inform them to update their ARP cache. Each client then transmits backup data along a route to the target MAC address on the backup server.

Fault Tolerance

If a network link fails during tape backup operations, all traffic between the backup server and client stops and backup jobs fail. If, however, the network topology was configured for both Marvell SLB and switch fault tolerance, this configuration would allow tape backup operations to continue without interruption during the link failure. All failover processes within the network are transparent to tape backup software applications.

To understand how backup data streams are directed during network failover process, consider the topology in Figure 12-10. Client-Server Red is transmitting data to the backup server through Path 1, but a link failure occurs between the backup server and the switch. Because the data can no longer be sent from Switch #1 to the Adapter A interface on the backup server, the data is redirected from Switch #1 through Switch #2, to the Adapter B interface on the backup server. This redirection occurs without the knowledge of the backup application because all fault tolerant operations are handled by the adapter team interface and trunk settings on the switches. From the client-server perspective, it still operates as if it is transmitting data through the original path.

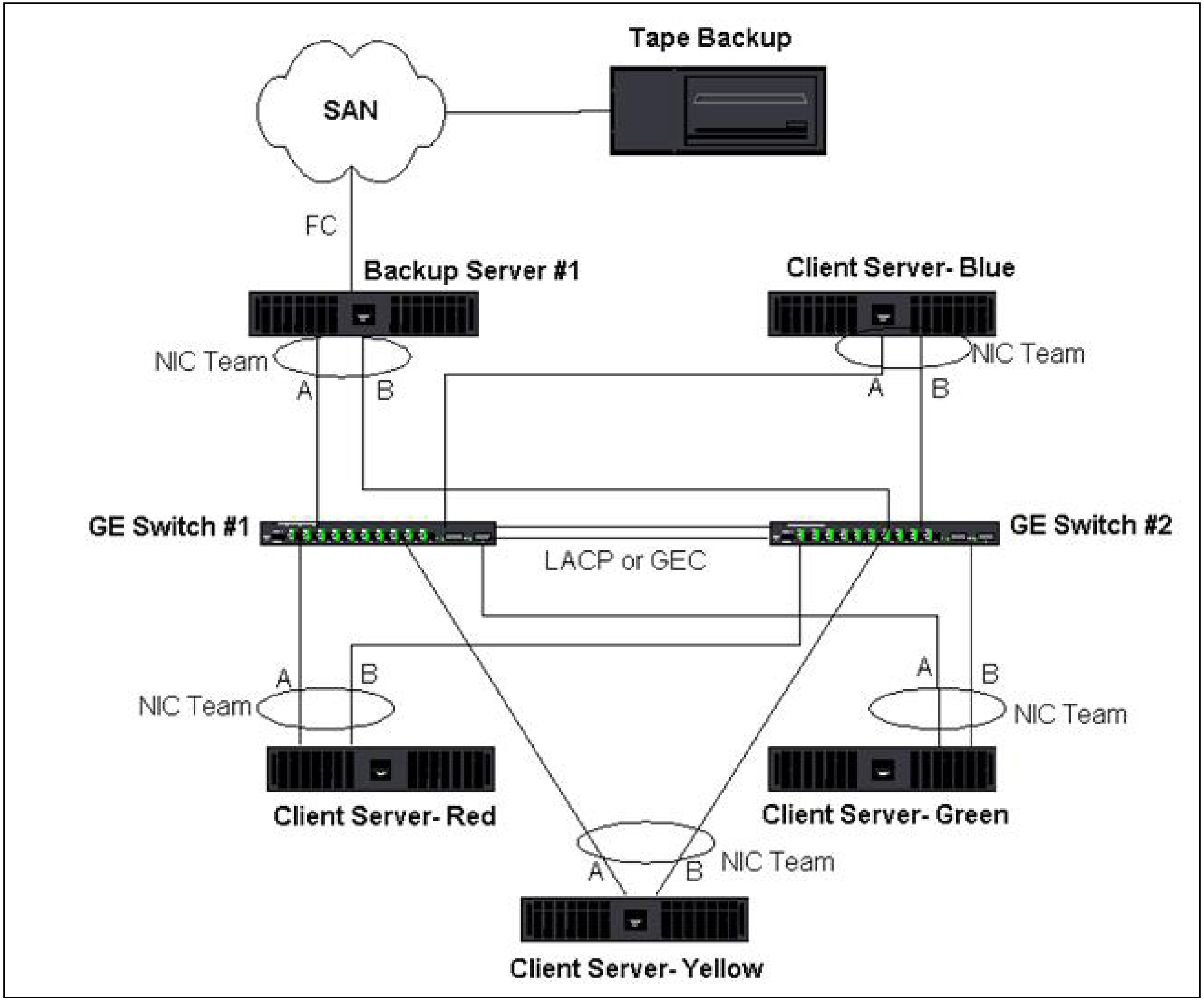


Figure 12-10. Network Backup with SLB Teaming Across Two Switches

Troubleshooting Teaming Problems

When running a protocol analyzer over a virtual adapter teamed interface, the MAC address shown in the transmitted frames may not be correct. The analyzer does not show the frames as constructed by QLASP and shows the MAC address of the team and not the MAC address of the interface transmitting the frame. Marvell suggests that you use the following process to monitor a team:

- Mirror all uplink ports from the team at the switch.
- If the team spans two switches, also mirror the interlink trunk.
- Sample all mirror ports independently.
- On the analyzer, use an adapter and driver that does not filter QoS and VLAN information.

Details for troubleshooting teaming problems are covered in:

- Teaming Configuration Tips
- Troubleshooting Guidelines

Teaming Configuration Tips

When troubleshooting network connectivity or teaming functionality issues, ensure that the following information is true for your configuration.

- Although mixed-speed SLB teaming is supported by Dell, Marvell recommends that all adapters in a team be the same speed (either all Gigabit Ethernet or all Fast Ethernet). For speeds of 10Gbps, Marvell highly recommends that all adapters in a team are the same speed.
- If LiveLink is not enabled, disable spanning tree protocol (STP), or enable an STP mode that bypasses the initial phases (for example, Port Fast, Edge Port) for the switch ports connected to a team.
- All switches that the team is directly connected to must have the same hardware revision, firmware revision, and software revision to be supported.
- To be teamed, adapters must be members of the same VLAN. In the event that multiple teams are configured, each team should be on a separate network.
- Do not assign a locally administered address on any physical adapter that is a member of a team.
- Verify that power management is disabled on all physical members of any team.
- Remove any static IP address from the individual physical team members before the team is built.

- A team that requires maximum throughput should use LACP or GEC\FEC. In these cases, the intermediate driver is only responsible for the outbound load balancing while the switch performs the inbound load balancing.
- Aggregated teams (802.3ad\LACP and GEC\FEC) must be connected to only a single switch that supports IEEE 802.3a, LACP, or GEC/FEC.
- Marvell does not recommend that you connect any team to a hub, because hubs only support half duplex. Connect hubs to a team for troubleshooting purposes only. Disabling the device driver of a network adapter participating in an LACP or GEC/FEC team may have adverse affects with network connectivity. To avoid a network connection loss, Marvell recommends that you first physically disconnect the adapter from the switch before disabling the device driver.
- Verify that the base (miniport) and team (intermediate) drivers are from the same release package. Dell does not test or support mixing base and teaming drivers from different releases.
- Before placing into a production environment, test the connectivity to each physical adapter prior to teaming.
- Test the failover and fallback behavior of the team.
- When moving from a nonproduction network to a production network, it is strongly recommended to test again for failover and fallback.
- Test the performance behavior of the team before placing into a production environment.
- Network teaming is not supported when running iSCSI traffic through the Microsoft iSCSI initiator or iSCSI offload. Use MPIO instead of Marvell network teaming for these ports.
- For information on iSCSI boot and iSCSI offload restrictions, see Chapter 11 iSCSI Protocol.

Troubleshooting Guidelines

Before you call Dell support, make sure you have completed the following steps for troubleshooting network connectivity problems when the server is using adapter teaming.

1. Make sure the link light is ON for every adapter and that all the cables are attached.
2. Check that the matching base and intermediate drivers belong to the same Dell release and are loaded correctly.
3. Check for a valid IP address using the Windows `ipconfig` command.
4. Check that STP is disabled or Edge Port or Port Fast is enabled on the switch ports connected to the team or that LiveLink is being used.

5. Check that the adapters and the switch are configured identically for link speed and duplex.
6. If possible, break the team and check for connectivity to each adapter independently to confirm that the problem is directly associated with teaming.
7. Check that all switch ports connected to the team are on the same VLAN.
8. Check that the switch ports are configured properly for Generic Trunking (FEC/GEC)/802.3ad-Draft Static type of teaming and that it matches the adapter teaming type. If the system is configured for an SLB type of team, make sure the corresponding switch ports *are not* configured for Generic Trunking (FEC/GEC)/802.3ad-Draft Static types of teams.

Frequently Asked Questions

Question: Under what circumstances is traffic not load balanced? Why is all traffic not load balanced evenly across the team members?

Answer: The bulk of traffic does not use IP, TCP, or UDP, or the bulk of the clients are in a different network. The receive load balancing is not a function of traffic load, but a function of the quantity of clients that are connected to the server.

Question: What network protocols are load balanced when in a team?

Answer: Marvell's teaming software only supports IP, TCP, and UDP traffic. All other traffic is forwarded to the primary adapter.

Question: Which protocols are load balanced with SLB and which ones are not?

Answer: Only IP, TCP, and UDP protocols are load balanced in both directions: send and receive. IPX is load balanced on the transmit traffic only.

Question: Can I team a port running at 100Mbps with a port running at 1000Mbps?

Answer: Mixing link speeds within a team is only supported for Smart Load Balancing teams and 802.3ad teams.

Question: Can I team a fiber adapter with a copper GbE adapter?

Answer: Yes with SLB, and yes if the switch allows for it in FEC and GEC and 802.3ad.

Question: What is the difference between adapter load balancing and Microsoft's Network Load Balancing?

Answer: Adapter load balancing is done at a network session level, whereas Network Load Balancing is done at the server application level.

Question: Can I connect the teamed adapters to a hub?

Answer: Teamed ports can be connected to a hub for troubleshooting purposes only. However, this practice is not recommended for normal operation because the performance would be degraded due to hub limitations. Connect the teamed ports to a switch instead.

Question: Can I connect the teamed adapters to ports in a router?

Answer: No. All ports in a team must be on the same network; in a router, however, each port is a separate network by definition. All teaming modes require that the link partner be a Layer 2 switch.

Question: Can I use teaming with Microsoft Cluster Services?

Answer: Yes. Teaming is supported on the public network only, not on the private network used for the heartbeat link.

Question: Can PXE work over a virtual adapter (team)?

Answer: A PXE client operates in an environment before the operating system is loaded; as a result, virtual adapters have not been enabled yet. If the physical adapter supports PXE, it can be used as a PXE client, whether or not it is part of a virtual adapter when the operating system loads. PXE servers may operate over a virtual adapter.

Question: Can WoL work over a virtual adapter (team)?

Answer: Wake on LAN functionality operates in an environment before the operating system is loaded. WoL occurs when the system is off or in standby, so no team is configured.

Question: What is the maximum quantity of ports that can be teamed together?

Answer: Up to 16 ports can be assigned to a team, of which one port can be a standby team member.

Question: What is the maximum quantity of teams that can be configured on the same server?

Answer: Up to 16 teams can be configured on the same server.

Question: Why does my team lose connectivity for the first 30 to 50 seconds after the Primary adapter is restored (fallback)?

Answer: Because Spanning Tree Protocol is bringing the port from blocking to forwarding. You must enable Port Fast or Edge Port on the switch ports connected to the team or use LiveLink to account for the STP delay.

Question: Can I connect a team across multiple switches?

Answer: Smart Load Balancing can be used with multiple switches because each physical adapter in the system uses a unique Ethernet MAC address. Link Aggregation and Generic Trunking cannot operate across switches because they require all physical adapters to share the same Ethernet MAC address.

Question: How do I upgrade the intermediate driver (QLASP)?

Answer: The intermediate driver cannot be upgraded through the Local Area Connection Properties. It must be upgraded using the Setup installer.

Question: How can I determine the performance statistics on a virtual adapter (team)?

Answer: In QLogic Control Suite, click the **Statistics** tab for the virtual adapter.

Question: Can I configure Network Load Balancing and teaming concurrently?

Answer: Yes, but only when running Network Load Balancing in a multicast mode (Network Load Balancing is not supported with MS Cluster Services).

Question: Should both the backup server and client servers that are backed up be teamed?

Answer: Because the backup server is under the most data load, it should always be teamed for link aggregation and failover. A fully redundant network, however, requires that both the switches and the backup clients be teamed for fault tolerance and link aggregation.

Question: During backup operations, does the adapter teaming algorithm load balance data at a byte-level or a session-level?

Answer: When using adapter teaming, data is only load balanced at a session level and not a byte level to prevent out-of-order frames. Adapter teaming load balancing does not work the same way as other storage load balancing mechanisms such as EMC PowerPath.

Question: Is there any special configuration required in the tape backup software or hardware to work with adapter teaming?

Answer: No special configuration is required in the tape software to work with teaming. Teaming is transparent to tape backup applications.

Question: How do I know what driver I am currently using?

Answer: In all operating systems, the most accurate method for checking the driver revision is to physically locate the driver file and check the properties.

Question: Can SLB detect a switch failure in a Switch Fault Tolerance configuration?

Answer: No. SLB can only detect the loss of link between the teamed port and its immediate link partner. SLB cannot detect link failures on other ports.

Question: Where can I get the latest supported drivers?

Answer: Go to Dell support at http://support.dell.com for driver package updates or support documents.

Question: Why does my team lose connectivity for the first 30 to 50 seconds after the primary adapter is restored (fall-back after a failover)?

Answer: During a fall-back event, link is restored causing Spanning Tree Protocol to configure the port for blocking until it determines that it can move to the forwarding state. You must enable Port Fast or Edge Port on the switch ports connected to the team to prevent the loss of communications caused by STP.

Question: Where do I monitor real time statistics for an adapter team in a Windows server?

Answer: Use QCC GUI or QCS CLI to monitor general, IEEE 802.3 and custom counters.

Question: What features are not supported on a multivendor team?

Answer: VLAN tagging, and RSS are not supported on a multivendor team.

Event Log Messages

Event log messages include the following:

- Windows System Event Log Messages
- Base Driver (Physical Adapter or Miniport)
- Intermediate Driver (Virtual Adapter or Team)
- Virtual Bus Driver (VBD)

Windows System Event Log Messages

The known base and intermediate Windows System Event Log status messages for the Marvell 57*xx* and 57*xxx* adapters are listed in Table 12-8 on page 196 and Table 12-9 on page 199. As a Marvell adapter driver loads, Windows places a status code in the system event viewer. There may be up to two classes of entries for these event codes depending on whether both drivers are loaded (one set for the base or miniport driver and one set for the intermediate or teaming driver).

Base Driver (Physical Adapter or Miniport)

The base driver is identified by source **L2ND**. Table 12-8 lists the event log messages supported by the base driver, explains the cause for the message, and provides the recommended action.

NOTE

In Table 12-8, message numbers 1 through 17 apply to both NDIS 5.*x* and NDIS 6.*x* drivers, message numbers 18 through 23 apply only to the NDIS 6.*x* driver.

Table 12-8. Base Driver Event Log Messages

Message Number	Severity	Message	Cause	Corrective Action
1	Error	Failed to allocate memory for the device block. Check system memory resource usage.	The driver cannot allocate memory from the operating system.	Close running applications to free memory.
2	Error	Failed to allocate map registers.	The driver cannot allocate map registers from the operating system.	Unload other drivers that may allocate map registers.
3	Error	Failed to access configuration information. Reinstall the network driver.	The driver cannot access PCI configuration space registers on the adapter.	For add-in adapters: reseat the adapter in the slot, move the adapter to another PCI slot, or replace the adapter.
4	Warning	The network link is down. Check to make sure the network cable is properly connected.	The adapter has lost its connection with its link partner.	Check that the network cable is connected, verify that the network cable is the right type, and verify that the link partner (for example, switch or hub) is working correctly.
5	Informational	The network link is up.	The adapter has established a link.	No action is required.

Table 12-8. Base Driver Event Log Messages (Continued)

Message Number	Severity	Message	Cause	Corrective Action
6	Informational	Network controller configured for 10Mb half-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
7	Informational	Network controller configured for 10Mb full-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
8	Informational	Network controller configured for 100Mb half-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
9	Informational	Network controller configured for 100Mb full-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
10	Informational	Network controller configured for 1Gb half-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
11	Informational	Network controller configured for 1Gb full-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
12	Informational	Network controller configured for 2.5Gb full-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
13	Error	Medium not supported.	The operating system does not support the IEEE 802.3 medium.	Reboot the operating system, run a virus check, run a disk check (chkdsk), and reinstall the operating system.
14	Error	Unable to register the interrupt service routine.	The device driver cannot install the interrupt handler.	Reboot the operating system; remove other device drivers that may be sharing the same IRQ.

Table 12-8. Base Driver Event Log Messages (Continued)

Message Number	Severity	Message	Cause	Corrective Action
15	Error	Unable to map I/O space.	The device driver cannot allocate memory-mapped I/O to access driver registers.	Remove other adapters from the system, reduce the amount of physical memory installed, and replace the adapter.
16	Informational	Driver initialized successfully.	The driver has successfully loaded.	No action is required.
17	Informational	NDIS is resetting the miniport driver.	The NDIS layer has detected a problem sending/receiving packets and is resetting the driver to resolve the problem.	Run QLogic Control Suite diagnostics; check that the network cable is good.
18	Error	Unknown PHY detected. Using a default PHY initialization routine.	The driver could not read the PHY ID.	Replace the adapter.
19	Error	This driver does not support this device. Upgrade to the latest driver.	The driver does not recognize the installed adapter.	Upgrade to a driver version that supports this adapter.
20	Error	Driver initialization failed.	Unspecified failure during driver initialization.	Reinstall the driver, update to a newer driver, run QLogic Control Suite diagnostics, or replace the adapter.
21	Informational	Network controller configured for 10Gb full-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
22	Error	Network controller failed initialization because it cannot allocate system memory.	Insufficient system memory prevented the initialization of the driver.	Increase system memory.

Table 12-8. Base Driver Event Log Messages (Continued)

Message Number	Severity	Message	Cause	Corrective Action
23	Error	Network controller failed to exchange the interface with the bus driver.	The driver and the bus driver are not compatible.	Update to the latest driver set, ensuring the major and minor versions for both NDIS and the bus driver are the same.

Intermediate Driver (Virtual Adapter or Team)

The intermediate driver is identified by source **BLFM**, regardless of the base driver revision. Table 12-9 lists the event log messages supported by the intermediate driver, explains the cause for the message, and provides the recommended action.

Table 12-9. Intermediate Driver Event Log Messages

System Event Message Number	Severity	Message	Cause	Corrective Action
1	Informational	Event logging enabled for QLASP driver.	—	No action is required.
2	Error	Unable to register with NDIS.	The driver cannot register with the NDIS interface.	Unload other NDIS drivers.
3	Error	Unable to instantiate the management interface.	The driver cannot create a device instance.	Reboot the operating system.
4	Error	Unable to create symbolic link for the management interface.	Another driver has created a conflicting device name.	Unload the conflicting device driver that uses the name *Blf*.
5	Informational	QLASP driver has started.	The driver has started.	No action is required.
6	Informational	QLASP driver has stopped.	The driver has stopped.	No action is required.

Table 12-9. Intermediate Driver Event Log Messages (Continued)

System Event Message Number	Severity	Message	Cause	Corrective Action
7	Error	Could not allocate memory for internal data structures.	The driver cannot allocate memory from the operating system.	Close running applications to free memory.
8	Warning	Could not bind to adapter.	The driver could not open one of the team physical adapters.	Unload and reload the physical adapter driver, install an updated physical adapter driver, or replace the physical adapter.
9	Informational	Successfully bind to adapter.	The driver successfully opened the physical adapter.	No action is required.
10	Warning	Network adapter is disconnected.	The physical adapter is not connected to the network (it has not established link).	Check that the network cable is connected, verify that the network cable is the right type, and verify that the link partner (switch or hub) is working correctly.
11	Informational	Network adapter is connected.	The physical adapter is connected to the network (it has established link).	No action is required.
12	Error	QLASP features driver is not designed to run on this version of operating system.	The driver does not support the operating system on which it is installed.	Consult the driver release notes and install the driver on a supported operating system or update the driver.
13	Informational	Hot-standby adapter is selected as the primary adapter for a team without a load balancing adapter.	A standby adapter has been activated.	Replace the failed physical adapter.

Table 12-9. Intermediate Driver Event Log Messages (Continued)

System Event Message Number	Severity	Message	Cause	Corrective Action
14	Informational	Network adapter does not support Advanced Failover.	The physical adapter does not support the Marvell NIC Extension (NICE).	Replace the adapter with one that does support NICE.
15	Informational	Network adapter is enabled through management interface.	The driver has successfully enabled a physical adapter through the management interface.	No action is required.
16	Warning	Network adapter is disabled through management interface.	The driver has successfully disabled a physical adapter through the management interface.	No action is required.
17	Informational	Network adapter is activated and is participating in network traffic.	A physical adapter has been added to or activated in a team.	No action is required.
18	Informational	Network adapter is deactivated and is no longer participating in network traffic.	The driver does not recognize the installed adapter.	No action is required.
19	Informational	The LiveLink feature in QLASP connected the link for the network adapter.	The connection with the remote target(s) for the LiveLink-enabled team member has been established or restored	No action is required.
20	Informational	The LiveLink feature in QLASP disconnected the link for the network adapter.	The LiveLink-enabled team member is unable to connect with the remote target(s).	No action is required.

Virtual Bus Driver (VBD)

Table 12-10 lists VBD event log messages.

Table 12-10. Virtual Bus Driver (VBD) Event Log Messages

Message Number	Severity	Message	Cause	Corrective Action
1	Error	Failed to allocate memory for the device block. Check system memory resource usage.	The driver cannot allocate memory from the operating system.	Close running applications to free memory.
2	Informational	The network link is down. Check to make sure the network cable is properly connected.	The adapter has lost its connection with its link partner.	Check that the network cable is connected, verify that the network cable is the right type, and verify that the link partner (for example, switch or hub) is working correctly.
3	Informational	The network link is up.	The adapter has established a link.	No action is required.
4	Informational	Network controller configured for 10Mb half-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
5	Informational	Network controller configured for 10Mb full-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
6	Informational	Network controller configured for 100Mb half-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
7	Informational	Network controller configured for 100Mb full-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.

Table 12-10. Virtual Bus Driver (VBD) Event Log Messages (Continued)

Message Number	Severity	Message	Cause	Corrective Action
8	Informational	Network controller configured for 1Gb half-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
9	Informational	Network controller configured for 1Gb full-duplex link.	The adapter has been manually configured for the selected line speed and duplex settings.	No action is required.
10	Error	Unable to register the interrupt service routine.	The device driver cannot install the interrupt handler.	Reboot the operating system; remove other device drivers that may be sharing the same IRQ.
11	Error	Unable to map I/O space.	The device driver cannot allocate memory-mapped I/O to access driver registers.	Remove other adapters from the system, reduce the amount of physical memory installed, and replace the adapter.
12	Informational	Driver initialized successfully.	The driver has successfully loaded.	No action is required.
13	Error	Driver initialization failed.	Unspecified failure during driver initialization.	Reinstall the driver, update to a newer driver, run QLogic Control Suite diagnostics, or replace the adapter.
14	Error	This driver does not support this device. Upgrade to the latest driver.	The driver does not recognize the installed adapter.	Upgrade to a driver version that supports this adapter.
15	Error	This driver fails initialization because the system is running out of memory.	Insufficient system memory prevented the initialization of the driver.	Increase system memory.

13 NIC Partitioning and Bandwidth Management

NIC partitioning and bandwidth management covered in this chapter includes:

Overview

NIC partitioning (NPAR) divides a Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet NIC into multiple virtual NICs by having multiple PCI physical functions per port. Each PCI function is associated with a different virtual NIC. To the OS and the network, each physical function appears as a separate NIC port.

The quantity of partitions for each port can range from one to four; thus, a dual-port NIC can have up to eight partitions. Each partition behaves as if it is an independent NIC port.

Benefits of a partitioned 10G NIC include:

- Reduced cabling and ports when used to replace many 1G NICs.
- Server segmentation with separate subnets and VLANs.
- High server availability with NIC failover and NIC link bandwidth aggregation.
- Server I/O virtualization with a virtual OS and monolithic OS support.
- No change to the OS is required.
- Switch-independent type teaming is supported.

Supported Operating Systems for NIC Partitioning

The Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet adapters support NIC partitioning on the following operating systems:

- Windows
  - 2016 Server
  - 2019 Server
  - Azure Stack HCI
- Linux
  - RHEL 8.*x* and later family
  - RHEL 7.*x* and later family
  - SLES 15.*x* and later family
- VMware
  - ESX 6.*x* and later family
  - ESX 7.*x* and later family

NOTE

32-bit Linux operating systems have a limited amount of memory space available for Kernel data structures. Therefore, Marvell recommends that you use only 64-bit Linux to configure NPAR.

Some older OS versions may require an earlier driver release.

Configuring for NIC Partitioning

When NIC partitioning is enabled on an adapter, by default, no offloads are enabled on any physical function (PF) or virtual NIC (VNIC). The user must explicitly configure storage offloads on a PF to use FCoE and/or iSCSI offload functionality on an adapter.

NIC partitioning can be configured using the **UEFI HII** menu. You can access the **UEFI HII** menu by pressing the Dell F2 key during the system boot (UEFI should be supported by most Dell server BIOSs). For more information on using the **UEFI HII** menu, refer to the Dell server documentation.

NIC partitioning can also be configured using pre-boot CCM, Linux and Windows QCC GUI, Linux and Windows QCS CLI, and the VMware QCC vSphere GUI plug-in. See the respective user's guides for more information.

NOTE

In NPAR mode, SR-IOV cannot be enabled on any partition or PF (VNIC) on which storage offload (FCoE or iSCSI) is configured. This does not apply to adapters in Single Function (SF) mode.

Configure NPAR mode (and reboot the system) before attempting to configure the SR-IOV settings on any NPAR-ed partitions of an adapter port. The NPAR mode configuration will take precedence over the SR-IOV configuration.

To configure a NIC for partitioning using the CCM utility:

1. Select the NIC from **Device List**.
2. From the **Main Menu**, select **Device Hardware Configuration**.
3. Change the **Multi-Function Mode** to **NPAR**.
4. Configure the NIC parameters for your configuration based on the options shown in Table 13-1, which lists the configuration parameters available on the NIC Partitioning Configuration window.

Table 13-1. Configuration Options

Parameter	Description	Options
Flow Control	Configures the Flow Control mode for this port.	■ Auto ■ TX Flow Control ■ RX Flow Control ■ TX/RX Flow Control ■ None
PF#0, PF#2, PF#4, PF#6	Displays the physical function (PF) information regarding the partition(s) on port 0. Select to configure.	See Table 13-2 for configuration options.
PF#1, PF#3, PF#5, PF#7	Displays the physical function (PF) information regarding the partition(s) on port 1. Select to configure.	See Table 13-2 for configuration options.
Reset Configuration to Default	Resets the NIC partition configuration to the factory default settings.	—

Table 13-2 describes the functions available from the PF# *X* window.

Table 13-2. Function Description

Function	Description	Option
Ethernet Protocol	Enables and disables the Ethernet protocol.	■ Enable ■ Disable
iSCSI Offload Protocol	Enables and disables the iSCSI protocol.	■ Enable ■ Disable
FCoE Offload protocol	Enables and disables the FCoE protocol.	■ Enable ■ Disable
Bandwidth Weight	Configures the weight or importance of a specific function. For the four functions per port, the weight is used to arbitrate between the functions in case of congestion.	The sum of all weights for the four functions are either 0 or 100.
Maximum Bandwidth	Configures the maximum bandwidth (in percentage) of the physical port link.	—
Network MAC Address [a]	Displays the network MAC address.	—
iSCSI MAC Address[a]	Displays the iSCSI MAC address.	—
FCoE FIP MAC Address	Displays the FCoE MAC address.	—
FCoE WWPN	Displays the FCoE world wide port name.	—
FCoE WWNN	Displays the FCoE world wide node name.	—

[a] Ensure that the **Network MAC Address** and the **iSCSI MAC Address** are not the same.

NOTE

For Linux, Citrix XenServer, and VMware ESXi OSs, the Ethernet protocol for all partitions is always enabled, even if you disable the Ethernet personality using the Marvell Comprehensive Configuration Management (CCM) tool.

Configuring equal **Bandwidth Weight** values for all functions has different effects depending on the actual values used for configuration. For example, when all functions are configured as “0” or “25”, offloads configured on these functions exhibit different bandwidth settings even though, logically, they would be expected to have the same effect.

Consider this example configuration: Four functions (or partitions) are configured with a total of six protocols, as shown in the following.

Function 0

- Ethernet
- FCoE

Function 1

- Ethernet

Function 2

- Ethernet

Function 3

- Ethernet
- iSCSI

1. If **Relative Bandwidth Weight** is configured as “0” for all four physical functions (PFs), all six offloads share the bandwidth equally. In this case, each offload are assigned roughly 16.67 percent of the total bandwidth.
2. If **Relative Bandwidth Weight** is configured as “25” for all four PFs, Ethernet and FCoE offloads on function 0 and Ethernet and iSCSI offloads on function 3 are assigned roughly 12.5 percent of the total bandwidth, whereas Ethernet offloads on function 1 and function 2 are assigned roughly 25 percent of the total bandwidth.

14 Fibre Channel Over Ethernet

Fibre Channel over Ethernet (FCoE) information includes:

Overview

In today's data center, multiple networks, including network attached storage (NAS), management, IPC, and storage, are used to achieve the performance and versatility that you require. In addition to iSCSI for storage solutions, Fibre Channel over Ethernet (FCoE) can now be used with capable Marvell C-NICs. FCoE is a standard that allows Fibre Channel protocol to be transferred over Ethernet by preserving existing Fibre Channel infrastructures and capital investments by classifying received FCoE and FCoE Initialization Protocol (FIP) frames.

The following FCoE features are supported:

- Receiver classification of FCoE and FIP frames. FIP is the FCoE Initialization Protocol used to establish and maintain connections.
- Receiver CRC offload
- Transmitter CRC offload
- Dedicated queue set for Fibre Channel traffic
- N_Port ID virtualization (NPIV) on Windows and Linux
- Virtual machine virtual Fibre Channel (vFC) Host Bus Adapters in Windows Server 2012 and later, and R2 Hyper-V

- Data center bridging (DCB) provides lossless behavior with priority flow control (PFC)
- DCB allocates a share of link bandwidth to FCoE traffic with enhanced transmission selection (ETS)

DCB supports storage, management, computing, and communications fabrics onto a single physical fabric that is simpler to deploy, upgrade, and maintain than in standard Ethernet networks. DCB technology allows the capable Marvell C-NICs to provide lossless data delivery, lower latency, and standards-based bandwidth sharing of data center physical links. The DCB supports FCoE, iSCSI, network attached storage (NAS), management, and IPC traffic flows. For more information on DCB, see Chapter 15 Data Center Bridging.

Configure NPIV in Windows QCC GUI by clicking an FCoE adapter instance and then selecting either **Create a Virtual Port** or **Create Multiple Virtual Ports**. You can also issue the QCS CLI **`createnpivport`** and **`createmultinpivport`** commands. Configure NPIV in Linux by issuing the **`vport_create`** command.

Add Windows Server vFCs by issuing the Windows Server 2012 R2 (and later) PowerShell **`Add-VMFibreChannelHBA`** command.

FCoE Boot from SAN

This section describes the install and boot procedures for the Windows, Linux, and ESXi operating systems.

The following section details the BIOS setup and configuration of the boot environment prior to the OS install.

Preparing System BIOS for FCoE Build and Boot

To prepare the system BIOS for the FCoE build and boot, modify the system boot order and specify the BIOS boot protocol, if required.

Modifying System Boot Order

The Marvell initiator must be the first entry in the system boot order. The second entry must be the OS installation media. It is important that you correctly set the boot order, otherwise the installation will not proceed correctly. Either the boot LUN that you want will not be discovered, or it will be discovered but marked offline.

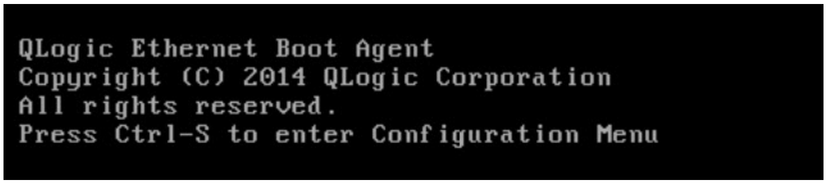
Specifying BIOS Boot Protocol (If Required)

On some platforms, the boot protocol must be configured through system BIOS configuration. On all other systems the boot protocol is specified through the Marvell Comprehensive Configuration Management (CCM), and for those systems this step is not required.

Preparing Marvell Multiple Boot Agent for FCoE Boot (CCM)

CCM is available only when the system is set to legacy boot mode; it is not available when the systems is set to UEFI boot mode. The UEFI device configuration pages are available in both modes.

1. Invoke the CCM utility during POST. At the QLogic Ethernet Boot Agent banner (Figure 14-1), press the CTRL+S keys.



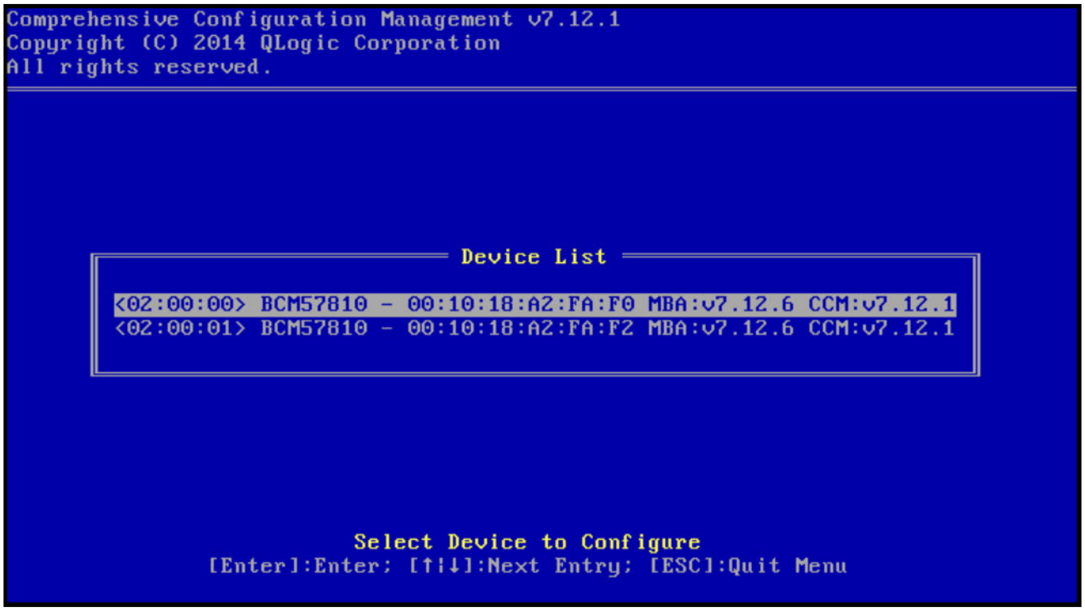
```
QLogic Ethernet Boot Agent
Copyright (C) 2014 QLogic Corporation
All rights reserved.
Press Ctrl-S to enter Configuration Menu
```

Figure 14-1. Invoking the CCM Utility

2. From the Device List (Figure 14-2), select the device through which boot is to be configured.

NOTE

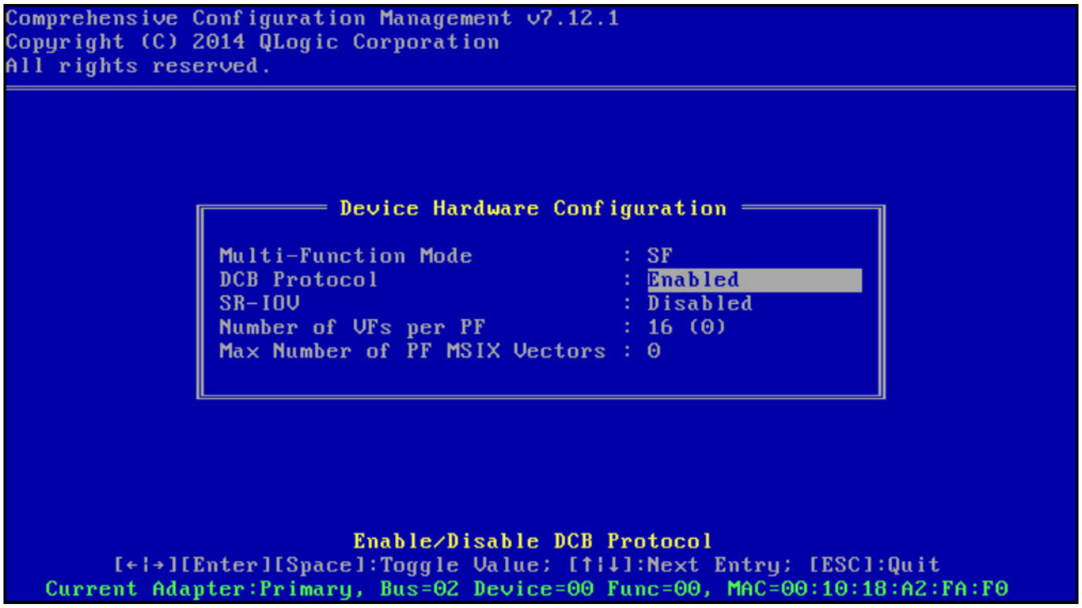
When running in NIC partitioning (NPAR) mode, FCoE boot is supported only when the first function of the booting port is assigned the FCoE personality. FCoE boot is not supported when the FCoE personality is assigned to any other function.



```
Comprehensive Configuration Management v7.12.1
Copyright (C) 2014 QLogic Corporation
All rights reserved.

                         Device List

     <02:00:00> BCM57810 - 00:10:18:A2:FA:F0 MBA:v7.12.6 CCM:v7.12.1
     <02:00:01> BCM57810 - 00:10:18:A2:FA:F2 MBA:v7.12.6 CCM:v7.12.1

                   Select Device to Configure
        [Enter]:Enter; [↑↓]:Next Entry; [ESC]:Quit Menu
```

Figure 14-2. CCM Device List

3. Ensure that DCB and DCBX are enabled on the device (Figure 14-3). FCoE boot is only supported on DCBX-capable configurations. As such, DCB and DCBX must be enabled, and the directly attached link peer must also be DCBX-capable with parameters that allow for full DCBX synchronization.

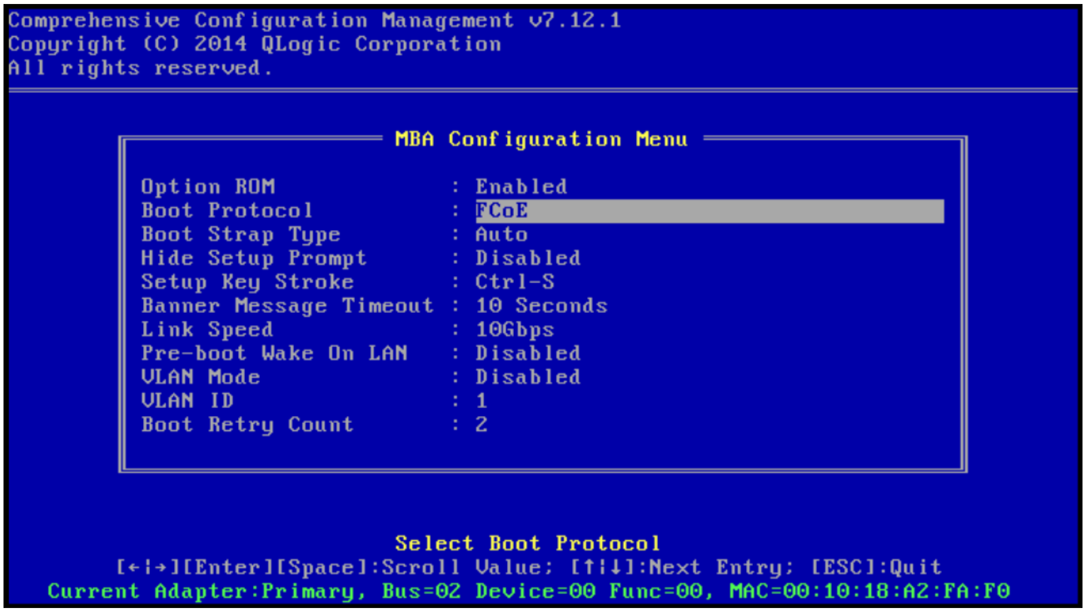


```
Comprehensive Configuration Management v7.12.1
Copyright (C) 2014 QLogic Corporation
All rights reserved.

                    Device Hardware Configuration

                Multi-Function Mode              : SF
                DCB Protocol                     : Enabled
                SR-IOV                           : Disabled
                Number of VFs per PF             : 16 (0)
                Max Number of PF MSIX Vectors    : 0

                        Enable/Disable DCB Protocol
        [←|→][Enter][Space]:Toggle Value; [↑|↓]:Next Entry; [ESC]:Quit
   Current Adapter:Primary, Bus=02 Device=00 Func=00, MAC=00:10:18:A2:FA:F0
```

Figure 14-3. CCM Device Hardware Configuration

4. On some platforms, you may need to set the boot protocol through system BIOS configuration in the integrated devices pane as described in the preceding.

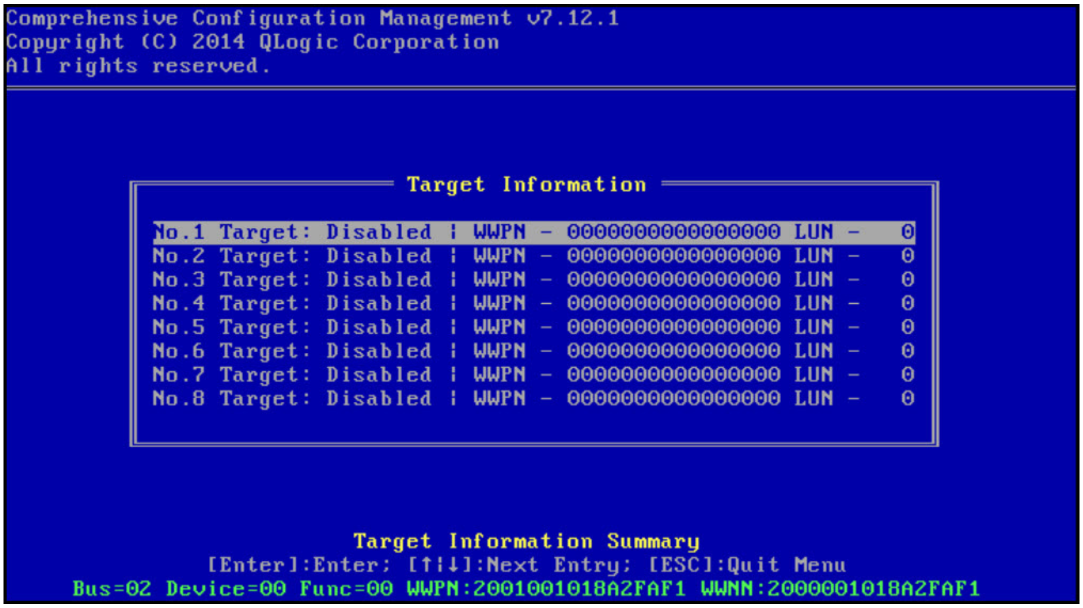
For all other devices, use the CCM **MBA Configuration Menu** to set the **Boot Protocol** option to **FCoE** (Figure 14-4).



```
Comprehensive Configuration Management v7.12.1
Copyright (C) 2014 QLogic Corporation
All rights reserved.

                         MBA Configuration Menu

     Option ROM                  : Enabled
     Boot Protocol               : FCoE
     Boot Strap Type             : Auto
     Hide Setup Prompt           : Disabled
     Setup Key Stroke            : Ctrl-S
     Banner Message Timeout      : 10 Seconds
     Link Speed                  : 10Gbps
     Pre-boot Wake On LAN        : Disabled
     VLAN Mode                   : Disabled
     VLAN ID                     : 1
     Boot Retry Count            : 2

                         Select Boot Protocol
    [←|→][Enter][Space]:Scroll Value; [↑|↓]:Next Entry; [ESC]:Quit
 Current Adapter:Primary, Bus=02 Device=00 Func=00, MAC=00:10:18:A2:FA:F0
```

Figure 14-4. CCM MBA Configuration Menu

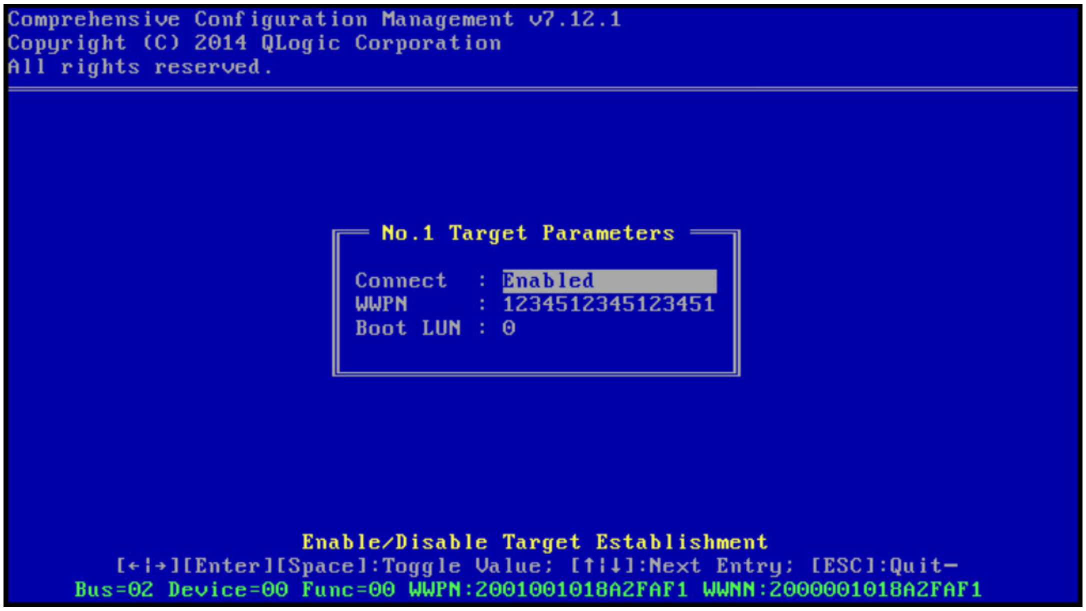
5. Configure the boot target and LUN. From the **Target Information** menu, select the first available path (Figure 14-5).



```
Comprehensive Configuration Management v7.12.1
Copyright (C) 2014 QLogic Corporation
All rights reserved.

                         Target Information

     No.1 Target: Disabled ¦ WWPN - 0000000000000000 LUN -   0
     No.2 Target: Disabled ¦ WWPN - 0000000000000000 LUN -   0
     No.3 Target: Disabled ¦ WWPN - 0000000000000000 LUN -   0
     No.4 Target: Disabled ¦ WWPN - 0000000000000000 LUN -   0
     No.5 Target: Disabled ¦ WWPN - 0000000000000000 LUN -   0
     No.6 Target: Disabled ¦ WWPN - 0000000000000000 LUN -   0
     No.7 Target: Disabled ¦ WWPN - 0000000000000000 LUN -   0
     No.8 Target: Disabled ¦ WWPN - 0000000000000000 LUN -   0

                      Target Information Summary
          [Enter]:Enter; [↑|↓]:Next Entry; [ESC]:Quit Menu
   Bus=02 Device=00 Func=00 WWPN:2001001018A2FAF1 WWNN:2000001018A2FAF1
```

Figure 14-5. CCM Target Information

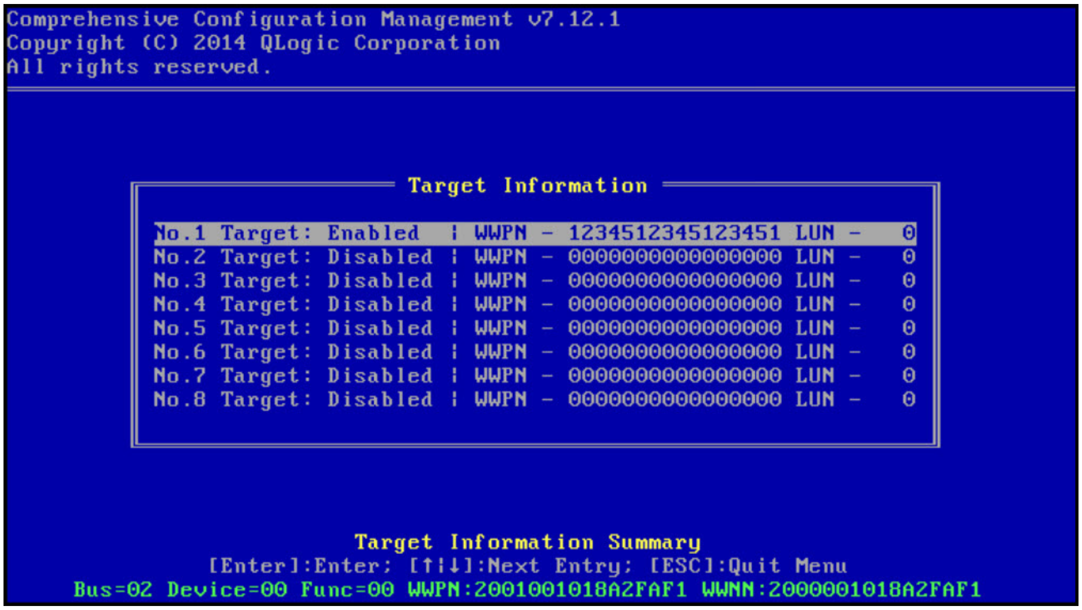
6. Enable the **Connect** option, and then the target WWPN and Boot LUN information for the target to be used for boot (Figure 14-6).



```
Comprehensive Configuration Management v7.12.1
Copyright (C) 2014 QLogic Corporation
All rights reserved.

                    No.1 Target Parameters

                    Connect   : Enabled
                    WWPN      : 1234512345123451
                    Boot LUN  : 0

              Enable/Disable Target Establishment
     [←|→][Enter][Space]:Toggle Value; [↑|↓]:Next Entry; [ESC]:Quit-
  Bus=02 Device=00 Func=00 WWPN:2001001018A2FAF1 WWNN:2000001018A2FAF1
```

Figure 14-6. CCM Target Parameters

The target information shows the changes (Figure 14-7).



```
Comprehensive Configuration Management v7.12.1
Copyright (C) 2014 QLogic Corporation
All rights reserved.

                      Target Information

          No.1 Target: Enabled  ¦ WWPN - 1234512345123451 LUN -  0
          No.2 Target: Disabled ¦ WWPN - 0000000000000000 LUN -  0
          No.3 Target: Disabled ¦ WWPN - 0000000000000000 LUN -  0
          No.4 Target: Disabled ¦ WWPN - 0000000000000000 LUN -  0
          No.5 Target: Disabled ¦ WWPN - 0000000000000000 LUN -  0
          No.6 Target: Disabled ¦ WWPN - 0000000000000000 LUN -  0
          No.7 Target: Disabled ¦ WWPN - 0000000000000000 LUN -  0
          No.8 Target: Disabled ¦ WWPN - 0000000000000000 LUN -  0

                   Target Information Summary
          [Enter]:Enter; [↑|↓]:Next Entry; [ESC]:Quit Menu
  Bus=02 Device=00 Func=00 WWPN:2001001018A2FAF1 WWNN:2000001018A2FAF1
```

Figure 14-7. CCM Target Information (After Configuration)

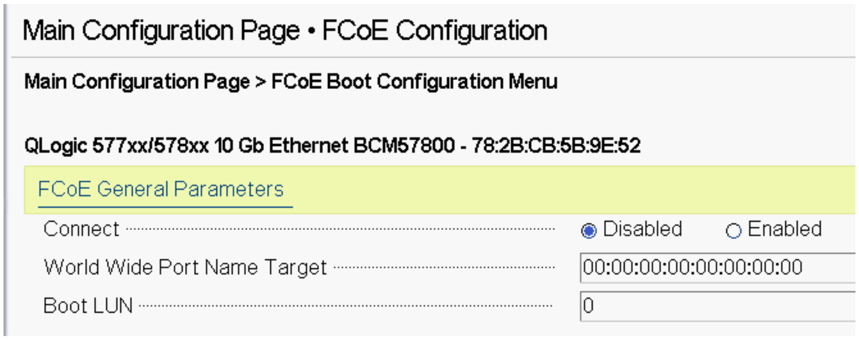
7. Press the ESC key until prompted to exit and save changes. To exit CCM, restart the system, and apply changes, press the CTRL+ALT+DEL keys.
8. Proceed to OS installation after storage access has been provisioned in the SAN.

Preparing Marvell Multiple Boot Agent for FCoE Boot (UEFI)

To prepare the Marvell multiple boot agent for FCOE boot (UEFI):

1. Enter the system BIOS UEFI device configuration page by pressing F2 during POST, and then select **Device Settings** (see Figure 6-2).
2. In the Device Settings menu (see Figure 6-3), select the desired device port.
3. In the Main Configuration Page menu, select **FCoE Configuration** (see Figure 6-4).

   The FCoE Boot Configuration Menu appears (see Figure 14-8).



Main Configuration Page • FCoE Configuration

Main Configuration Page > FCoE Boot Configuration Menu

QLogic 577xx/578xx 10 Gb Ethernet BCM57800 - 78:2B:CB:5B:9E:52

FCoE General Parameters

Connect ● Disabled ○ Enabled

World Wide Port Name Target 00:00:00:00:00:00:00:00

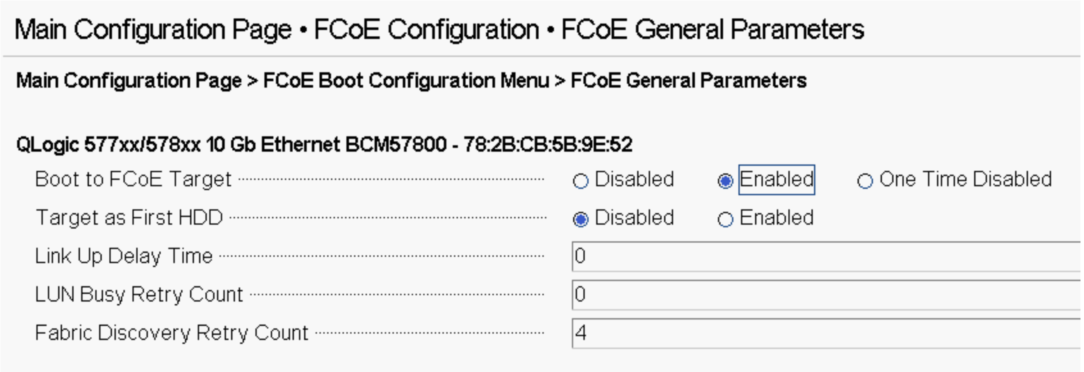
Boot LUN 0

Figure 14-8. FCoE Boot Configuration Menu

4. In the FCoE Boot Configuration menu:
   a. Select **Enabled** for the Connect field.
   b. Enter the World Wide Port Name Target.
   c. Enter the Boot LUN.

5. In the FCoE Configuration menu, select **FCoE General Parameters**.

   The FCoE General Parameters menu appears (see Figure 14-9).



Main Configuration Page • FCoE Configuration • FCoE General Parameters

Main Configuration Page > FCoE Boot Configuration Menu > FCoE General Parameters

QLogic 577xx/578xx 10 Gb Ethernet BCM57800 - 78:2B:CB:5B:9E:52

Boot to FCoE Target	○ Disabled	● Enabled	○ One Time Disabled
Target as First HDD	● Disabled	○ Enabled	
Link Up Delay Time	0		
LUN Busy Retry Count	0		
Fabric Discovery Retry Count	4		

Figure 14-9. FCoE Boot Configuration Menu, FCoE General Parameters

6. In the FCoE General Parameters menu:

   a. Select the desired Boot to FCoE Target mode (see One-Time Disabled). For initial OS installation to a blank FCoE target LUN from a CD/DVD-ROM or mounted bootable OS installation image, set **Boot from Target** to **One Time Disabled**.

      This setting prevents the system from booting from the configured FCoE target after establishing a successful login and connection. This setting reverts to Enabled after the next system reboot.

      The Enabled setting allows the system to connect to an FCoE target and attempt to boot from it.

      The Disabled setting allows the system to connect to an FCoE target and restricts it from booting from that device. Instead, it hands off the boot vector to the next bootable device in the boot sequence.

   b. Select the desired Target as First HDD mode.

      This setting specifies that the selected FCoE target drive appears as the first hard drive in the system.

   c. Select the desired LUN Busy Retry Count value.

      This value controls the number of connection retries the FCoE boot initiator attempts if the FCoE target LUN is busy.

   d. Select the desired Fabric Discovery Retry Count value.

      This value controls the number of connection retries the FCoE Boot initiator attempts if the FCoE Fabric is busy.

Provisioning Storage Access in the SAN

Storage access consists of zone provisioning and storage selective LUN presentation, each of which is commonly provisioned per initiator WWPN. Two main paths are available for approaching storage access:

- Pre-provisioning
- CTRL+R Method

Pre-provisioning

With pre-provisioning, note the initiator WWPN and manually modify fabric zoning and storage selective LUN presentation to allow the appropriate access for the initiator.

The initiator WWPN can be seen at the bottom of the pane in the FCoE boot target configuration window.

The initiator WWPN can also be directly inferred from the FIP MAC address associated with the interface(s) planned for boot. Two MAC addresses are printed on stickers attached to the SFP+ cage on your adapter. The FIP MAC ends in an odd digit. The WWPN is 20:00: + <FIP MAC>. For example, if the FIP MAC is 00:10:18:11:22:33, the WWPN is 20:00:00:10:18:11:22:33.

NOTE

The default WWPN is 20:00: + <FIP MAC>. The default WWNN is 10:00: + <FIP MAC>.

NOTE

In Dell FlexAddress™ configurations, the SAN or FIP MAC may be overridden by the blade chassis management system.

CTRL+R Method

The CTRL+R method allows you to use the boot initiator to bring up the link and log in to all available fabrics and targets. Using this method, you can ensure that the initiator is logged into the fabric or target before making provisioning changes, and as such, can provision without manually typing in WWPNs.

1. Configure at least one boot target through CCM as described in Pre-provisioning.
2. Allow the system to attempt to boot through the selected initiator.

When the initiator boot starts, it begins DCBX sync, FIP Discovery, Fabric Login, Target Login, and LUN readiness checks. As each of these phases completes, if the initiator is unable to proceed to the next phase, MBA presents the option to press the CTRL+R keys.

3. Press the CTRL+R keys.
4. When CTRL+R has been activated, the boot initiator maintains a link in whatever phase has most recently succeeded and allows you time to make the necessary provisioning corrections to proceed to the next phase.
5. If the initiator logs into the fabric, but is unable to log into the target, a CTRL+R pauses the boot process and allows you to configure fabric zoning.

   When zoning is complete, the initiator automatically logs into all visible targets.
6. If the initiator is unable to discover the designated LUN on the designated target as provisioned in Step 1, CTRL+R pauses the boot process and allows you to configure selective LUN presentation.
7. The boot initiator periodically polls the LUN for readiness, and when you have provisioned access to the LUN, the boot process automatically proceeds.

NOTE

Ensure that you also put the boot initiator into one-time disabled mode as described in One-Time Disabled.

One-Time Disabled

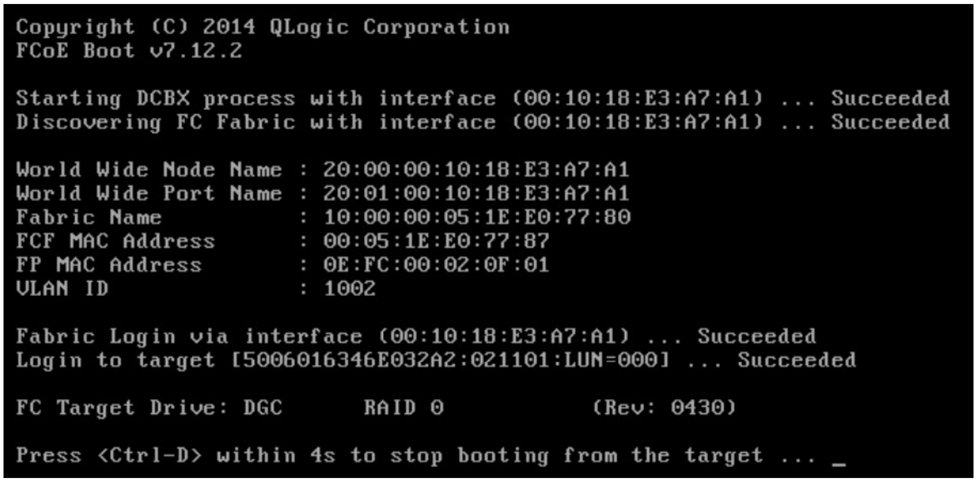
The Marvell FCoE ROM is implemented as Boot Entry Vector (BEV). In this implementation, the Option ROM only connects to the target when it has been selected by BIOS as the chosen boot device. This approach is different from other implementations that will connect to the boot device even if another device has been selected by the system BIOS.

For OS installation over the FCoE path, you must instruct the Option ROM to bypass FCoE and skip to CD or DVD installation media. As instructed in "Preparing Marvell Multiple Boot Agent for FCoE Boot (CCM)" on page 211, the boot order must be configured with Marvell boot first and installation media second. Furthermore, during OS installation, it is necessary to bypass the FCoE boot and pass through to the installation media for boot. To do so, perform one-time disable of the FCoE boot ROM from booting, and not by simply allowing the FCoE ROM to attempt to boot and allowing the BIOS to fail through and boot the installation media. Finally, the FCoE ROM must successfully discover and test the readiness of the boot LUN for installation to proceed successfully. Failure to allow the boot ROM to discover the LUN and do a coordinated bypass will result in a failure to properly install the OS to the LUN.

The two choices for the coordinated bypass include:

- When the FCoE boot ROM discovers a ready target LUN, it prompts you to press the CTRL+D keys within four seconds to **Stop booting from the target**. Press CTRL+D, and proceed to boot from the installation media.
- From CCM, set the **Option ROM** setting under MBA settings to **One Time Disabled**. With this setting, the FCoE ROM loads and automatically bypasses when the ready LUN is discovered. On the subsequent reboot after installation, the option ROM will automatically revert to **Enabled**.

Wait through all option ROM banners. When FCoE Boot is invoked, it connects to the target, and provides a four-second window to press the CTRL+D keys to invoke the bypass, as shown in Figure 14-10. Press CTRL+D to proceed to installation.



```
Copyright (C) 2014 QLogic Corporation
FCoE Boot v7.12.2

Starting DCBX process with interface (00:10:18:E3:A7:A1) ... Succeeded
Discovering FC Fabric with interface (00:10:18:E3:A7:A1) ... Succeeded

World Wide Node Name : 20:00:00:10:18:E3:A7:A1
World Wide Port Name : 20:01:00:10:18:E3:A7:A1
Fabric Name          : 10:00:00:05:1E:E0:77:80
FCF MAC Address      : 00:05:1E:E0:77:87
FP MAC Address       : 0E:FC:00:02:0F:01
VLAN ID              : 1002

Fabric Login via interface (00:10:18:E3:A7:A1) ... Succeeded
Login to target [5006016346E032A2:021101:LUN=000] ... Succeeded

FC Target Drive: DGC      RAID 0           (Rev: 0430)

Press <Ctrl-D> within 4s to stop booting from the target ... _
```

Figure 14-10. FCoE Boot

Windows Server 2016/2019/Azure Stack HCI FCoE Boot Installation

Windows Server 2016/2019/Azure Stack HCI boot from SAN installation requires the use of a “slipstream” DVD or ISO image with the latest Marvell drivers injected (see “Injecting (Slipstreaming) Marvell Drivers into Windows Image Files” on page 122). Also, refer to the Microsoft Knowledge Base topic KB974072 at support.microsoft.com, which is also helpful for Windows Server 2016/2019/Azure Stack HCI FCoE Boot from SAN. Microsoft's procedure injects only OIS, VBD, and NDIS drivers. Marvell strongly recommends injecting all drivers, especially those listed below in **bold**:

- **EVBD (Core)**
- VBD (Core)
- BXND (Ethernet or NDIS)
- BXOIS (iSCSI offload)
- **BXFCoE** (FCoE offload)

When you have a properly slipstreamed ISO, you can use that ISO for normal Windows Server 2016/2019/Azure Stack HCI and later installation, without needing USB-provided drivers.

Linux FCoE Boot Installation

Configure the adapter boot parameters and Target Information (press CTRL+S and enter the CCM utility) as detailed in “Preparing System BIOS for FCoE Build and Boot” on page 210. Then, use the guidelines in the following sections for FCoE boot installation with the appropriate Linux version:

- SLES 12 Through SLES 15 Installation
- Booting from RHEL 7.x Installation Media With the FCoE Target Already Connected

SLES 12 Through SLES 15 Installation

1. To start the installation:

   a. Boot from the SLES installation medium.

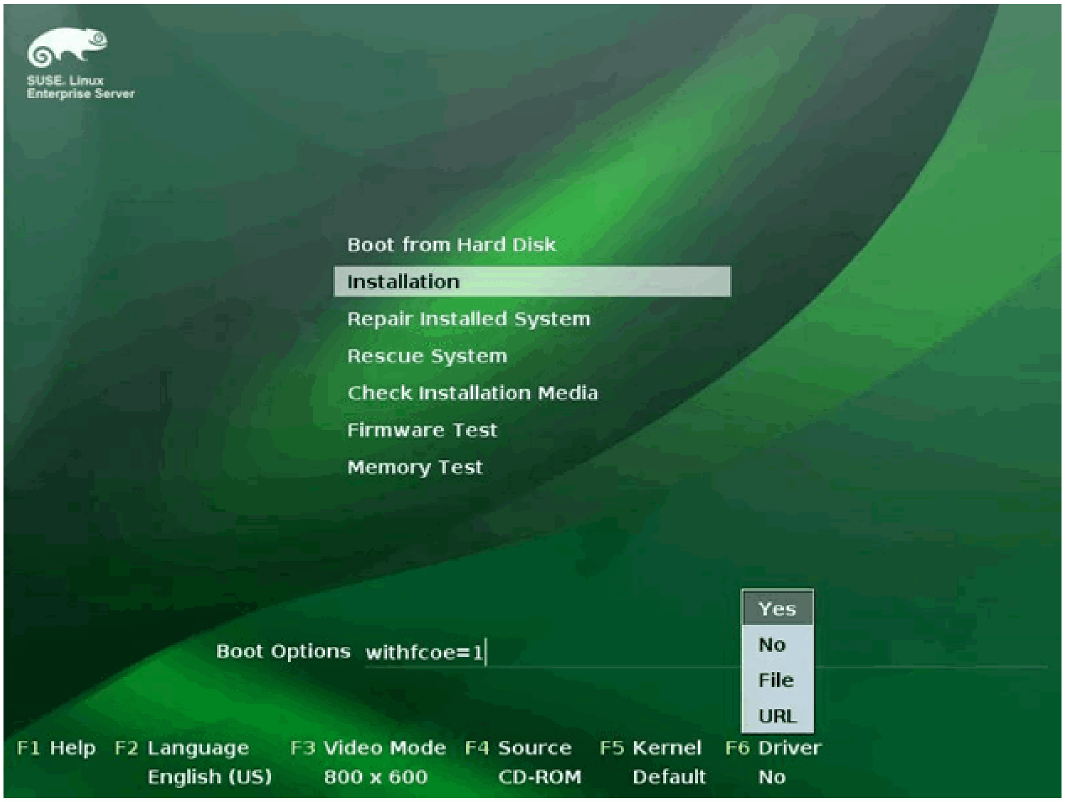
   b. On the installation splash window, press the F6 key for driver update disk.

   c. Select **Yes**.

   d. In **Boot Options**, type one of the following:

      - **`withfcoe=1`** (before SLES 15)
      - **`withfoce=1 dud=1`** (SLES 15)

e. Click **Installation** to proceed (Figure 14-11).



Boot Options withfcoe=1

F2 Language English (US)

F3 Video Mode 800 x 600

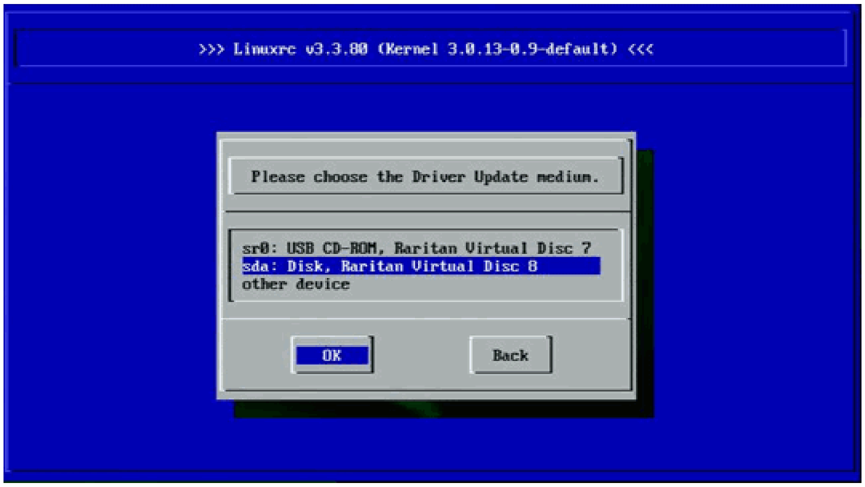
F4 Source CD-ROM

F5 Kernel Default

F6 Driver No

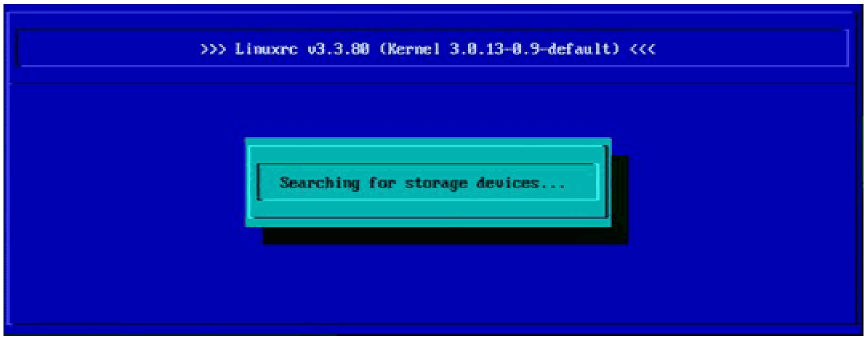
Figure 14-11. Starting SLES Installation

2. Follow the prompts to choose the driver update medium (Figure 14-12) and load the drivers (Figure 14-13).



```
>>> Linuxrc v3.3.80 (Kernel 3.0.13-0.9-default) <<<

Please choose the Driver Update medium.

sr0: USB CD-ROM, Raritan Virtual Disc 7
sda: Disk, Raritan Virtual Disc 8
other device
```

Figure 14-12. Selecting Driver Update Medium



```
>>> Linuxrc v3.3.80 (Kernel 3.0.13-0.9-default) <<<

Searching for storage devices...
```

Figure 14-13. Loading the Drivers

3. After the driver update is complete, select **Next** to continue with OS installation.

4. When requested, click **Configure FCoE Interfaces** (Figure 14-14).

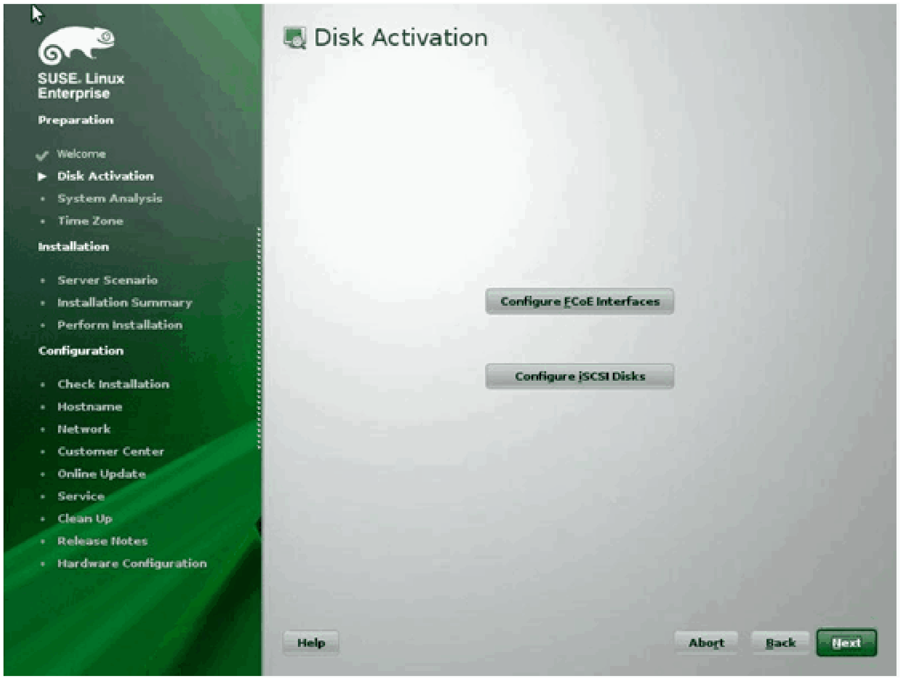
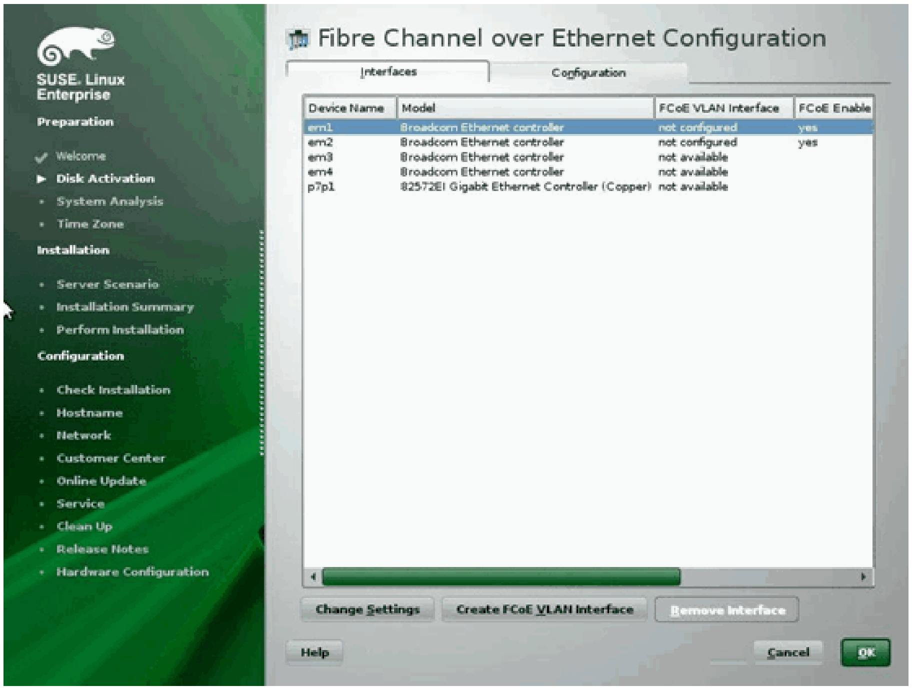


Figure 14-14. Activating the Disk

5. Ensure that **FCoE Enable** is set to **yes** on the 10GbE Marvell initiator ports that you want to use as the SAN boot paths (Figure 14-15).



Device Name	Model	FCoE VLAN Interface	FCoE Enable
em1	Broadcom Ethernet controller	not configured	yes
em2	Broadcom Ethernet controller	not configured	yes
em3	Broadcom Ethernet controller	not available	
em4	Broadcom Ethernet controller	not available	
p7p1	82572EI Gigabit Ethernet Controller (Copper)	not available	

Figure 14-15. Enabling FCoE

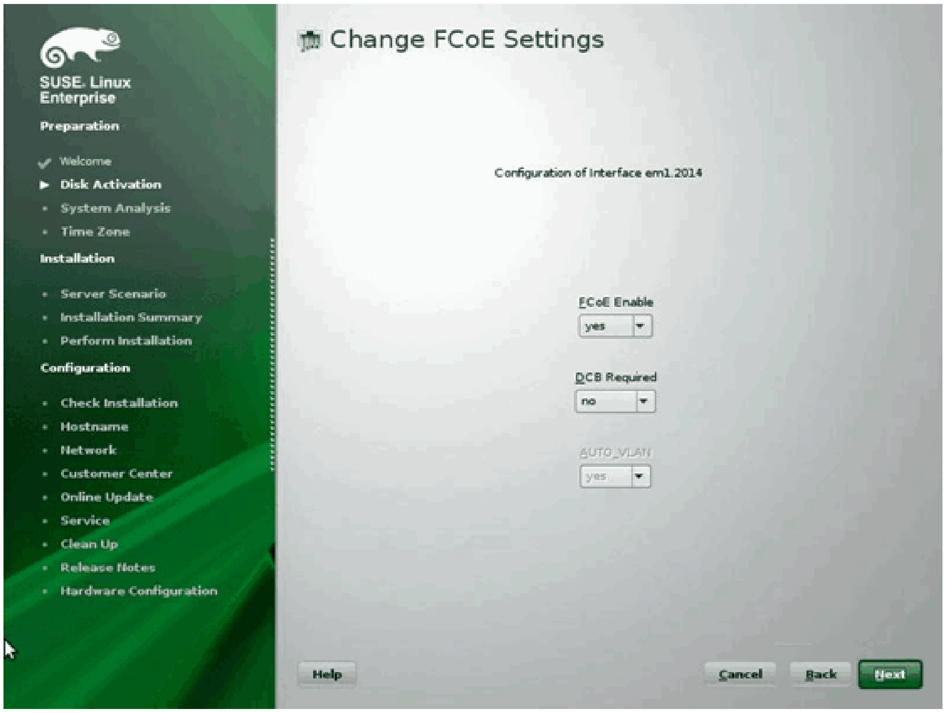
6. For each interface to be enabled for FCoE boot:

   a. Click **Change Settings**.

   b. On the Change FCoE Settings window (Figure 14-16), ensure that **FCoE Enable** and **Auto_VLAN** are set to **yes**.

   c. Ensure that **DCB Required** is set to **no**.

   d. Click **Next** to save the settings.



Change FCoE Settings

Configuration of Interface em1.2014

FCoE Enable: yes

DCB Required: no

AUTO_VLAN: yes

Figure 14-16. Changing FCoE Settings

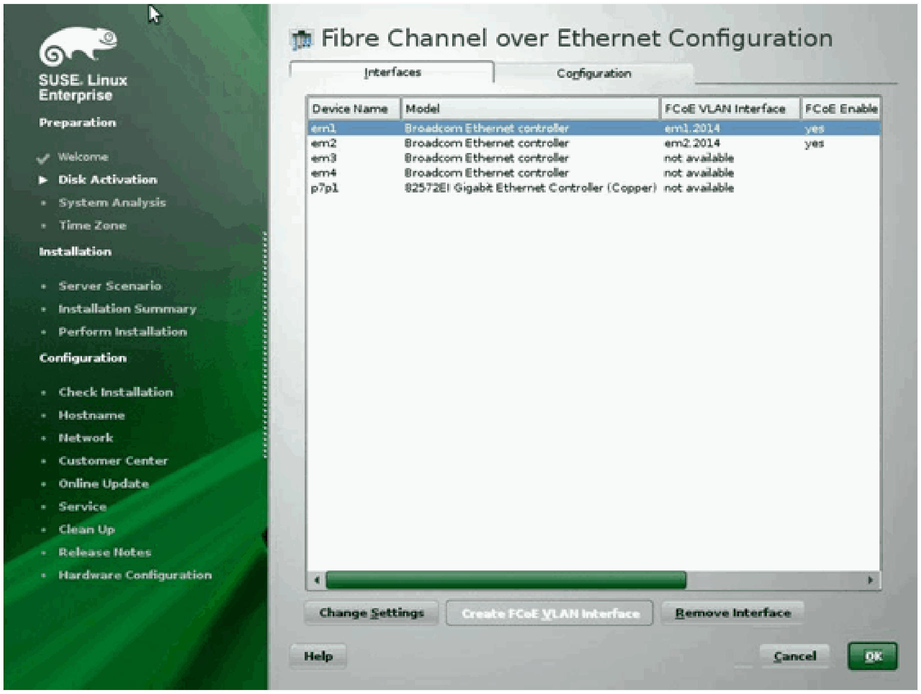
7. For each interface to be enabled for FCoE boot:

   a. Click **Create FCoE VLAN Interface**.

   b. On the VLAN interface creation dialog box, click **Yes** to confirm and trigger automatic FIP VLAN discovery.

   If successful, the VLAN is displayed under **FCoE VLAN Interface**. If no VLAN is visible, check your connectivity and switch configuration.

8. After completing the configuration of all interfaces, click **OK** to proceed (Figure 14-17).



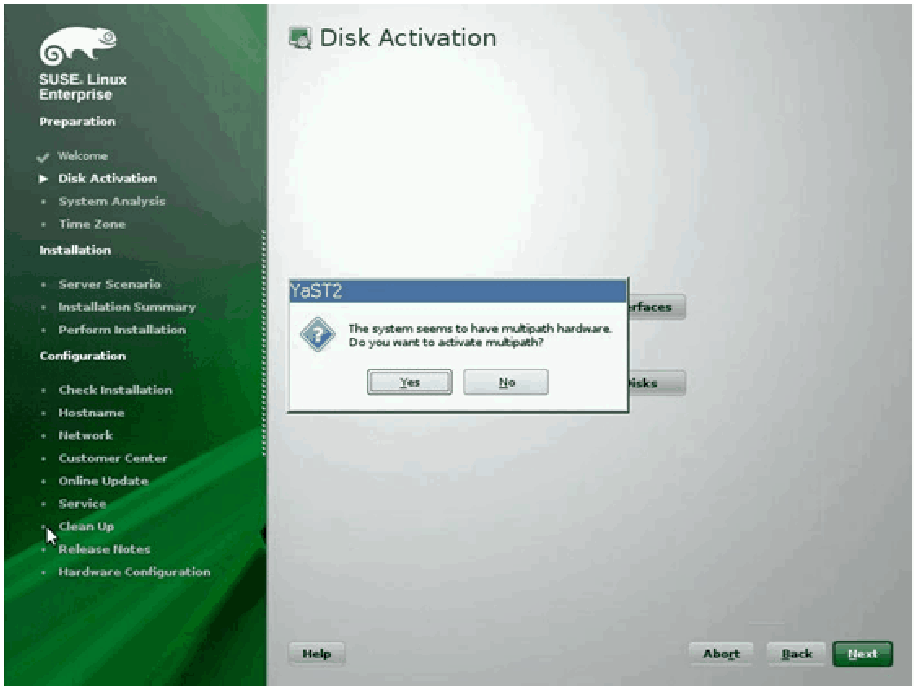
Fibre Channel over Ethernet Configuration

Device Name	Model	FCoE VLAN Interface	FCoE Enable
em1	Broadcom Ethernet controller	em1.2014	yes
em2	Broadcom Ethernet controller	em2.2014	yes
em3	Broadcom Ethernet controller	not available	
em4	Broadcom Ethernet controller	not available	
p7p1	82572EI Gigabit Ethernet Controller (Copper)	not available	

Figure 14-17. FCoE Interface Configuration

9. Click **Next** to continue installation.

10. YaST2 prompts you to activate multipath. Answer as appropriate (Figure 14-18).

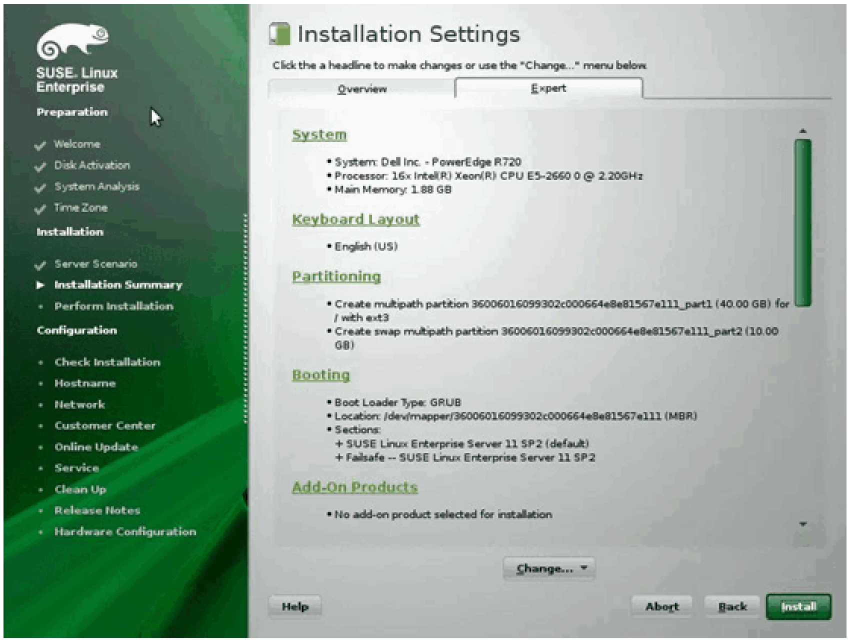


The system seems to have multipath hardware.
Do you want to activate multipath?

Figure 14-18. Disk Activation

11. Continue installation as usual.

12. On the Expert page on the Installation Settings window, click **Booting** (Figure 14-19).



Installation Settings

Click the a headline to make changes or use the "Change..." menu below.

System

- System: Dell Inc. - PowerEdge R720
- Processor: 16x Intel(R) Xeon(R) CPU E5-2660 0 @ 2.20GHz
- Main Memory: 1.88 GB

Keyboard Layout

- English (US)

Partitioning

- Create multipath partition 36006016099302c000664e8e81567e111_part1 (40.00 GB) for / with ext3
- Create swap multipath partition 36006016099302c000664e8e81567e111_part2 (10.00 GB)

Booting

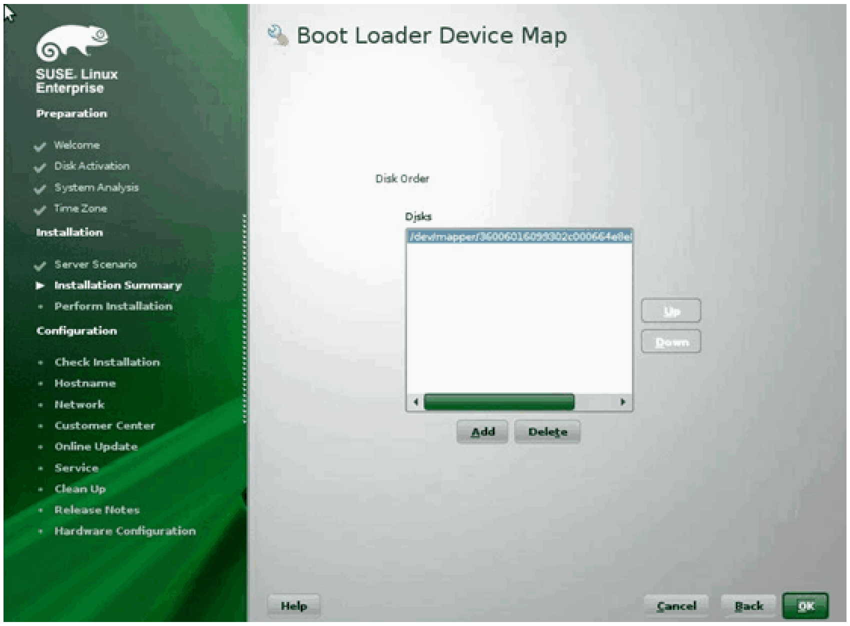
- Boot Loader Type: GRUB
- Location: /dev/mapper/36006016099302c000664e8e81567e111 (MBR)
- Sections:
  + SUSE Linux Enterprise Server 11 SP2 (default)
  + Failsafe -- SUSE Linux Enterprise Server 11 SP2

Add-On Products

- No add-on product selected for installation

Figure 14-19. Installation Settings

13. Click the **Boot Loader Installation** tab, and then select **Boot Loader Installation Details**. Make sure you have one boot loader entry here; delete all redundant entries (Figure 14-20).



/dev/mapper/36006016099302c000664e8e

Figure 14-20. Boot Loader Device Map

14. Click **OK** to proceed and complete installation.

Booting from RHEL 7.*x* Installation Media With the FCoE Target Already Connected

To install Linux FCoE boot on RHEL 7.*x*:

1. Boot from the RHEL 7.*x* installation media with the FCoE target already connected.

```
Install Red Hat Enterprise Linux 7.x
Test this media & install Red Hat Enterprise 7.x
Troubleshooting -->

Use the UP and DOWN keys to change the selection
Press 'e' to edit the selected item or 'c' for a command
Prompt
```

2. Select **Troubleshooting**. Use the UP and DOWN keys to change the selection. Select **e** to edit the selected item or **c** for a command prompt.

3. To install an out-of-box driver, press the E key.

4. Select the kernel line, and then press the E key to edit the line.
5. Issue the following command, and then press ENTER:

   ```
   inst.dd modprobe.blacklist=bnx2x,bnx2fc,bnx2i,cnic
   ```

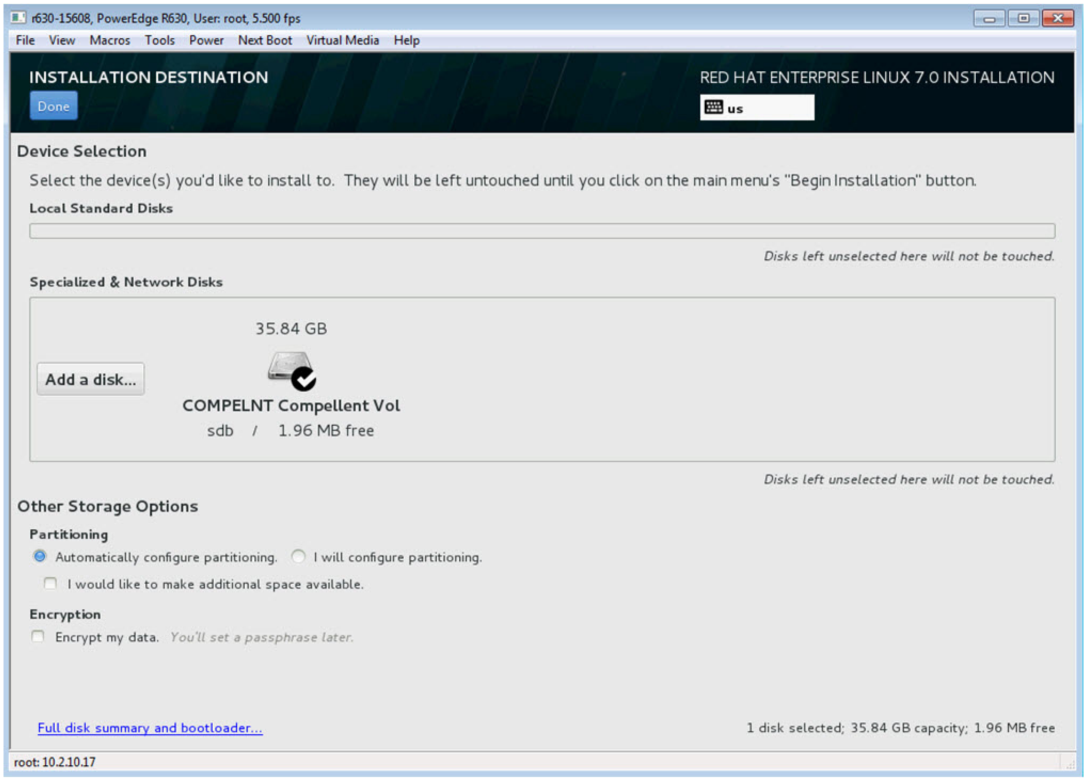
6. At the **Driver disk device selection** prompt:
   a. Refresh the device list by pressing the R key.
   b. Type the appropriate number for your media.
   c. Press the C key to continue.

   NOTE

   RHEL does not allow the driver update media to be loaded through the network when installing driver updates for network devices. Use local media.

7. After the drivers are loaded, proceed with the installation by pressing C.
8. On the Installation Summary window, click **Installation Destination**.
9. On the Installation Destination window under **Specialized & Network Disks**, click **Add a disk**.
10. On the Search page, click **Add FCoE SAN**.
11. Complete the Please Select the Network Interface... dialog box as follows:
    a. Select the appropriate **NIC**.
    b. Clear the **Use DCB** check box.
    c. Click **Add FCoE Disk(s)**.
12. On the Search page, select the newly added disk, and then click **Done**.

13. On the Installation Destination window (Figure 14-21) under **Other Storage Options**, select your **Partitioning** options, and then click **Done**.


r630-15608, PowerEdge R630, User: root, 5.500 fps
File View Macros Tools Power Next Boot Virtual Media Help
Done
us
Add a disk...
root: 10.2.10.17

INSTALLATION DESTINATION

RED HAT ENTERPRISE LINUX 7.0 INSTALLATION

Device Selection

Select the device(s) you'd like to install to. They will be left untouched until you click on the main menu's "Begin Installation" button.

Local Standard Disks

Disks left unselected here will not be touched.

Specialized & Network Disks

35.84 GB
COMPELNT Compellent Vol
sdb / 1.96 MB free

Disks left unselected here will not be touched.

Other Storage Options

Partitioning

● Automatically configure partitioning. ○ I will configure partitioning.

□ I would like to make additional space available.

Encryption

□ Encrypt my data. *You'll set a passphrase later.*

Full disk summary and bootloader...

1 disk selected; 35.84 GB capacity; 1.96 MB free

Figure 14-21. Selecting Partitioning Options

14. On the Installation Summary window, click **Begin Installation**.

Linux: Adding Boot Paths

RHEL requires updates to the network configuration when adding new boot through an FCoE initiator that was not configured during installation.

RHEL 6.2 and Later

On RHEL 6.2 and later, if the system is configured to boot through an initiator port that has not previously been configured in the OS, the system automatically boots successfully, but will encounter problems during shutdown. All new boot path initiator ports must be configured in the OS before updating pre-boot FCoE boot parameters.

1. Identify the network interface names for the newly added interfaces through **`ifconfig -a`**.

2. Edit `/boot/grub/menu.lst` by adding **`ifname=<INTERFACE>:<MAC_ADDRESS>`** to the line `kernel /vmlinuz …` for each new interface. The MAC address must be all lower case and separated by a colon. (for example, `ifname=em1:00:00:00:00:00:00`)

3. Create a **/etc/fcoe/cfg-<INTERFACE>** file for each new FCoE initiator by duplicating the /etc/fcoe/cfg-<INTERFACE> file that was already configured during initial installation.

4. Issue the following command:

```
nm-connection-editor
```

   a. Open **Network Connection** and choose each new interface.

   b. Configure each interface as needed, including DHCP settings.

   c. Click **Apply** to save.

5. For each new interface, edit **/etc/sysconfig/network-scripts/ifcfg-<INTERFACE>** to add the line **NM_CONTROLLED="no"**. Modifying these files automatically causes a restart to the network service, which may cause the system to appear to hang briefly. Marvell recommends that you ensure that redundant multipath paths are available before performing this operation.

VMware ESXi FCoE Boot Installation

FCoE Boot from SAN requires that the latest Marvell 57*xx* and 57*xxx* async drivers be included into the ESXi install image. Refer to Image_builder_doc.pdf from VMware on how to slipstream drivers. Table 14-1 shows the supported Legacy BFS and uEFI BFS.

Table 14-1. Supported Legacy BFS and uEFI BFS

Version	Legacy BFS	uEFI BFS
ESXi 6.7	Supported	Not Supported
ESXi 7.0	Supported	Not Supported

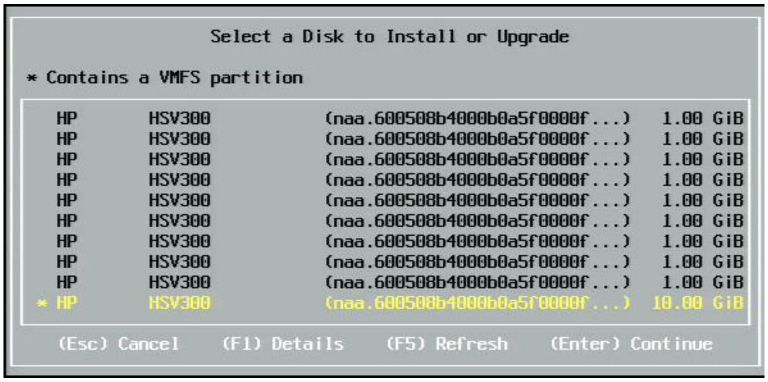
To install ESXi FCoE boot:

1. Boot from the updated ESXi installation image and select the appropriate ESXi installer when prompted.

2. On the Welcome to the VMware ESXi installation window, press the ENTER key to continue.

3. On the EULA window, press the F11 key to accept the agreement and continue.

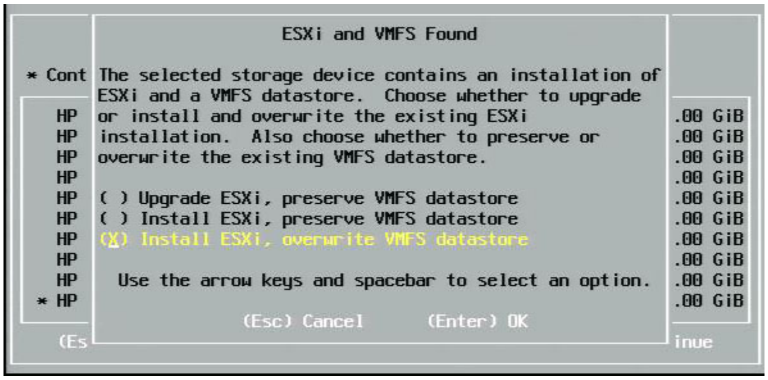
4. On the Select a Disk window (Figure 14-22), scroll to the boot LUN for installation, and then press ENTER to continue.



```
                Select a Disk to Install or Upgrade

 * Contains a VMFS partition

    HP       HSV300           (naa.600508b4000b0a5f0000f...)   1.00 GiB
    HP       HSV300           (naa.600508b4000b0a5f0000f...)   1.00 GiB
    HP       HSV300           (naa.600508b4000b0a5f0000f...)   1.00 GiB
    HP       HSV300           (naa.600508b4000b0a5f0000f...)   1.00 GiB
    HP       HSV300           (naa.600508b4000b0a5f0000f...)   1.00 GiB
    HP       HSV300           (naa.600508b4000b0a5f0000f...)   1.00 GiB
    HP       HSV300           (naa.600508b4000b0a5f0000f...)   1.00 GiB
    HP       HSV300           (naa.600508b4000b0a5f0000f...)   1.00 GiB
    HP       HSV300           (naa.600508b4000b0a5f0000f...)   1.00 GiB
  * HP       HSV300           (naa.600508b4000b0a5f0000f...)  10.00 GiB

    (Esc) Cancel    (F1) Details    (F5) Refresh    (Enter) Continue
```

Figure 14-22. ESXi Disk Selection

5. On the ESXi and VMFS Found window (Figure 14-23), select the installation method.



```
                    ESXi and VMFS Found

 * Cont The selected storage device contains an installation of
        ESXi and a VMFS datastore.  Choose whether to upgrade
    HP  or install and overwrite the existing ESXi             .00 GiB
    HP  installation.  Also choose whether to preserve or      .00 GiB
    HP  overwrite the existing VMFS datastore.                 .00 GiB
    HP                                                         .00 GiB
    HP  ( ) Upgrade ESXi, preserve VMFS datastore              .00 GiB
    HP  ( ) Install ESXi, preserve VMFS datastore              .00 GiB
    HP  (X) Install ESXi, overwrite VMFS datastore             .00 GiB
    HP                                                         .00 GiB
    HP    Use the arrow keys and spacebar to select an option. .00 GiB
  * HP                                                         .00 GiB
                  (Esc) Cancel        (Enter) OK
    (Es                                                        inue
```

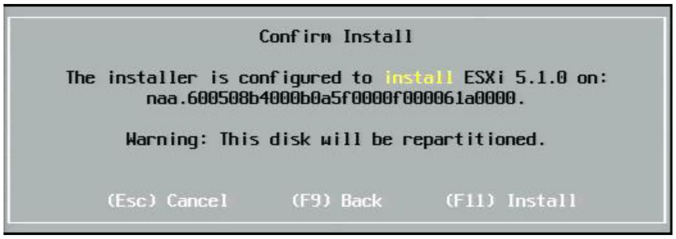
Figure 14-23. ESXi and VMFS Found

6. Follow the prompts to:

   a. Select the keyboard layout.

   b. Enter and confirm the root password.

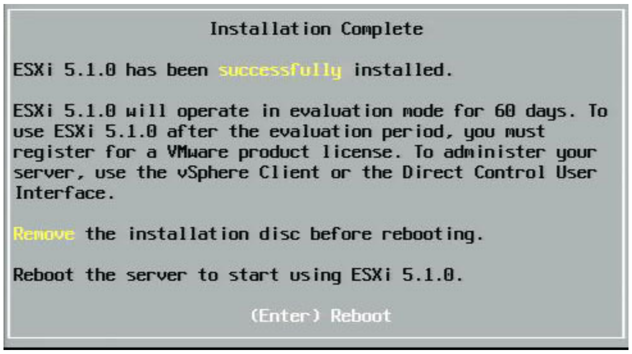
7. On the Confirm Install window (Figure 14-24), press the F11 key to confirm the installation and repartition.



```
                         Confirm Install

      The installer is configured to install ESXi 5.1.0 on:
             naa.600508b4000b0a5f0000f000061a0000.

             Warning: This disk will be repartitioned.


         (Esc) Cancel        (F9) Back        (F11) Install
```

Figure 14-24. ESXi Confirm Install

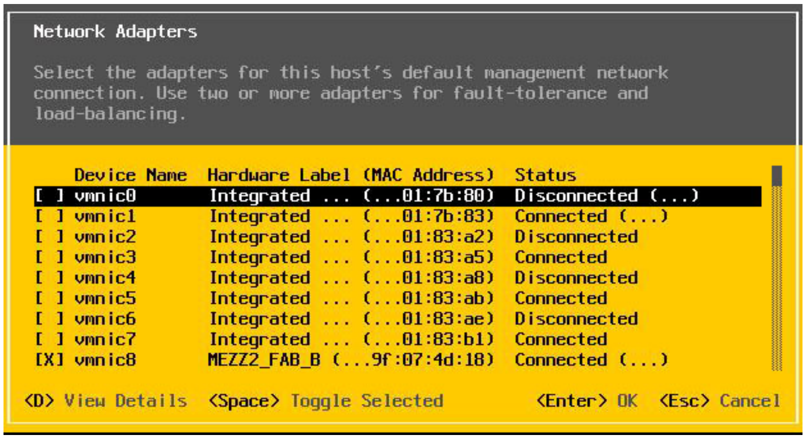
8. After successful installation (Figure 14-25), press ENTER to reboot.



```
                   Installation Complete

ESXi 5.1.0 has been successfully installed.

ESXi 5.1.0 will operate in evaluation mode for 60 days. To
use ESXi 5.1.0 after the evaluation period, you must
register for a VMware product license. To administer your
server, use the vSphere Client or the Direct Control User
Interface.

Remove the installation disc before rebooting.

Reboot the server to start using ESXi 5.1.0.

                       (Enter) Reboot
```

Figure 14-25. ESXi Installation Complete

9. On 57800 and 57810 boards, the management network is not vmnic0. After booting, open the GUI console and view the Configure Management Network, and then the Network Adapters window (Figure 14-26) to select the NIC to be used as the management network device.

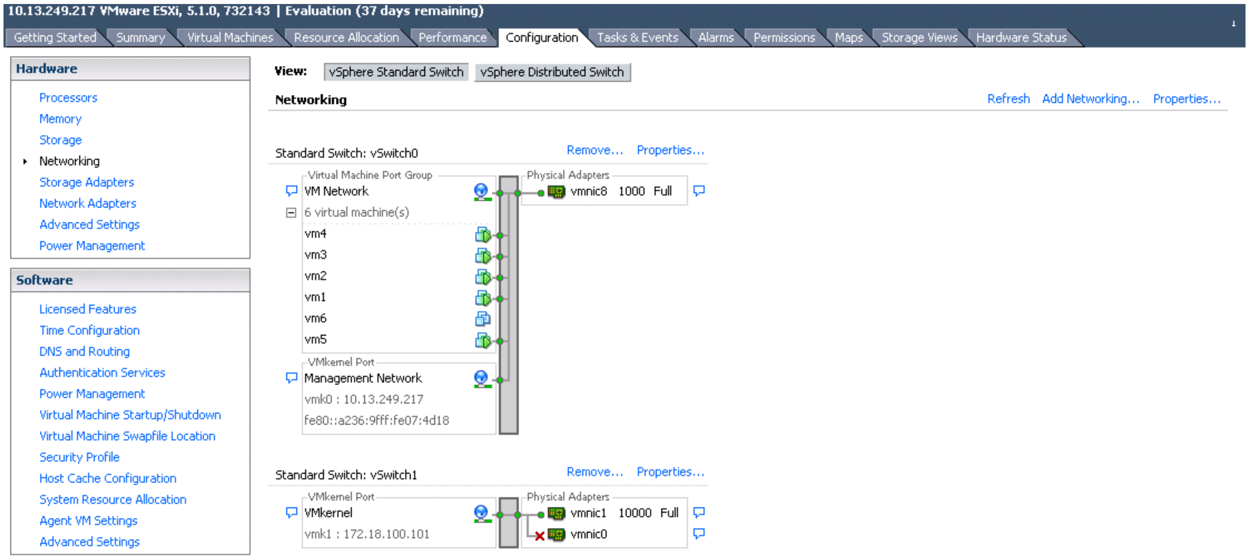


```
 Network Adapters

 Select the adapters for this host's default management network
 connection. Use two or more adapters for fault-tolerance and
 load-balancing.

     Device Name  Hardware Label (MAC Address)  Status
 [ ] vmnic0       Integrated ... (...01:7b:80)  Disconnected (...)
 [ ] vmnic1       Integrated ... (...01:7b:83)  Connected (...)
 [ ] vmnic2       Integrated ... (...01:83:a2)  Disconnected
 [ ] vmnic3       Integrated ... (...01:83:a5)  Connected
 [ ] vmnic4       Integrated ... (...01:83:a8)  Disconnected
 [ ] vmnic5       Integrated ... (...01:83:ab)  Connected
 [ ] vmnic6       Integrated ... (...01:83:ae)  Disconnected
 [ ] vmnic7       Integrated ... (...01:83:b1)  Connected
 [X] vmnic8       MEZZ2_FAB_B (...9f:07:4d:18)  Connected (...)

<D> View Details  <Space> Toggle Selected      <Enter> OK  <Esc> Cancel
```

Figure 14-26. ESXi Management Network Selection

10. For 57800 and 57810 boards, the FCoE boot devices must have a separate vSwitch other than vSwitch0. This switch allows DHCP to assign the IP address to the management network rather than to the FCoE boot device. To create a vSwitch for the FCoE boot devices, add the boot device vmnics in vSphere Client on the Configuration page under **Networking**.

    Figure 14-27 shows an example.



10.13.249.217 VMware ESXi, 5.1.0, 732143 | Evaluation (37 days remaining)

Getting Started | Summary | Virtual Machines | Resource Allocation | Performance | Configuration | Tasks & Events | Alarms | Permissions | Maps | Storage Views | Hardware Status

Hardware
- Processors
- Memory
- Storage
- Networking
- Storage Adapters
- Network Adapters
- Advanced Settings
- Power Management

Software
- Licensed Features
- Time Configuration
- DNS and Routing
- Authentication Services
- Power Management
- Virtual Machine Startup/Shutdown
- Virtual Machine Swapfile Location
- Security Profile
- Host Cache Configuration
- System Resource Allocation
- Agent VM Settings
- Advanced Settings

View: vSphere Standard Switch | vSphere Distributed Switch

Networking — Refresh | Add Networking... | Properties...

Standard Switch: vSwitch0 — Remove... | Properties...

Virtual Machine Port Group
VM Network
6 virtual machine(s)
vm4
vm3
vm2
vm1
vm6
vm5

VMkernel Port
Management Network
vmk0 : 10.13.249.217
fe80::a236:9fff:fe07:4d18

Physical Adapters
vmnic8 1000 Full

Standard Switch: vSwitch1 — Remove... | Properties...

VMkernel Port
VMkernel
vmk1 : 172.18.100.101

Physical Adapters
vmnic1 10000 Full
vmnic0

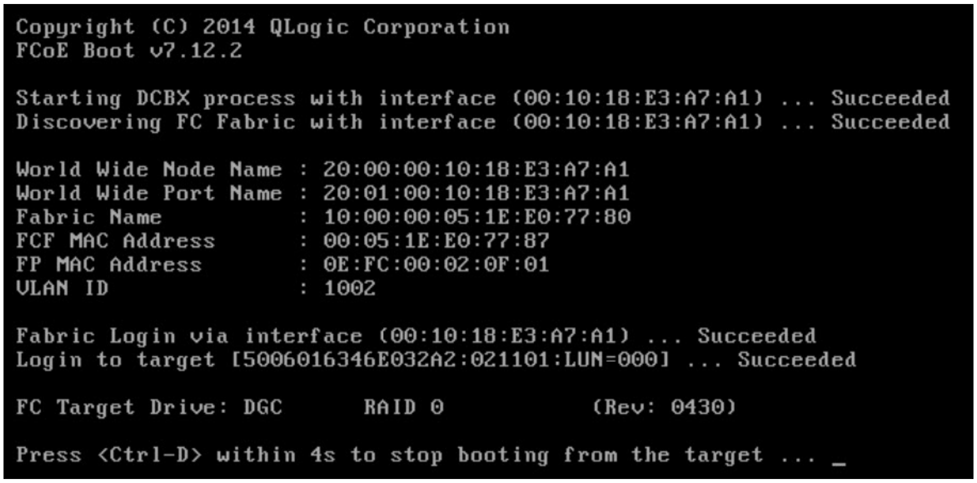
Figure 14-27. VMware vSphere Client Network Configuration

Configuring FCoE Boot from SAN on VMware

Note that each host must have access only to its own boot LUN—not to the boot LUNs of other hosts. Use storage system software to ensure that the host accesses only the designated LUNs.

Booting from SAN After Installation

After boot configuration and OS installation are complete, you can reboot and test the installation. On this and all future reboots, no other user interactivity is required. Ignore the CTRL+D prompt and allow the system to boot through to the FCoE SAN LUN, as shown in Figure 14-28.



```
Copyright (C) 2014 QLogic Corporation
FCoE Boot v7.12.2

Starting DCBX process with interface (00:10:18:E3:A7:A1) ... Succeeded
Discovering FC Fabric with interface (00:10:18:E3:A7:A1) ... Succeeded

World Wide Node Name : 20:00:00:10:18:E3:A7:A1
World Wide Port Name : 20:01:00:10:18:E3:A7:A1
Fabric Name          : 10:00:00:05:1E:E0:77:80
FCF MAC Address      : 00:05:1E:E0:77:87
FP MAC Address       : 0E:FC:00:02:0F:01
VLAN ID              : 1002

Fabric Login via interface (00:10:18:E3:A7:A1) ... Succeeded
Login to target [5006016346E032A2:021101:LUN=000] ... Succeeded

FC Target Drive: DGC      RAID 0           (Rev: 0430)

Press <Ctrl-D> within 4s to stop booting from the target ... _
```

Figure 14-28. Booting from SAN After Installation

If additional redundant failover paths are then needed, you can configure those paths through CCM, and the MBA will automatically failover to secondary paths if the first path is not available. In addition, the redundant boot paths yield redundant paths visible through host MPIO software to provide a fault-tolerant configuration.

Driver Upgrade on Linux Boot from SAN Systems

1. Remove the existing installed 57*xx* and 57*xxx* package as follows:

   a. Log in as root.

   b. Query for the existing 57*xx* and 57*xxx* package.

   c. Remove it by issuing the following commands:

      # rpm -e <57*xx* and 57*xxx* package name>

      For example:

      rpm -e netxtreme2

      or:

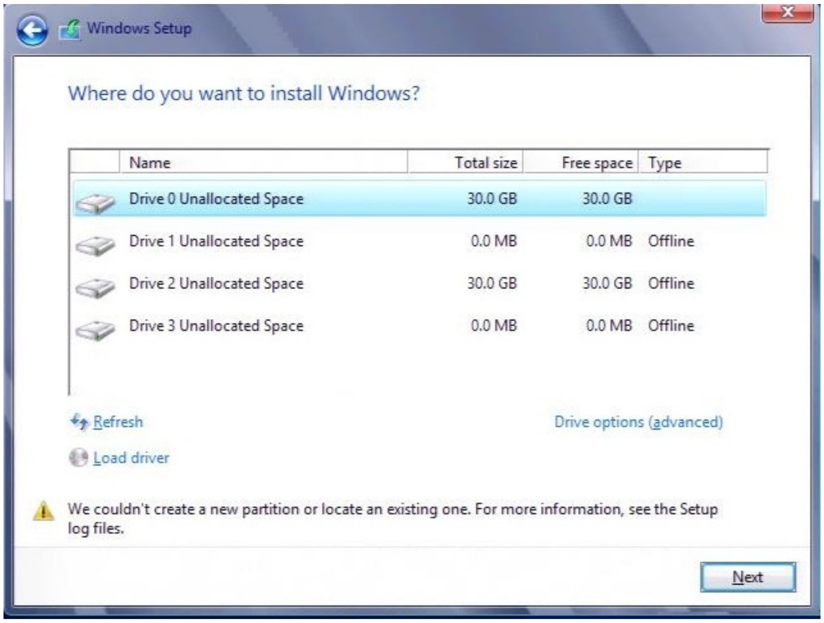
      rpm -e netxtreme2-x.y.z-1.x86_64

2. Install the binary RPM containing the new driver version. Refer to the linux-nx2 package `README` file for instructions on how to prepare a binary driver RPM.

3. Issue the following command to update the ramdisk (not required for SLES 12 and later, or RHEL 7.*x* and later):
   - On RHEL 6.*x* systems, issue: **`dracut -force`**
   - On SLES 11 SPX systems, issue: **`mkinitrd`**
4. If you are using different name for the initrd under `/boot`:
   a. Overwrite it with the default, because `dracut/mkinitrd` updates the ramdisk with the default original name.
   b. Verify that your appropriate entry for the boot from SAN setup uses the correct or updated intrd name in `/boot/grub/menu.lst`.
5. To complete your driver upgrade, reboot the system and select the modified grub boot entry that contains the updated initrd.

Errors During Windows FCoE Boot from SAN Installation

If any USB flash drive is connected while Windows setup is loading files for installation, an error message will appear when you provide the drivers and then select the SAN disk for the installation. The most common error message that Windows OS installer reports is, "*We couldn't create a new partition or locate an existing one. For more information, see the setup log files*" (see Figure 14-29).



Where do you want to install Windows?

Name	Total size	Free space	Type
Drive 0 Unallocated Space	30.0 GB	30.0 GB	
Drive 1 Unallocated Space	0.0 MB	0.0 MB	Offline
Drive 2 Unallocated Space	30.0 GB	30.0 GB	Offline
Drive 3 Unallocated Space	0.0 MB	0.0 MB	Offline

We couldn't create a new partition or locate an existing one. For more information, see the Setup log files.

Figure 14-29. Windows Partition Error Message

In other cases, the error message may indicate a need to ensure that the disk's controller is enabled in the computer's BIOS menu.

To avoid any of the preceding error messages, you must ensure that there is no USB flash drive attached until the setup asks for the drivers. When you load the drivers and see your SAN disks, detach or disconnect the USB flash drive immediately before selecting the disk for further installation.

Configuring FCoE

By default, DCB is enabled on 57712/578*xx* FCoE-, DCB-compatible C-NICs. The 57712/578*xx* FCoE requires a DCB-enabled interface. For Windows operating systems, use one of the following to configure the DCB parameters:

- QCC GUI
- QCC PowerKit
- QLogic Control Suite (QCS) CLI
- Server BIOS UEFI HII device configuration page
- Marvell Comprehensive Configuration Management (CCM) utility

For more information on QCS CLI, see the *User's Guide, QLogic Control Suite CLI*, part number BC0054511-00, available from Marvell.

For FCoE offload, the 57712/578*xx* adapters should have FCoE offload and DCB enabled.

- For all OSs, use Marvell's pre-boot CCM utility or the server's pre-boot BIOS UEFI HII device configuration page to configure the DCB parameters.
  - ❑ For FCoE on the VMware OS, see the FCoE Support section in the *User's Guide, Converged Network Adapters and Intelligent Ethernet Adapters, QLogic FastLinQ 3400 and 8400 Series* (part number 83840-546-00). To locate this document, see "Laser Safety Information" on page xxiii.
  - ❑ For FCoE on the Linux OS, see the Installing Linux Driver Software section in the *User's Guide, Converged Network Adapters and Intelligent Ethernet Adapters, QLogic FastLinQ 3400 and 8400 Series*. To locate this document, see "Laser Safety Information" on page xxiii.
  - ❑ For FCoE on the Windows OS, use the QCC GUI, QCS CLI, or QCC PowerKit to enable or disable the FCoE-offload instance per port on Windows in single function mode.

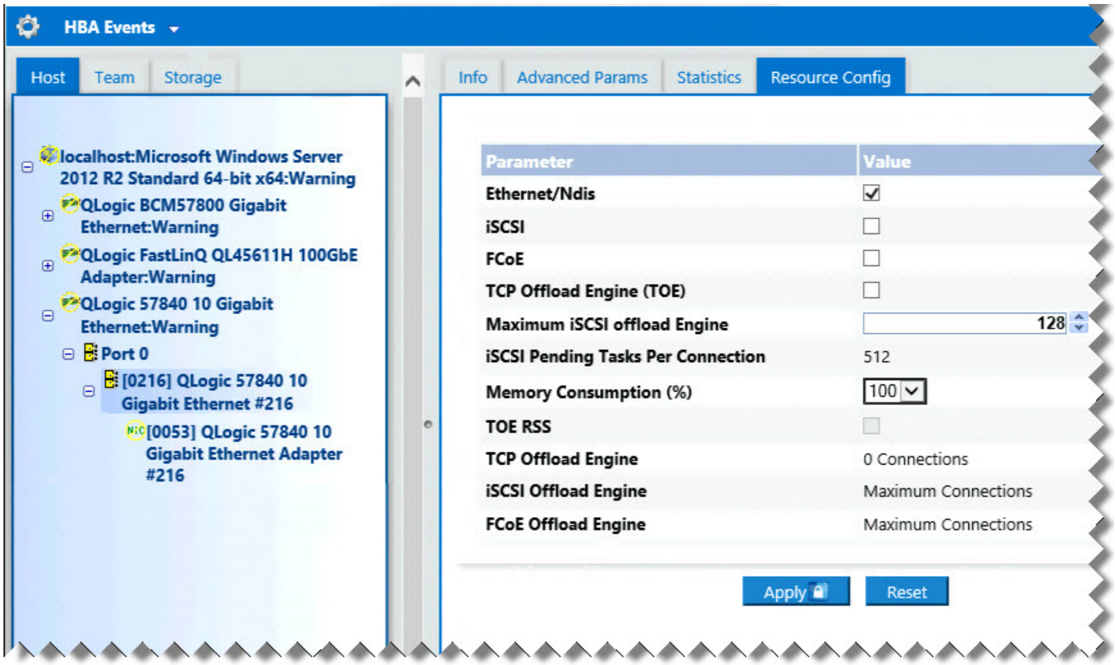
    To configure iSCSI offload in NPAR mode, use the NPAR configuration page in any of the following applications:

    - QCC GUI
    - QCS CLI
    - QCC PowerKit
    - Pre-boot server UEFI HII
    - Pre-boot CCM

To enable and disable the FCoE-offload instance on Windows using QCC GUI:

1. Open QCC GUI.
2. In the tree pane on the left, under the port node, select the port's virtual bus device instance.
3. In the configuration pane on the right, click the **Resource Config** tab.

   The Resource Config page appears (see Figure 14-30).


HBA Events
Host
Team
Storage
localhost:Microsoft Windows Server 2012 R2 Standard 64-bit x64:Warning
QLogic BCM57800 Gigabit Ethernet:Warning
QLogic FastLinQ QL45611H 100GbE Adapter:Warning
QLogic 57840 10 Gigabit Ethernet:Warning
Port 0
[0216] QLogic 57840 10 Gigabit Ethernet #216
[0053] QLogic 57840 10 Gigabit Ethernet Adapter #216
Info
Advanced Params
Statistics
Resource Config
Apply
Reset

Parameter	Value
Ethernet/Ndis	☑
iSCSI	□
FCoE	□
TCP Offload Engine (TOE)	□
Maximum iSCSI offload Engine	128
iSCSI Pending Tasks Per Connection	512
Memory Consumption (%)	100
TOE RSS	□
TCP Offload Engine	0 Connections
iSCSI Offload Engine	Maximum Connections
FCoE Offload Engine	Maximum Connections

Figure 14-30. Resource Config Page

4. Complete the Ethernet/NDIS and/or iSCSI and/or FCoE and/or TOE settings on the Resource Config page for each selected port as follows:

   a. To enable FCoE offload for the port, for the FCoE parameter, select the **Value** check box.

   b. To disable FCoE offload for the port, for the FCoE parameter, clear the **Value** check box.

   c. Click the **Apply** button.

5. (optional) To enable or disable FCoE-Offload or iSCSI-Offload in single function or NPAR mode on Windows or Linux using QCS CLI, see the *User's Guide, QLogic Control Suite CLI* (part number BC0054511-00). To enable or disable FCoE-Offload or iSCSI-Offload in single function or NPAR mode on Windows or Linux using the QCC PowerKit, see the *User's Guide, PowerShell* (part number BC0054518-00). To locate these documents, see "Laser Safety Information" on page xxiii.

N_Port ID Virtualization (NPIV)

NPIV is a Fibre Channel protocol that allows multiple, virtual N_Ports to be instantiated on a single physical N_Port.

- Each NPIV port is provided with a unique identification in the fabric and appears as a distinct initiator port at the operating system level.
- The 57712/578*xx* FCoE drivers support NPIV by default, without requiring any user inputs.
- The quantity of NPIV ports that can be created depends on the individual operating system drivers and the capabilities/limits of the fabric (FCoE/FC switch). Operating system driver limits for the 57712/578*xx* FCoE adapters are:
  - ❑ Microsoft Windows: 256
  - ❑ Linux: 64
  - ❑ ESXi 6.7/7.0: 64 (using only the native qfle3f driver)

15 Data Center Bridging

This chapter provides the following information about the data center bridging feature:

Overview

Data center bridging (DCB) is a collection of IEEE specified standard extensions to Ethernet to provide lossless data delivery, low latency, and standards-based bandwidth sharing of data center physical links. DCB supports storage, management, computing, and communications fabrics onto a single physical fabric that is simpler to deploy, upgrade, and maintain than in standard Ethernet networks. DCB has a standards-based bandwidth sharing at its core, allowing multiple fabrics to coexist on the same physical fabric. The various capabilities of DCB allow for LAN traffic (large quantity of flows and not latency-sensitive), SAN traffic (large packet sizes and requires lossless performance), and IPC (latency-sensitive messages) to bandwidth share the same physical converged connection and achieve the necessary individual traffic performance.

DCB includes the following capabilities:

- Enhanced transmission selection (ETS)
- Priority-based flow control (PFC)
- Data center bridging exchange (DCBX) protocol

DCB Capabilities

DCB capabilities include ETS, PFC, and DCBX, as described in this section.

Enhanced Transmission Selection (ETS)

Enhanced transmission selection (ETS) provides a common management framework for assignment of bandwidth to traffic classes. Each traffic class or priority can be grouped in a priority group (PG), and it can be considered as a virtual link or virtual interface queue. The transmission scheduler in the peer is responsible for maintaining the allocated bandwidth for each PG. For example, a user can configure FCoE traffic to be in PG 0 and iSCSI traffic in PG 1. The user can then allocate each group a specific bandwidth. For example, 60 percent to FCoE and 40 percent to iSCSI. The transmission scheduler in the peer will ensure that in the event of congestion, the FCoE traffic will be able to use at least 60 percent of the link bandwidth and iSCSI to use 40 percent. See additional references at:

http://www.ieee802.org/1/pages/802.1az.html

Priority Flow Control (PFC)

Priority flow control (PFC) provides a link-level flow control mechanism that can be controlled independently for each traffic type. The goal of this mechanism is to ensure zero loss due to congestion in DCB networks. Traditional IEEE 802.3 Ethernet does not guarantee that a packet transmitted on the network will reach its intended destination. Upper-level protocols are responsible to maintain the reliability by way of acknowledgment and retransmission. In a network with multiple traffic classes, it becomes very difficult to maintain the reliability of traffic in the absence of feedback. This is traditionally tackled with the help of link-level flow control.

When PFC is used in a network with multiple traffic types, each traffic type can be encoded with a different priority value and a pause frame can refer to this priority value while instructing the transmitter to stop and restart the traffic. The value range for the priority field is from 0 to 7, allowing eight distinct types of traffic that can be individually stopped and started. See additional references at:

http://www.ieee802.org/1/pages/802.1bb.html

Data Center Bridging Exchange (DCBX)

Data center bridging exchange (DCBX) is a discovery and capability exchange protocol that is used for conveying capabilities and configuration of ETS and PFC between link partners to ensure consistent configuration across the network fabric. In order for two devices to exchange information, one device must be willing to adopt network configuration from the other device. For example, if a C-NIC is configured to willingly adopt ETS and PFC configuration information from a connected switch, and the switch acknowledges the C-NIC's willingness, the switch will send the C-NIC the recommended ETS and PFC parameter settings. The DCBX protocol uses the link level discovery protocol (LLDP) to exchange PFC and ETS configurations between link partners.

Configuring DCB

By default, DCB is enabled on 57712/578*xx* DCB-compatible C-NICs. DCB configuration is rarely required, as the default configuration should satisfy most scenarios. DCB parameters can be configured through QCS CLI. For more information on QCS CLI, see the *User's Guide, QLogic Control Suite CLI*.

NOTE

FCoE operation depends on successful VLAN discovery. All switches that support FCoE support VLAN discovery, but some switches may require specific configuration. Refer to the switch configuration guides for information on how to configure a port for successful VLAN discovery.

DCB Conditions

The following conditions allow DCB technology to function on the network:

- If DCB is enabled on the interface, DCBX is automatically enabled and carried out automatically when a link is established.
- If DCBX fails to synchronize with a compatible peer, the adapter will automatically fall back to default NIC behavior (no priority tagging, no PFC, no ETS).
- By default, the port will advertise itself as willing, and as such, will accept all DCB settings as advertised by the switch.
- If PFC is operational, PFC settings supersede link level flow control settings. If PFC is not operational, link level flow control settings prevail.

- In NIC partitioned enabled configurations, ETS (if operational) overrides the Bandwidth Relative (minimum) Weights assigned to each function. Transmission selection weights are per protocol per ETS settings instead. Maximum bandwidths per function are still honored in the presence of ETS.
- In the absence of an iSCSI or FCoE application TLV advertised through the DCBX peer, the adapter will use the settings taken from the local Admin MIB.

Data Center Bridging in Windows Server 2012 and Later

Starting with Windows Server 2012, Microsoft introduced a new way of managing quality of service (QoS) at the OS level. The two main aspects of Windows QoS include:

- A vendor-independent method for managing DCB settings on NICs, both individually and across an entire domain. The management interface is provided by Windows PowerShell cmdlets.
- The ability to tag specific types of Layer 2 networking traffic, such as SMB traffic, so that hardware bandwidth can be managed using ETS.

All Marvell Converged Network Adapters that support DCB are capable of interoperating with Windows QoS.

To enable the QoS Windows feature, ensure that the Marvell device is DCB-capable:

1. Using CCM or another management utility, enable data center bridging.
2. Using Windows Device Manager or another management utility, select the NDIS driver, display **Advanced** properties, and enable the **Quality of Service** property.

When QoS is enabled, administrative control over DCB-related settings is relinquished to the operating system (that is, QCS CLI or QCC GUI can no longer be used for administrative control of the DCB). You can use PowerShell to configure and manage the QoS feature. Using PowerShell Cmdlets, you can configure various QoS-related parameters, such as traffic classification, priority flow control, and traffic class throughput scheduling. You should ensure that the PowerShell-configured DCB settings are compatible with the attached DCB-enabled switch.

For more information on using PowerShell Cmdlets, see the DCB Windows PowerShell User Scripting Guide in the Microsoft Technet Library.

To revert to standard QCS CLI or QCC GUI control over the Marvell DCB feature set, uninstall the Microsoft QoS feature or disable quality of service in the QCS CLI, QCC GUI, or Device Manager NDIS advance properties page.

NOTE

Marvell recommends that you do not install the DCB feature if SR-IOV will be used. If you install the DCB feature, be aware that selecting **Enable single-root I/O virtualization (SR-IOV)** in Virtual Switch Manager forces the underlying adapter into a DCB state in which OS DCB configuration to be ignored, and DCB configuration from QCS CLI or QCC GUI becomes in effect. However, the user-configured **Networking Priority** value (nonzero) does not take effect, even though it appears that it is from QCS CLI or QCC GUI.

16 SR-IOV

This chapter provides information about single-root I/O virtualization (SR-IOV):

NOTE

See the VMware documentation for enabling SR-IOV on a pNIC at the hypervisor/driver level.

Overview

Virtualization of network controllers allows users to consolidate their networking hardware resources and run multiple virtual machines concurrently on consolidated hardware. Virtualization also provides the user a rich set of features such as I/O sharing, consolidation, isolation and migration, and simplified management with provisions for teaming and failover.

Virtualization can come at the cost of reduced performance due to hypervisor overhead. The PCI-SIG introduced the SR-IOV specification to address these performance issues by creating a virtual function (VF), a lightweight PCIe function that can be directly assigned to a virtual machine (VM), bypassing the hypervisor layer for the main data movement.

Not all Marvell adapters support SR-IOV; refer to your product documentation for details.

Enabling SR-IOV

Before attempting to enable SR-IOV, ensure that:

- The adapter hardware supports SR-IOV.
- SR-IOV is supported and enabled in the system BIOS.
- Configure NPAR mode (if using).

To enable SR-IOV:

1. Enable the feature on the adapter using either QCC GUI, QCS CLI, QCC PowerKit, Dell pre-boot UEFI, or pre-boot CCM.

   If using Windows QCC GUI:

   a. Select the network adapter in the Explorer View pane. Click the **Configuration** tab and select **SR-IOV Global Enable**.

   b. In the **SR-IOV VFs per PF** box, configure the quantity of SR-IOV virtual functions (VFs) that the adapter can support per physical function, from 0 to 64 (57810/57800) or 32 (57840) in increments of 8 (default = 16).

      In NPAR mode, the total VFs that can be enabled over all of the partitions of a single 578*xx* port is limited to either 64 (dual-port 57810) or 32 (quad-port 57840). The 2x10G + 2x1G 57800 adapter supports up to 64 VFs only on the two 10G ports. Be sure to configure NPAR before configuring SR-IOV.

   c. In the **SR-IOV Max Chains per VF** box, configure the maximum quantity of transmit and receive queues (such as receive side scaling (RSS) queues) that can be used for each virtual function. The maximum is 16.

   If using pre-boot UEFI:

   a. During power up, press F2 at the prompt to enter the Dell System Setup.

   b. Select the **Device Settings** menu.

   c. Select the SR-IOV-capable adapter port from the Device Settings menu.

   d. Select the **Device Level Configuration Menu** on the Main Configuration Page.

   e. In the Virtualization Mode list, select **SR-IOV** or **NPar+SR-IOV** (if you want SR-IOV-over-NPAR mode) control.

f. If in SR-IOV mode (without NPAR mode), select the desired number of VFs for this port in the **Number of VFs Per PF** control window.

   The 2x1G+2x10G 57800 allows up to 64 VFs per 10G port (the 57800's two 1G ports do not support SR-IOV). The 2x10G 57810 allows up to 64 VFs per port. The 4x10G 57840 allows up to 32 VFs per port.

g. If in SR-IOV (with NPAR mode), each partition has a separate Number of VFs Per PF control window. Press ESC to return to the Main Configuration Page, and then select the **NIC Partitioning Configuration** menu (which appears only if NPAR mode is selected in the Virtualization Mode control). In the NIC Partitioning Configuration page, select each **Partition "N" Configuration** menu and set the **Number of VFs per PF** control. The total number of VFs assigned per PF on a single physical port cannot exceed the numbers assigned in Step f.

If using pre-boot CCM:

a. During power up, press CTRL+S at the prompt to enter CCM.

b. Select the SR-IOV-capable adapter from the Device List. On the Main Menu, select **Device Hardware Configuration**, and then select **SR-IOV Enabled**.

c. To configure the quantity of VFs that the adapter can support:

   - If **Multi-Function Mode** to is set to **SF** (Single Function), the **Number of VFs per PF** box appears, which you can set from 0 to 64 in increments of 8 (default is 16).
   - If **Multi-Function Mode** is set to **NPAR**, display the Main Menu and select **NIC Partition Configuration**. Then, select the NPAR Function to configure and enter the appropriate value in the **Number of VFs per PF** box.

2. Using either the Windows Device Manager, QCS CLI, or QCC GUI, enable SR-IOV in the advanced properties of the Windows driver.

3. In Virtual Switch Manager, create a virtual NIC using the appropriate procedure for either Windows or ESX.

   In Windows:

   a. Select **Allow Management operating system to share the network adapter** if the host will use this vSwitch to connect to the associated VMs.

   b. Create a vSwitch and select the **Enable Single root I/O Virtualization** option.

   c. In Virtual Switch Manager, select the virtual adapter and select **Hardware Acceleration** in the navigation pane. In the **Single-root I/O virtualization** section, select **Enable SR-IOV**. SR-IOV must be done now and cannot be enabled after the vSwitch is created.

   In ESX:

   a. Install the qfle3 driver.

   b. Ensure that the `lspci` command output on ESXi lists the desired adapter.

   c. From `lspci`, select the 10G NIC sequence number for which SR-IOV is required. For example:

      ```
      ~ # lspci | grep -i Broadcom 0000:03:00.0 Network
      Controllers: Broadcom Corporation NetXtreme II BCM57810
      10 Gigabit Ethernet [vmnic0]
      ```

      Following is a sample output.

      ```
      0000:03:00.1 Network Controllers: Broadcom Corporation
      NetXtreme II BCM57810 10 Gigabit Ethernet [vmnic1]
      ~ #
      ```

   d. In the driver, enable SR-IOV by using the `max_vfs` parameter and passing a list containing the quantity of VFs for each port. In the BIOS, ensure that the quantity of VFs per PF parameter is configured with a minimum of the required quantity of VFs. Each PF port supports a maximum of 64 VFs; the minimum quantity is 1. For example:

      ```
      ~ # esxcli system module parameters set -m bnx2x -p
      "max_vfs=64, 64"
      ```

   e. Restart the system.

4. Install the Marvell drivers for the adapters detected in the VM. Use the latest drivers available from your vendor for the host OS (do not use the inbox drivers). The same driver version should be installed on the host and the VM.

Verifying that SR-IOV is Operational

Follow the appropriate steps for Hyper-V, VMware vSphere, or ESXi CLI.

To verify SR-IOV in Hyper-V Manager:

1. Start the VM.
2. In Hyper-V Manager, select the adapter and select the VM in the **Virtual Machines** list.
3. Click the **Networking** tab at the bottom of the window and view the adapter status.

To verify SR-IOV in VMware vSphere 6.0 U2 Web Client:

1. Confirm that the VFs appear as regular VMDirectPath devices by selecting **Host**, **Manage, Settings**, **Hardware**, and then **PCI Devices**.
2. Right-click **VM**, **Edit settings**, **New Device**, **Select Network**, and **Add**. Click **New Network** and then select **SR-IOV** as the adapter type. Click **OK**.

To verify SR-IOV in ESXi CLI:

1. Issue the `lspci` command:

   ```
   ~ # lspci | grep -i ether
   ```

   Following is a sample output.

   ```
   0000:03:01.0 Network controller: Broadcom Corporation
   NetXtreme II BCM57810 10 Gigabit Ethernet Virtual Function
   [PF_0.3.0_VF_0]
   ```

2. To list the SR-IOV-enabled NIC, issue the `esxcli` command:

   ```
   ~ # esxcli network sriovnic list
   ```

   Following is a sample output.

```
Name   PCI Device    Driver Link Speed Duplex MAC Address       MTU  Description
------ ------------- ------ ---- ----- ------ ----------------  ---- -----------
vmnic0 0000:003:00.0 bnx2x  Up   10000 Full   3c:d9:2b:f6:71:50 1500 Broadcom Corpo
vmnic1 0000:003:00.1 bnx2x  Down 0     Full   3c:d9:2b:f6:71:54 1500 Broadcom Corpo
```

SR-IOV and Storage Functionality

You can enable storage functionality (FCoE or iSCSI) on an SR-IOV-enabled adapter. However, if storage is used on an NPAR-enabled physical function (PF), the quantity of virtual functions for that PF is set to zero; therefore, SR-IOV is disabled on that specific PF.

This limitation applies only when the adapter is configured in NPAR mode. It is not relevant when the adapter is configured in single-function (SF) mode.

In ESX, after enabling SR-IOV in the OS for SF mode, the storage adapter will not be discovered.

SR-IOV and Jumbo Packets

If SR-IOV is enabled on a virtual function (VF) on the adapter, ensure that the same jumbo packet settings is configured on both the VF and the Microsoft synthetic adapter. You can configure these values using Windows Device Manager, Advanced properties.

If there is a mismatch in the values, the SR-IOV function is shown as being in the degraded state in Hyper-V, Networking Status.

17 Specifications

Specifications, characteristics, and requirements include:

10/100/1000BASE-T and 10GBASE-T Cable Specifications

Table 17-1. 10/100/1000BASE-T Cable Specifications

Port Type	Connector	Media	Maximum Distance
10BASE-T	RJ45	CAT-3, 4, or 5 unshielded twisted pairs (UTP)	328ft (100m)
100/1000BASE-T [a]	RJ45	CAT-5 [b] UTP	328ft (100m)

[a] 1000BASE-T signaling requires four twisted pairs of CAT-5 balanced cabling, as specified in ISO/IEC 11801:2002 and ANSI/EIA/TIA-568-B.

[b] CAT-5 is the minimum requirement. CAT-5e and CAT-6 are fully supported.

Table 17-2. 10GBASE-T Cable Specifications

Port Type	Connector	Media	Maximum Distance
10GBASE-T	RJ45	CAT-6 [a] UTP CAT-6A[a] UTP	131ft (40m) 328ft (100m)

[a] 10GBASE-T signaling requires four twisted pairs of CAT-6 or CAT-6A (augmented CAT-6) balanced cabling, as specified in ISO/IEC 11801:2002 and ANSI/TIA/EIA-568-B

Supported SFP+ Modules Per NIC

Table 17-3. 57710 Supported Modules

Module Type	Module Vendor	Module Part Number
Optic Modules (SR)	Finisar Corp. Avago Avago Intel Corp.	FTLX8571D3BCL AFBR-707SDZ-D1 AFBR-703SDZ-D1 FTLX8571D3BCV-IT
Direct Attach Cables	Cisco-Molex Inc. Cisco-Molex Inc. Cisco-Molex Inc. Cisco-Molex Inc.	74752-9093 74752-9094 74752-9096 74752-9098

Table 17-4. 57810 Supported Modules

Module Type	Dell Part Number	Module Vendor	Module Part Number
Optic Modules (SR)	W365M	Avago	AFBR-703SDZ-D1
	N743D	Finisar Corp.	FTLX8571D3BCL
	R8H2F	Intel Corp.	AFBR-703SDZ-IN2
	R8H2F	Intel Corp.	FTLX8571D3BCV-IT
Direct Attach Cables	K585N	Cisco-Molex Inc.	74752-9093
	J564N	Cisco-Molex Inc.	74752-9094
	H603N	Cisco-Molex Inc.	74752-9096
	G840N	Cisco-Molex Inc.	74752-9098
	1539W	Brocade	58-1000026-01
	V239T	Brocade	58-1000027-01
	48V40	Brocade	58-1000023-01
	C4D08 - Force10 1m DAC	Amphenol	599700002
	C4D08 - Force10 1m DAC	Amphenol	616740001
	53HVN - Force10 3m DAC	Amphenol	599700006
	53HVN - Force10 3m DA	Amphenol	616740003
	5CN56 - Force10 5m DAC	Amphenol	599700004
	5CN56 - Force10 5m DAC	Amphenol	616740005

Table 17-5. 57840 Supported Modules

Module Type	Dell Part Number	Module Vendor	Module Part Number
Optic Modules (SR)	R8H2F	Intel Corp. Intel Corp.	AFBR-703SDZ-IN2 FTLX8571D3BCV-IT
Direct Attach Cables	K585N	Cisco-Molex Inc.	74752-9093
	J564N	Cisco-Molex Inc.	74752-9094
	H603N	Cisco-Molex Inc.	74752-9096
	G840N	Cisco-Molex Inc.	74752-9098
	1539W	Brocade	58-1000026-01
	V239T	Brocade	58-1000027-01
	48V40	Brocade	58-1000023-01
	C4D08 - Force10 1m DAC	Amphenol	599700002
	C4D08 - Force10 1m DAC	Amphenol	616740001
	53HVN - Force10 3m DAC	Amphenol	599700006
	53HVN - Force10 3m DAC	Amphenol	616740003
	5CN56 - Force10 5m DAC	Amphenol	599700004
	5CN56 - Force10 5m DAC	Amphenol	616740005

Interface Specifications

Table 17-6. 10, 100, and 1000BASE-T Performance Specifications

Feature	Specification
PCI Express Interface	x4 link width
10/100/1000BASE-T	10/100/1000Mbps

Table 17-7. 10GBASE-T Performance Specifications

Feature	Specification
PCI Express Interface	x8 link width
10GBASE-T	10Gbps

NIC Physical Characteristics

Table 17-8. NIC Physical Characteristics

NIC Type	NIC Length	NIC Width
57810S PCI Express x8 low profile	6.6in (16.8cm)	2.54in (6.5cm)

NIC Power Requirements

Table 17-9. 957810A1006G NIC Power Requirements

Link	NIC 12V Current Draw (A)	NIC 3.3V Current Draw (A)	NIC Power (W) [a]
10G SFP Module	1.00	0.004	12.0

[a] Power, measured in watts (W), is a direct calculation of total current draw (A) multiplied by voltage (V). The maximum power consumption for the adapter will not exceed 25W.

Table 17-10. 957810A1008G NIC Power Requirements

Link	NIC 12V Current Draw (A)	NIC 3.3V Current Draw (A)	NIC Power (W) [a]
Idle (no link)	0.9	0.004	11.0
100BASE-T link	1.0	0.004	12.0
1000BASE-T link	1.3	0.004	15.5
10GBASE-T link	1.8	0.004	20.0

[a] Power, measured in watts (W), is a direct calculation of total current draw (A) multiplied by voltage (V). The maximum power consumption for the adapter will not exceed 25W.

Table 17-11. 957840A4006G Mezzanine Card Power Requirements

Link	Total Power (12V and 3.3VAUX) (W) [a]
10G SFP+	12.0
Standby WoL Enabled	5.0
Standby WoL Disabled	0.5

[a] Power, measured in watts (W), is a direct calculation of total current draw (A) multiplied by voltage (V). The maximum power consumption for the adapter will not exceed 25W.

Table 17-12. 957840A4007G Mezzanine Card Power Requirements

Link	Total Power (3.3V) (W) [a]
10G KR interface	10.0
WoL enabled	3.5

[a] Power, measured in watts (W), is a direct calculation of total current draw (A) multiplied by voltage (V). The maximum power consumption for the adapter will not exceed 25W.

Wake on LAN Power Requirements

Nominal power for WoL:

- 957810A1006G: 9.0W
- 957810A1008G: 16.0W

Environmental Specifications

Table 17-13. 5709 and 5716 Environmental Specifications

Parameter	Condition
Operating Temperature	32°F to 131°F (0°C to 55°C)
Air Flow Requirement (LFM)	0
Storage Temperature	–40°F to 149°F (–40°C to 65°C)
Storage Humidity	5% to 95% condensing
Vibration and Shock	IEC 68, FCC Part 68.302, NSTA, 1A
Electrostatic/Electromagnetic Susceptibility	EN 61000-4-2, EN 55024

Table 17-14. 957810A1006G Environmental Specifications

Parameter	Condition
Operating Temperature	32°F to 131°F (0°C to 55°C)
Air Flow Requirement (LFM)	100
Storage Temperature	–40°F to 149°F (–40°C to 65°C)
Storage Humidity	5% to 95% condensing
Vibration and Shock	IEC 68, FCC Part 68.302, NSTA, 1A
Electrostatic/Electromagnetic Susceptibility	IEC 801-2, 3, 4, 5

Table 17-15. 957810A1008G Environmental Specifications

Parameter	Condition
Operating Temperature	32°F to 131°F (0°C to 55°C)
Air Flow Requirement (LFM)	50
Storage Temperature	–40°F to 149°F (–40°C to 65°C)
Storage Humidity	5% to 95% condensing
Vibration and Shock	IEC 68, FCC Part 68.302, NSTA, 1A
Electrostatic/Electromagnetic Susceptibility	IEC 801-2, 3, 4, 5

Table 17-16. 957840A4007G Environmental Specifications

Parameter	Condition
Operating Temperature	32°F to 131°F (0°C to 65°C)
Air Flow Requirement (LFM)	200
Storage Temperature	–40°F to 149°F (–40°C to 65°C)
Storage Humidity	5% to 95% condensing
Vibration and Shock	IEC 68, FCC Part 68.302, NSTA, 1A
Electrostatic/Electromagnetic Susceptibility	IEC 801-2, 3, 4, 5

18 Regulatory Information

Regulatory information covered in this chapter includes the following:

- Product Safety
- AS/NZS (C-Tick)
- "FCC Notice" on page 261
- "VCCI Notice" on page 263
- "CE Notice" on page 268
- "Canadian Regulatory Information (Canada Only)" on page 269
- "Korea Communications Commission (KCC) Notice (Republic of Korea Only)" on page 271
- "BSMI" on page 274
- "Certifications for 95709SA0908G, 957710A1023G (E02D001), and 957711A1123G (E03D001)" on page 274

Product Safety

⚠ **WARNING**

Before installing adapter hardware, power off the computer and all attached devices such as monitors, printers, and external components.

Use 57*xx* and 57*xxx* adapters only with the listed ITE or equivalent. UL and TUV standard number and CB certifications:

- UL 60950-1 (2nd Edition) 2007
- CSA C22.2 No.60950-1-07 (2nd Edition) 2007
- TUV EN60950-1:2006+A11+A1+A12 2nd Edition
- TUV IEC 60950-1:2005 2nd Edition Am 1:2009 CB
- TUV IEC 62368-1 2nd Edition, 3rd Edition

AS/NZS (C-Tick)

AS/NZS; CISPR 22:2009+A1:2010 Class A

FCC Notice

FCC, Class B

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95708A0804F
- 95709A0907G
- 95709A0906G
- 957810A1008G

Marvell Semiconductor, Inc.
15485 San Canyon Ave
Irvine, CA 92618 USA

The equipment complies with Part 15 of the FCC Rules. Operation is subject to the following two conditions: 1) The device may not cause harmful interference, and 2) This equipment must accept any interference received, including interference that may cause unwanted operation.

The equipment has been tested and found to comply with the limits for a Class B digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference in a residential installation. The equipment generates, uses and can radiate radio-frequency energy and, if not installed and used in accordance with the instructions, may cause harmful interference to radio communications. However, there is no guarantee that interference will not occur in a specific installation. If the equipment does cause harmful interference to radio or television reception, which can be determined by turning the equipment off and on, the user is encouraged to try to correct the interference by one or more of the following measures:

- Reorient or relocate the receiving antenna.
- Increase the separation between the equipment and the receiver.
- Connect the equipment into an outlet on a circuit different from that to which the receiver is connected.
- Consult the dealer or an experienced radio/TV technician for assistance.

Do not make mechanical or electrical modifications to the equipment.

NOTE

If the device is changed or modified without permission of Marvell, the user may void his or her authority to operate the equipment.

FCC, Class A

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller:

- 95709A0916G

Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet controller:

- 957800
- 957710A1022G
- 957710A1021G
- 957711A1113G
- 957711A1102G
- 957810A1006G (BC0410401)
- 957840A4006G
- 957840A4007G

Marvell Semiconductor, Inc.
15485 San Canyon Ave
Irvine, CA 92618 USA

This device complies with Part 15 of the FCC Rules. Operations is subject to the following two conditions: 1) This device may not cause harmful interference, and 2) This device must accept any interference received, including interference that may cause unwanted operation.

This product has been tested and found to comply with the limits for a Class A digital device pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This product generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the manufacturer's instruction manual, may cause harmful interference with radio communications. Operation of this product in a residential area is likely to cause harmful interference, in which case you will be required to correct the interference at your own expense.

These limits are designed to provide reasonable protection against harmful interference in a non-residential installation. However, there is no guarantee that interference will not occur in a specific installation. If this equipment does cause harmful interference with radio or television reception, which can be determined by turning the equipment off and on, you are encouraged to try to correct the interference by one or more of the following measures:

- Reorient the receiving antenna.
- Relocate the system with respect to the receiver.
- Move the system away from the receiver.
- Plug the system into a different outlet so that the system and receiver are on different branch circuits.

Do not make mechanical or electrical modifications to the equipment.

NOTE

If the device is changed or modified without permission of Marvell, the user may void his or her authority to operate the equipment.

VCCI Notice

The following tables provide the VCCI notice physical specifications for the Marvell 57*xx* and 57*xxx* adapters for Dell.

Table 18-1. Marvell 57800S 1GB and 10GBASE-T Rack Network Daughter Card Physical Characteristics

Item	Description
Ports	Dual 1Gbps Ethernet and dual 10Gbps Ethernet
Form Factor	Network daughter card 3.66in × 2.93in (92.9mm × 74.4mm)
Supported Servers	13th Generation: R630, R730, R730xd, and T630 12th Generation: R620, R720, R720xd, R820, and R920
Connector	10G BASE-T and RJ45
Cable	CAT6a and 7 up to 100 meters CAT6 up to 40 meters
Certifications	RoHS, FCC A, UL, CE, VCCI, BSMI, C-Tick, KCC, TUV, and ICES-003

Table 18-2. Marvell 57800S Quad RJ-45, SFP+, or Direct Attach Rack Network Daughter Card Physical Characteristics

Item	Description
Ports	Dual 1Gbps Ethernet and dual 10Gbps Ethernet
Form Factor	Network daughter card 3.66in × 2.93in (92.9mm × 74.4mm)
Supported Servers	13th Generation: R630, R730, R730xd, and T630 12th Generation: R620, R720, R720xd, R820, and R920

Table 18-2. Marvell 57800S Quad RJ-45, SFP+, or Direct Attach Rack Network Daughter Card Physical Characteristics (Continued)

Item	Description
Connectors	Two ports SFP+ (10GbE) Two ports RJ45 (1GbE)
Certifications	RoHS, FCC A, UL, CE, VCCI, BSMI, C-Tick, KCC, TUV, and ICES-003

Table 18-3. Marvell 57810S Dual 10GBASE-T PCI-e Card Physical Characteristics

Item	Description
Ports	Dual 10Gbps BASE-T Ethernet ports
Form Factor	PCI Express short, low-profile card 6.60in × 2.71in (167.64mm × 68.91mm)
Supported Servers	13th Generation: R630, R730, R730xd, and T630 12th Generation: R320, R420, R520, R620, R720, R720xd, R820, T420, and T620
Connector	RJ45
Cable	CAT6a and 7 up to 100 meters CAT6 up to 40 meters
Certifications	RoHS, FCC A, UL, CE, VCCI, BSMI, C-Tick, KCC, TUV, and ICES-003

Table 18-4. Marvell 57810S Dual SFP+ or Direct Attach PCIe Physical Characteristics

Item	Description
Ports	Dual 10Gbps Ethernet
Form Factor	PCI Express short, low-profile card 6.60in × 2.71in (67.64mm × 68.91mm)

Table 18-4. Marvell 57810S Dual SFP+ or Direct Attach PCIe Physical Characteristics (Continued)

Item	Description
Supported Servers	13th Generation: R630, R730, R730xd, and T630 12th Generation: R220, R320, R420, R520, R620, R720, R720xd, R820, R920, T420, and T620
Certifications	RoHS, FCC A, UL, CE, VCCI, BSMI, C-Tick, KCC, TUV, and ICES-003

Table 18-5. Marvell 57810S-K Dual KR Blade Mezzanine Adapter Physical Characteristics

Item	Description
Ports	Dual 10Gbps Ethernet
Form Factor	Mezzanine adapter 3.13in × 2.85in (79.5mm × 72.4mm)
Supported Servers	13th Generation: M630 12th Generation: M420, M520, M620, and M820
Certifications	RoHS, FCC A, UL, CE, VCCI, C-Tick, KCC, TUV, and ICES-003

Table 18-6. Marvell 57810S-K Dual KR Blade Network Daughter Card Physical Characteristics

Item	Description
Ports	Dual 10Gbps Ethernet
Form Factor	Network daughter card 2.45in × 3.0 in. (62.2 mm × 76.2 mm)
Supported Servers	13th Generation: M630 12th Generation: M620 and M820
Certifications	RoHS, FCC A, UL, CE, VCCI, C-Tick, KCC, TUV, and ICES-003

Table 18-7. Marvell 57840S Quad 10GbE SFP+ or Direct Attach Rack Network Daughter Card Physical Characteristics

Item	Description
Ports	Dual 10Gbps Ethernet
Form Factor	PCI Express short, low-profile card 6.60in × 2.71in (67.64mm × 68.91mm)
Supported Servers	13th Generation: R630, R730, R730xd, and T630 12th Generation: R320, R420, R520, R620, R720, R720xd, R820, T420, and T620
Certifications	RoHS, FCC A, UL, CE, VCCI, BSMI, C-Tick, KCC, TUV, and ICES-003

Table 18-8. Marvell 57840S-K Quad KR Blade Network Daughter Card Physical Characteristics

Item	Description
Ports	Quad 10Gbps Ethernet
Form Factor	Network daughter card 2.45in × 3.00in (62.2mm × 76.2mm)
Supported Servers	13th Generation: M630 12th Generation: M420, M520, M620, and M820
Certifications	RoHS, FCC A, UL, CE, VCCI, BSMI, C-Tick, KCC, TUV, and ICES-003

VCCI, Class B

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95708A0804F
- 95709A0907G
- 95709A0906G
- 957810A1008G

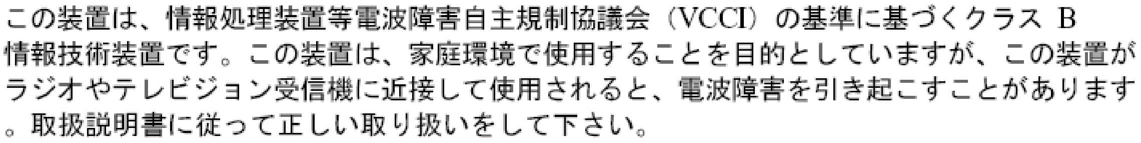
Marvell Semiconductor, Inc.
15485 San Canyon Ave
Irvine, CA 92618 USA

The equipment is a Class B product based on the standard of the Voluntary Control Council for Interference from Information Technology Equipment (VCCI). If used near a radio or television receiver in a domestic environment, it may cause radio interference. Install and use the equipment according to the instruction manual.

CAUTION

The potential exists for this equipment to become impaired in the presence of conducted radio frequency energy between the frequency range of 59–66 MHz. Normal operation will return upon removal of the RF energy source.

VCCI Class B Statement (Japan)



この装置は、情報処理装置等電波障害自主規制協議会（VCCI）の基準に基づくクラス B
情報技術装置です。この装置は、家庭環境で使用することを目的としていますが、この装置が
ラジオやテレビジョン受信機に近接して使用されると、電波障害を引き起こすことがあります
。取扱説明書に従って正しい取り扱いをして下さい。

VCCI, Class A

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95709A0916G

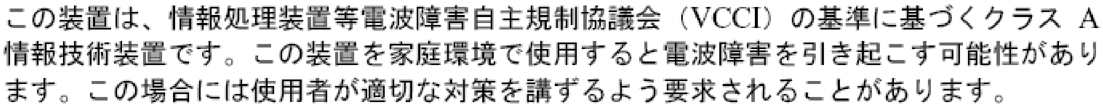
Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet controller

- 957710A1022G
- 957710A1021G
- 957711A1113G
- 957711A1102G
- 957840A4006G
- 957840A4007G

Marvell Semiconductor, Inc.
15485 San Canyon Ave
Irvine, CA 92618 USA

This equipment is a Class A product based on the standard of the Voluntary Control Council for interference by Information Technology Equipment (VCCI). If used in a domestic environment, radio disturbance may arise. Install and use the equipment according to the instruction manual.

VCCI Class A Statement (Japan)



この装置は、情報処理装置等電波障害自主規制協議会（VCCI）の基準に基づくクラス A 情報技術装置です。この装置を家庭環境で使用すると電波障害を引き起こす可能性があります。この場合には使用者が適切な対策を講ずるよう要求されることがあります。

CE Notice

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95708A0804F
- 95709A0907G
- 95709A0906G
- 95709A0916G
- 957810A1008G

Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet controller

- 957710A1022G
- 957710A1021G
- 957711A1113G
- 957711A1102G
- 957840A4006G
- 957840A4007G

This product has been determined to be in compliance with 2006/95/EC (Low Voltage Directive), 2004/108/EC (EMC Directive), and amendments of the European Union.

A “Declaration of Conformity” in accordance with the preceding directives and standards has been made and is on file at QLogic Corporation, 26650 Aliso Viejo Parkway, Aliso Viejo, California 92656, USA.

European Union, Class B
This QLogic device is classified for use in a typical Class B domestic environment.

European Union, Class A
WARNING: This is a Class A product. In a domestic environment this product may cause radio frequency interference in which case the user may be required to take adequate measures.

Canadian Regulatory Information (Canada Only)

Industry Canada, Class B

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95708A0804F
- 95709A0907G
- 95709A0906G

Marvell Semiconductor, Inc.
15485 San Canyon Ave
Irvine, CA 92618 USA

This Class B digital apparatus complies with Canadian ICES-003.

Notice: The Industry Canada regulations provide that changes or modifications not expressly approved by Marvell could void your authority to operate this equipment.

Industry Canada, Class A

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95709A0916G

Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet controller

- 957710A1022G
- 957710A1021G
- 957711A1113G
- 957711A1102G
- 957810A1008G
- 957840A4006G
- 957840A4007G

Marvell Semiconductor, Inc.
15485 San Canyon Ave
Irvine, CA 92618 USA

This Class A digital apparatus complies with Canadian ICES-003.

Notice: The Industry Canada regulations provide that changes or modifications not expressly approved by Marvell could void your authority to operate this equipment.

Industry Canada, classe B

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95708A0804F
- 95709A0907G
- 95709A0906G

Marvell Semiconductor, Inc.
15485 San Canyon Ave
Irvine, CA 92618 USA

Cet appareil numérique de la classe B est conforme à la norme canadienne ICES-003.

Avis : Dans le cadre des réglementations d'Industry Canada, vos droits d'utilisation de cet équipement peuvent être annulés si des changements ou modifications non expressément approuvés par Marvell y sont apportés.

Industry Canada, classe A

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95709A0916G

Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet controller

- 957710A1022G
- 957710A1021G
- 957711A1113G
- 957711A1102G
- 957810A1008G
- 957840A4006G
- 957840A4007G

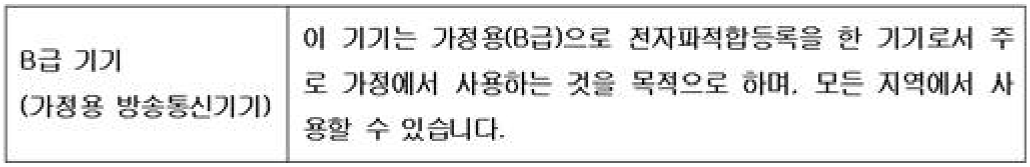
Marvell Semiconductor, Inc.
15485 San Canyon Ave
Irvine, CA 92618 USA

Cet appareil numérique de classe A est conforme à la norme canadienne ICES-003.

Avis : Dans le cadre des réglementations d'Industry Canada, vos droits d'utilisation de cet équipement peuvent être annulés si des changements ou modifications non expressément approuvés par Marvell y sont apportés.

Korea Communications Commission (KCC) Notice (Republic of Korea Only)

B Class Device

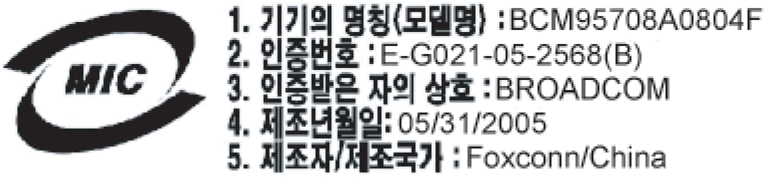


B급 기기 (가정용 방송통신기기)	이 기기는 가정용(B급)으로 전자파적합등록을 한 기기로서 주로 가정에서 사용하는 것을 목적으로 하며, 모든 지역에서 사용할 수 있습니다.

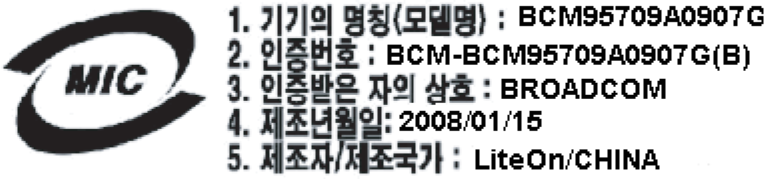
Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95708A0804F
- 95709A0907G
- 95709A0906G

Marvell Semiconductor, Inc.
15485 San Canyon Ave
Irvine, CA 92618 USA



1. 기기의 명칭(모델명) :BCM95708A0804F
2. 인증번호 :E-G021-05-2568(B)
3. 인증받은 자의 상호 :BROADCOM
4. 제조년월일: 05/31/2005
5. 제조자/제조국가 :Foxconn/China



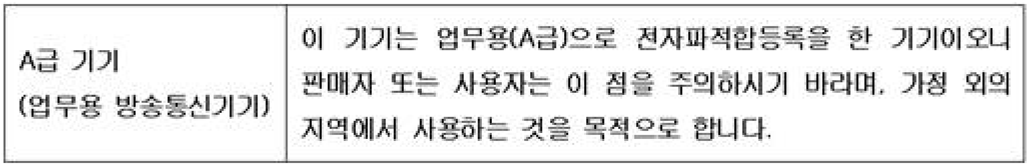
1. 기기의 명칭(모델명) : BCM95709A0907G
2. 인증번호 : BCM-BCM95709A0907G(B)
3. 인증받은 자의 상호 : BROADCOM
4. 제조년월일: 2008/01/15
5. 제조자/제조국가 : LiteOn/CHINA

방송통신위원회

BCM-BCM95709A0906G (B)

Note that this device has been approved for non-business purposes and may be used in any environment, including residential areas.

A Class Device



A급 기기 (업무용 방송통신기기)	이 기기는 업무용(A급)으로 전자파적합등록을 한 기기이오니 판매자 또는 사용자는 이 점을 주의하시기 바라며, 가정 외의 지역에서 사용하는 것을 목적으로 합니다.

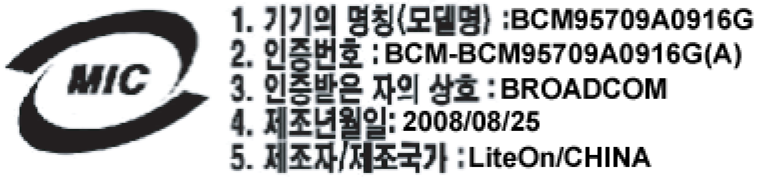
Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95709A0916G

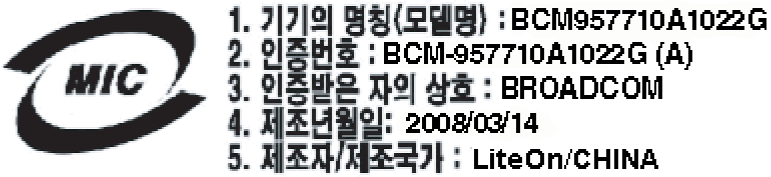
Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet controller

- 957710A1022G
- 957710A1021G
- 957711A1113G
- 957711A1102G
- 957810A1008G
- 957840A4006G
- 957840A4007G

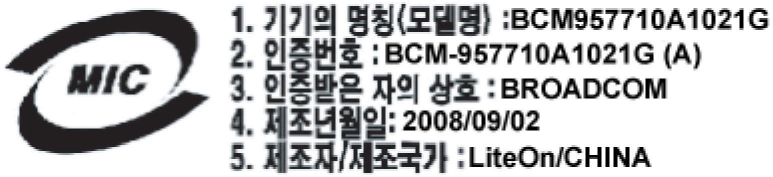
Marvell Semiconductor, Inc.
15485 San Canyon Ave
Irvine, CA 92618 USA



1. 기기의 명칭(모델명) :BCM95709A0916G
2. 인증번호 : BCM-BCM95709A0916G(A)
3. 인증받은 자의 상호 : BROADCOM
4. 제조년월일: 2008/08/25
5. 제조자/제조국가 :LiteOn/CHINA



1. 기기의 명칭(모델명) : BCM957710A1022G
2. 인증번호 : BCM-957710A1022G (A)
3. 인증받은 자의 상호 : BROADCOM
4. 제조년월일: 2008/03/14
5. 제조자/제조국가 : LiteOn/CHINA

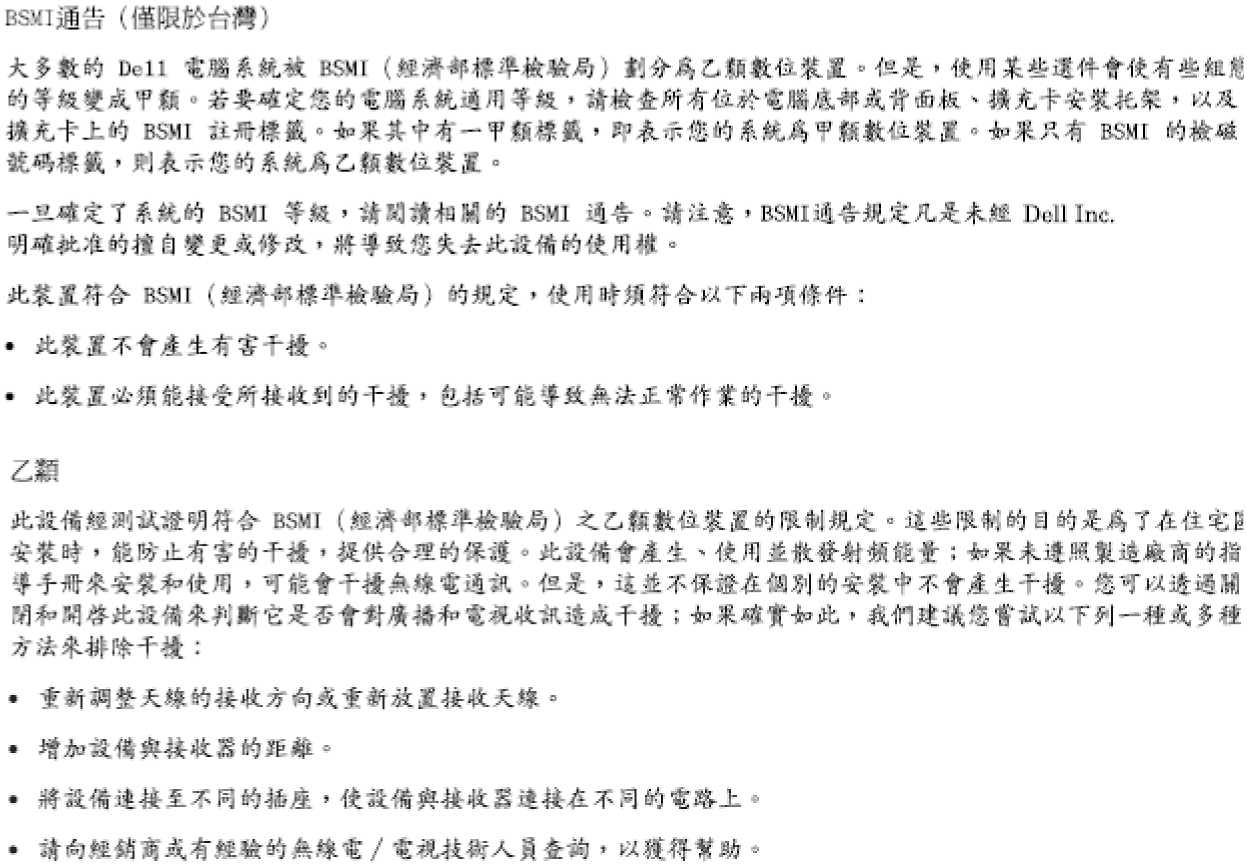


1. 기기의 명칭(모델명) :BCM957710A1021G
2. 인증번호 : BCM-957710A1021G (A)
3. 인증받은 자의 상호 : BROADCOM
4. 제조년월일: 2008/09/02
5. 제조자/제조국가 :LiteOn/CHINA

방송통신위원회
BCM957711A1113G (A)

방송통신위원회
BCM-957711A1102G (A)

BSMI



BSMI通告（僅限於台灣）

大多數的 Dell 電腦系統被 BSMI（經濟部標準檢驗局）劃分爲乙類數位裝置。但是，使用某些選件會使有些組態的等級變成甲類。若要確定您的電腦系統適用等級，請檢查所有位於電腦底部或背面板、擴充卡安裝托架，以及擴充卡上的 BSMI 註冊標籤。如果其中有一甲類標籤，即表示您的系統爲甲類數位裝置。如果只有 BSMI 的檢驗號碼標籤，則表示您的系統爲乙類數位裝置。

一旦確定了系統的 BSMI 等級，請閱讀相關的 BSMI 通告。請注意，BSMI通告規定凡是未經 Dell Inc.
明確批准的擅自變更或修改，將導致您失去此設備的使用權。

此裝置符合 BSMI（經濟部標準檢驗局）的規定，使用時須符合以下兩項條件：

- 此裝置不會產生有害干擾。
- 此裝置必須能接受所接收到的干擾，包括可能導致無法正常作業的干擾。

乙類

此設備經測試證明符合 BSMI（經濟部標準檢驗局）之乙類數位裝置的限制規定。這些限制的目的是爲了在住宅區安裝時，能防止有害的干擾，提供合理的保護。此設備會產生、使用並散發射頻能量；如果未遵照製造廠商的指導手冊來安裝和使用，可能會干擾無線電通訊。但是，這並不保證在個別的安裝中不會產生干擾。您可以透過關閉和開啓此設備來判斷它是否會對廣播和電視收訊造成干擾；如果確實如此，我們建議您嘗試以下列一種或多種方法來排除干擾：

- 重新調整天線的接收方向或重新放置接收天線。
- 增加設備與接收器的距離。
- 將設備連接至不同的插座，使設備與接收器連接在不同的電路上。
- 請向經銷商或有經驗的無線電／電視技術人員查詢，以獲得幫助。

Certifications for 95709SA0908G, 957710A1023G (E02D001), and 957711A1123G (E03D001)

This section is included on behalf of Dell, and Marvell is not responsible for the validity or accuracy of the information.

The 95709SA0908G Marvell 57*xx* and 57*xxx* gigabit Ethernet controller and the 957710A1023G, E02D001, and 957711A1123G (E03D001) Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet controllers have received the following regulatory compliance certifications:

- FCC, Class A (USA)
- VCCI, Class A (Japan)
- Canadian Regulatory Information, Class A (Canada)
- Korea Communications Commission (KCC) Notice (Republic of Korea)

FCC Notice

FCC, Class A

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95709SA0908G

Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet controller

- 957710A1023G
- 957711A1123G (E03D001)
- E02D001

Dell Inc.
Worldwide Regulatory Compliance, Engineering and Environmental Affairs
One Dell Way PS4-30
Round Rock, Texas 78682, USA
512-338-4400

This device complies with Part 15 of the FCC Rules. Operations is subject to the following two conditions: 1) This device may not cause harmful interference, and 2) This device must accept any interference received, including interference that may cause unwanted operation.

This product has been tested and found to comply with the limits for a Class A digital device pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This product generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the manufacturer’s instruction manual, may cause harmful interference with radio communications. Operation of this product in a residential area is likely to cause harmful interference, in which case you will be required to correct the interference at your own expense.

These limits are designed to provide reasonable protection against harmful interference in a non-residential installation. However, there is no guarantee that interference will not occur in a specific installation. If this equipment does cause harmful interference with radio or television reception, which can be determined by turning the equipment off and on, you are encouraged to try to correct the interference by one or more of the following measures:

- Reorient the receiving antenna.
- Relocate the system with respect to the receiver.
- Move the system away from the receiver.
- Plug the system into a different outlet so that the system and receiver are on different branch circuits.

Do not make mechanical or electrical modifications to the equipment.

NOTE

If the device is changed or modified without permission of Dell Inc, the user may void his or her authority to operate the equipment.

VCCI Notice

Class A

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95709SA0908G

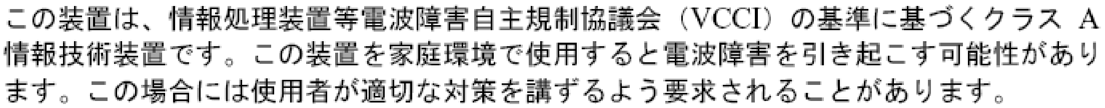
Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet controller

- 957710A1023G
- 957711A1123G (E03D001)
- E02D001

Dell Inc.
Worldwide Regulatory Compliance, Engineering and Environmental Affairs
One Dell Way PS4-30
Round Rock, Texas 78682, USA
512-338-4400

This equipment is a Class A product based on the standard of the Voluntary Control Council for interference by Information Technology Equipment (VCCI). If used in a domestic environment, radio disturbance may arise. Install and use the equipment according to the instruction manual.

VCCI Class A Statement (Japan)



この装置は、情報処理装置等電波障害自主規制協議会（VCCI）の基準に基づくクラス A 情報技術装置です。この装置を家庭環境で使用すると電波障害を引き起こす可能性があります。この場合には使用者が適切な対策を講ずるよう要求されることがあります。

CE Notice

Class A

Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95709SA0908G

Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet controller

- 957710A1023G

- 957711A1123G (E03D001)
- E02D001

Dell Inc.
Worldwide Regulatory Compliance, Engineering and Environmental Affairs
One Dell Way PS4-30
Round Rock, Texas 78682, USA
512-338-4400

This product has been determined to be in compliance with 2006/95/EC (Low Voltage Directive), 2004/108/EC (EMC Directive), and amendments of the European Union.

A “Declaration of Conformity” in accordance with the preceding directives and standards has been made and is on file at Dell Inc., Worldwide Regulatory Compliance, Engineering and Environmental Affairs, One Dell Way PS4-30, Round Rock, Texas 78682, USA.

European Union, Class A
WARNING: This is a Class A product. In a domestic environment this product may cause radio frequency interference in which case the user may be required to take adequate measures.

Canadian Regulatory Information (Canada Only)

Industry Canada, Class A

Marvell 57*xx* and 57*xxx* gigabit Ethernet Controller

- 95709SA0908G

Marvell 57*xx* and 57*xxx* 10Gbt Ethernet Controller

- 957710A1023G
- 957711A1123G (E03D001)
- E02D001

Dell Inc.
Worldwide Regulatory Compliance, Engineering and Environmental Affairs
One Dell Way PS4-30
Round Rock, Texas 78682, USA
512-338-4400

This Class A digital apparatus complies with Canadian ICES-003.

Notice: The Industry Canada regulations provide that changes or modifications not expressly approved by Dell Inc. could void your authority to operate this equipment.

Industry Canada, classe A

Marvell 57*xx* and 57*xxx* gigabit Ethernet Controller

- 95709SA0908G

Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet Controller

- 957710A1023G
- 957711A1123G (E03D001)
- E02D001

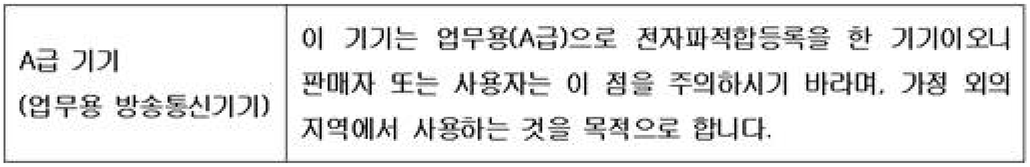
Dell Inc.
Worldwide Regulatory Compliance, Engineering and Environmental Affairs
One Dell Way PS4-30
Round Rock, Texas 78682, USA
512-338-4400

Cet appareil numérique de classe A est conforme à la norme canadienne ICES-003.

Avis : Dans le cadre des réglementations d'Industry Canada, vos droits d'utilisation de cet équipement peuvent être annulés si des changements ou modifications non expressément approuvés par Dell Inc. y sont apportés.

Korea Communications Commission (KCC) Notice (Republic of Korea Only)

A Class Device



A급 기기 (업무용 방송통신기기)	이 기기는 업무용(A급)으로 전자파적합등록을 한 기기이오니 판매자 또는 사용자는 이 점을 주의하시기 바라며. 가정 외의 지역에서 사용하는 것을 목적으로 합니다.

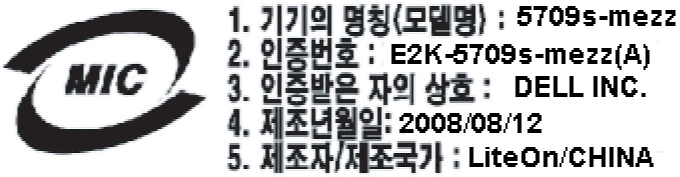
Marvell 57*xx* and 57*xxx* gigabit Ethernet controller

- 95709SA0908G (5709s-mezz)

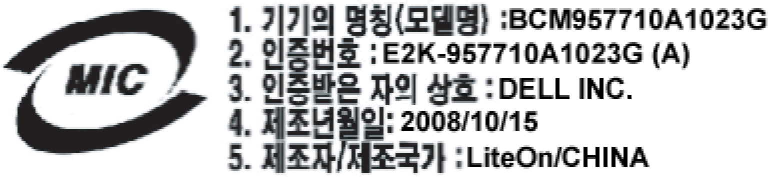
Marvell 57*xx* and 57*xxx* 10-gigabit Ethernet controller

- 957710A1023G
- 957711A1123G (E03D001)
- E02D001

Dell Inc.
Worldwide Regulatory Compliance, Engineering and Environmental Affairs
One Dell Way PS4-30
Round Rock, Texas 78682, USA
512-338-4400



1. 기기의 명칭(모델명) : 5709s-mezz
2. 인증번호 : E2K-5709s-mezz(A)
3. 인증받은 자의 상호 : DELL INC.
4. 제조년월일: 2008/08/12
5. 제조자/제조국가 : LiteOn/CHINA



1. 기기의 명칭(모델명) :BCM957710A1023G
2. 인증번호 :E2K-957710A1023G (A)
3. 인증받은 자의 상호 :DELL INC.
4. 제조년월일: 2008/10/15
5. 제조자/제조국가 :LiteOn/CHINA

방송통신위원회

E2K-E03D001 (A)

방송통신위원회

E2K-E02D001 (A)

19 Troubleshooting

Troubleshooting topics cover the following:

Hardware Diagnostics

Loopback diagnostic tests are available for testing the adapter hardware. These tests provide access to the adapter internal and external diagnostics, where packet information is transmitted across the physical link (for instructions and information on running tests in Windows environments, see the QCC GUI online help).

QCS CLI and QCC GUI Diagnostic Tests Failures

If any of the following tests fail while running the diagnostic tests from QCS CLI or QCC GUI, this may indicate a hardware issue with the NIC or LOM that is installed in the system.

- Control Registers
- MII Registers
- EEPROM
- Internal Memory
- On-Chip CPU
- Interrupt
- Loopback - MAC
- Loopback - PHY
- Test LED

Troubleshooting steps that may help correct the failure:

1. Remove the failing device and reseat it in the slot, ensuring the card is firmly seated in the slot from front to back.
2. Rerun the test.
3. If the card still fails, replace it with a different card of the same model and run the test. If the test passes on the known good card, contact your hardware vendor for assistance on the failing device.
4. Power down the machine, remove AC power from the machine, and then reboot the system.
5. Remove and re-install the diagnostic software.
6. Contact your hardware vendor.

QCS CLI and QCC GUI Network Test Failures

Typically, QCS CLI or QCC GUI network test failures are the result of a configuration problem on the network or with the IP addresses. The following steps are commonly performed when troubleshooting the network.

1. Verify that the cable is attached and you have proper link.
2. Verify that the drivers are loaded and enabled.
3. Replace the cable that is attached to the NIC or LOM.
4. Verify that the IP address is assigned correctly by issuing the command **`ipconfig`** or by checking the OS IP assigning tool.
5. Verify that the IP address is correct for the network to which the adapter or adapters are connected.

Checking Port LEDs

To check the state of the network link and activity, see "Network Link and Activity Indication" on page 7.

Troubleshooting Checklist

CAUTION

Before you open the cabinet of your server to add or remove the adapter, review "Safety Precautions" on page 19.

The following checklist provides recommended actions to take to resolve problems installing the Marvell 57*xx* and 57*xxx* adapter or running it in your system.

- Inspect all cables and connections. Verify that the cable connections at the network adapter and the switch are attached properly. Verify that the cable length and rating comply with the requirements listed in "Connecting the Network Cables" on page 20.
- Check the adapter installation by reviewing "Installation of the Add-In NIC" on page 20. Verify that the adapter is properly seated in the slot. Check for specific hardware problems, such as obvious damage to board components or the PCI edge connector.
- Check the configuration settings and change them if they are in conflict with another device.
- Verify that your server is using the latest BIOS.
- Try inserting the adapter in another slot. If the new position works, the original slot in your system may be defective.
- Replace the failed adapter with one that is known to work properly. If the second adapter works in the slot where the first one failed, the original adapter is probably defective.
- Install the adapter in another functioning system and run the tests again. If the adapter passed the tests in the new system, the original system may be defective.
- Remove all other adapters from the system and run the tests again. If the adapter passes the tests, the other adapters may be causing contention.

Checking if Current Drivers Are Loaded

Follow the appropriate procedure for your operating system to confirm if the current drivers are loaded.

Windows

See the QCC GUI online help for information on viewing vital information about the adapter, link status, and network connectivity.

Linux

To verify that the `bnx2.o` driver is loaded properly, issue the following command:

```
lsmod | grep -i <module name>
```

If the driver is loaded, the output of this command shows the size of the driver in bytes, the quantity of adapters configured, and their names. The following example shows the drivers loaded for the bnx2 module:

```
[root@test1]# lsmod | grep -i bnx2
bnx2                    199238  0
bnx2fc                  133775  0
libfcoe                  39764  2 bnx2fc,fcoe
libfc                   108727  3 bnx2fc,fcoe,libfcoe
scsi_transport_fc        55235  3 bnx2fc,fcoe,libfc
bnx2i                    53488  11
cnic                     86401  6 bnx2fc,bnx2i
libiscsi                 47617  8
be2iscsi,bnx2i,cxgb4i,cxgb3i,libcxgbi,ib_iser,iscsi_tcp,libiscsi_tcp
scsi_transport_iscsi     53047  8
be2iscsi,bnx2i,libcxgbi,ib_iser,iscsi_tcp,libiscsi
bnx2x                  1417947  0
libcrc32c                 1246  1 bnx2x
mdio                      4732  2 cxgb3,bnx2x
```

If you reboot after loading a new driver, you can issue the following command to verify that the currently loaded driver is the correct version.

```
modinfo bnx2
```

Following is a sample output.

```
[root@test1]# lsmod | grep -i bnx2
bnx2                    199238  0
```

Or, you can issue the following command:

```
[root@test1]# ethtool -i eth2
```

Following is a sample output.

```
driver: bnx2x
version: 1.78.07
firmware-version: bc 7.8.6
bus-info: 0000:04:00.2
```

If you loaded a new driver but have not yet booted, the `modinfo` command does not show the updated driver information. Instead, you can view the logs to verify that the proper driver is loaded and will be active upon reboot by issuing the following command:

```
dmesg | grep -i "Marvell" | grep -i "bnx2"
```

Running a Cable Length Test

For Windows operating systems, see the QCC GUI online help for information on running a cable length test. Cable analysis is not available for the 71*x*/578*xx* network adapters.

Testing Network Connectivity

NOTE

When using forced link speeds, verify that both the adapter and the switch are forced to the same speed.

Windows

Network connectivity can be tested using the feature in QCS CLI and QCC GUI.

An alternate method is to issue the `ping` command to determine if the network connection is working.

To test network connectivity in Windows:

1. Click **Start**, and then click **Run**.
2. In the **Open** box, type **`cmd`**, and then click **OK**.
3. To view the network connection to be tested, issue the following command:

   ```
   ipconfig /all
   ```

4. Issue the following command, and then press ENTER.

   ```
   ping <IP address>
   ```

The ping statistics that are displayed indicate whether the network connection is working or not.

Linux

To verify that the Ethernet interface is up and running, issue **`ifconfig`** to check the status of the Ethernet interface. It is possible to use **`netstat -i`** to check the statistics on the Ethernet interface. For information on `ifconfig` and `netstat`, see Chapter 7 Linux Driver Software.

Ping an IP host on the network to verify connection has been established.

From the command line, issue the **`ping <IP address>`** command, and then press ENTER.

The ping statistics that are displayed indicate whether or not the network connection is working.

Microsoft Virtualization with Hyper-V

Microsoft Virtualization is a hypervisor virtualization system for Windows Server. This section is intended for those who are familiar with Hyper-V, and it addresses issues that affect the configuration of 57*xx* and 57*xxx* network adapters and teamed network adapters when Hyper-V is used. For more information on Hyper-V, see:

http://www.microsoft.com/windowsserver2008/en/us/hyperv.aspx

Table 19-1 identifies Hyper-V supported features that are configurable for 57*xx* and 57*xxx* network adapters. This table is not an all-inclusive list of Hyper-V features.

Table 19-1. Configurable Network Adapter Hyper-V Features

Feature	Supported in Windows Server Version 2012 and Later	Comments and Limitations
IPv4	Yes	—
IPv6	Yes	—
IPv4 large send offload (LSO) (parent and child partition)	Yes	—
IPv4 checksum offload (CO) (parent and child partition)	Yes	—
IPv6 LSO (parent and child partition)	Yes	* When bound to a virtual network, OS limitation.
IPv6 CO (parent and child partition)	Yes	* When bound to a virtual network, OS limitation.

Table 19-1. Configurable Network Adapter Hyper-V Features (Continued)

Feature	Supported in Windows Server Version 2012 and Later	Comments and Limitations
Jumbo frames	Yes	* OS limitation.
RSS	Yes	* OS limitation.
RSC	Yes	* OS limitation.
SR-IOV	Yes	* OS limitation.

NOTE

For full functionality, ensure that Integrated Services, which is a component of Hyper-V, is installed in the guest operating system (child partition).

Single Network Adapter

Configuration of the Microsoft virtualization with Hyper-V for a single network adapter differs depending on the Windows Server version used.

Windows Server 2012, 2012 R2, 2016, 2019, and Azure Stack HCI

When configuring a 57*xx* and 57*xxx* network adapter on a Hyper-V system, be aware of the following:

- An adapter that is to be bound to a virtual network must not be configured for VLAN tagging through the driver's advanced properties. Instead, Hyper-V should manage VLAN tagging exclusively.
- The locally administered address (LAA) set by Hyper-V takes precedence over the address set in the adapter's advanced properties.
- The LSO and CO features in the guest OS are independent of the network adapter properties.
- To allow jumbo frame functionality from the guest OS, both the network adapter and the virtual adapter must have jumbo frames enabled. Set the Jumbo MTU property for the network adapter to allow traffic of large MTU from within the guest OS. Set the jumbo packet of the virtual adapter to segment the sent and received packets.

Teamed Network Adapters

Table 19-2 identifies Hyper-V supported features that are configurable for 57*xx* and 57*xxx* teamed network adapters. This table is not an all-inclusive list of Hyper-V features.

Table 19-2. Configurable Teamed Network Adapter Hyper-V Features

Feature	Supported in Windows Server Version 2012	Comments and Limitations
Smart Load Balancing and Failover (SLB) team type	Yes	Multimember SLB team allowed with latest QLASP6 version. NOTE: VM MAC is not presented to external switches.
Link aggregation (IEEE 802.3ad LACP) team type	Yes	—
Generic trunking (FEC/GEC) 802.3ad draft static team type	Yes	—
Failover	Yes	—
LiveLink	Yes	—
Large send offload (LSO)	Yes	* Conforms to miniport limitations outlines in Table 19-1.
Checksum offload (CO)	Yes	* Conforms to miniport limitations outlines in Table 19-1.
Hyper-V VLAN over an adapter	Yes	—
Hyper-V VLAN over a teamed adapter	Yes	—
Hyper-V VLAN over a VLAN	Limited*	Only an untagged VLAN.
Hyper-V virtual switch over an adapter	Yes	—
Hyper-V virtual switch over a teamed adapter	Yes	—
Hyper-V virtual switch over a VLAN	Yes	—
iSCSI boot	No*	* Remote boot to SAN is supported.
Virtual machine queue (VMQ)	Yes	See “Configuring VMQ with SLB Teaming” on page 288.

Table 19-2. Configurable Teamed Network Adapter Hyper-V Features (Continued)

Feature	Supported in Windows Server Version 2012	Comments and Limitations
RSC	Yes	—

Configuring VMQ with SLB Teaming

When a Hyper-V server is installed on a system configured to use Smart Load Balance and Failover (SLB) type teaming, you can enable virtual machine queue (VMQ) to improve overall network performance. VMQ enables delivering packets from an external virtual network directly to virtual machines defined in the SLB team, eliminating the need to route these packets and, thereby, reducing overhead.

To create a VMQ-capable SLB team:

1. Create an SLB team. If using the Teaming Wizard, when you select the SLB team type, also select **Enable HyperV Mode**. If using Expert mode, enable the property on the **Create Team** or **Edit Team** pages.
2. Follow these instructions to add the required registry entries in Windows:

   http://technet.microsoft.com/en-us/library/gg162696%28v=ws.10%29.aspx
3. For each team member on which you want to enable VMQ, modify the following registry entry and configure a unique instance number (in the following example, it is set to `0026`):

```
[HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\
      {4D36E972-E325-11CE-BFC1-08002BE10318}\0026]
"*RssOrVmqPreference"="1"
```

Removing the Marvell 57*xx* and 57*xxx* Device Drivers

Uninstall the Marvell 57*xx* and 57*xxx* device drivers from your system only through the InstallShield wizard. Uninstalling the device drivers with Device Manager or any other means may not provide a clean uninstall and may cause the system to become unstable. For information on uninstalling Marvell 57*xx* and 57*xxx* device drivers, see "Removing the Device Drivers" on page 95. When removing the device drivers, QLogic Control Suite is also removed as well as all other management applications.

If you manually uninstalled the device drivers with Device Manager and attempted to reinstall the device drivers but could not, run the **Repair** option from the InstallShield wizard. For information on repairing Marvell 57*xx* and 57*xxx* device drivers, see “Repairing or Reinstalling the Driver Software” on page 94.

Upgrading Windows Operating Systems

This section covers Windows upgrades from Windows Server 2008 R2 to Windows Server 2012.

Prior to performing an OS upgrade when a Marvell 57*xx* and 57*xxx* adapter is installed on your system, Marvell recommends the following procedure.

1. Save all team and adapter IP information.
2. Uninstall all Marvell drivers using the installer.
3. Perform the Windows upgrade.
4. Reinstall the latest Marvell adapter drivers and the QLogic Control Suite application.

Marvell Boot Agent

Problem: Unable to obtain network settings through DHCP using PXE.

Solution: For proper operation make sure that the Spanning Tree Protocol (STP) is disabled or that portfast mode (for Cisco) is enabled on the port to which the PXE client is connected. For instance, set **`spantree portfast 4/12 enable`**.

Linux

Problem: 57*xx* and 57*xxx* devices with SFP+ Flow Control default to **Off** rather than **Rx/Tx Enable**.

Solution: The Flow Control default setting for revision 1.6.*x* and later has been changed to **Rx Off and Tx Off** because SFP+ devices do not support auto-negotiation for flow control.

Problem: On kernels older than 2.6.16 when 16 partitions are created on a server containing two 57711 network adapters, not all partitions come up and an error indicating a shortage of space appears.

Solution: On architectures where the default `vmalloc` size is relatively small and not sufficient to load many interfaces, use `vmalloc=<size>` during boot to increase the size.

Problem: Routing does not work for 57*xx* and 57*xxx* 10GbE network adapters installed in Linux systems.

Solution: For 57*xx* and 57*xxx* 10GbE network adapters installed in systems with Linux kernels older than 2.6.26, disable TPA with either ethtool (if available) or with the driver parameter (see "disable_tpa" on page 45). Use ethtool to disable TPA (LRO) for a specific 57*xx* and 57*xxx* 10GbE network adapter.

Problem: On a 57*xx* and 57*xxx* 1GbE network adapter in a C-NIC environment, flow control does not work.

Solution: Flow control is working, but in a C-NIC environment, it has the appearance that it is not. The network adapter is capable of sending pause frames when the on-chip buffers are depleted, but the adapter also prevents the head-of-line blocking of other receive queues. Because the head-of-line blocking causes the on-chip firmware to discard packets inside the on-chip receive buffers, in the case a specific host queue is depleted, the on-chip receive buffers are rarely depleted, therefore, it may appear that flow control is not functioning.

Problem: Errors appear when compiling driver source code.

Solution: Some installations of Linux distributions do not install the development tools by default. Before compiling driver source code, ensure that the development tools for the Linux distribution you are using are installed.

Problem: L4 iSCSI offload boot from SAN fails (iscsiuio crashes). This problem is seen on Linux OSs based on 4.5 kernel and later.

Solution: To override the kernel config option `CONFIG_IO_STRICT_DEVMEM` and avoid crash of iscsiuio on OS boot, edit the OS grub and add the kernel command line parameter `iomem=relaxed` during the start of OS installation or OS boot.

Problem: iSCSI-Offload boot from SAN fails to boot after installation.
The iSCSI boot from SAN process is divided into two parts: pre switch-root and post switch root.

During pre switch-root, when the drivers load, the open-iSCSI tool iscsistart establishes the connection with the target and discovers the remote LUN. Then iscsistart starts a session using the iBFT information.

The iscsistart utility program is not run to manage connection with target. (Its primary use is to start sessions used for iSCSI root boot.)

After the post switch-root, as a part of initialize boot process, the open-iscsi tool iscsid takes over the pre switch-root iSCSI connection. Therefore, iscsid manages the iscsi connection with target during recovery.

There is gap between when the pre switch-root iscsistart establishes the connection and when iscsid takes over iSCSI connection. During this time, the OS boot process no way to recovery the iSCSI connection. In some cases, the bnx2x NIC interface’s link 'flaps' during this gap, the iSCSI connection is interrupted, and the iSCSI connection recovery or retries fail.

Solution: Avoid the bnx2x NIC interface’s link flap, load the bnx2x driver with the module parameter `disable_tpa=1`. Set this parameter through either a kernel grub command line or `/etc/modprobe.d/bnx2x.conf` file configuration.

NPAR

Problem: The following error message appears if the storage configurations are not consistent for all four ports of the device in NPAR mode:

```
PXE-M1234: NPAR block contains invalid configuration during boot.
```

A software defect can cause the system to be unable to BFS boot to an iSCSI or FCoE target if an iSCSI personality is enabled on the first partition of one port, whereas an FCoE personality is enabled on the first partition of another port. The MBA driver performs a check for this configuration and prompts the user when it is found.

Solution: If using the 7.6.*x* firmware and driver, to workaround this error, configure the NPAR block such that if iSCSI or FCoE is enabled on the first partition, the same must be enabled on all partitions of all four ports of that device.

Kernel Debugging Over Ethernet

Problem: When attempting to perform kernel debugging over an Ethernet network on a Windows 8.0 or Windows Server 2012 system, the system does not boot. This problem may occur with some adapters on systems where the Windows 8.0 or Windows Server 2012 OS is configured for unified extensible firmware interface (UEFI) mode. You may see a firmware error indicating that a Non Maskable Interrupt exception was encountered during the UEFI pre-boot environment.

Solution: Refer to the Microsoft knowledge base topic number 2920163, Non Maskable Interrupt error during boot on a system which has been configured for kernel debugging over Ethernet.

Miscellaneous

Problem: The 57810 10GbE NIC does not support 10Gbps or 1Gbps WoL link speed.

Solution: The 57810 10GbE NIC can only support 100Mbps WoL link speed due to power consumption limitations.

Problem: iSCSI Crash Dump is not working in Windows.

Solution: After upgrading the device drivers using the installer, the iSCSI crash dump driver is also upgraded, and **iSCSI Crash Dump** must be re-enabled from the **Advanced** section of the QCS Configurations page.

Problem: The Marvell 57*xx* and 57*xxx* adapter may not perform at optimal level on some systems if it is added after the system has booted.

Solution: The system BIOS in some systems does not set the cache line size and the latency timer if the adapter is added after the system has booted. Reboot the system after the adapter has been added.

Problem: Cannot configure Resource Reservations in QCC after SNP is uninstalled.

Solution: Reinstall SNP. Prior to uninstalling SNP from the system, ensure that NDIS is enabled by selecting the check box on the Resource Configuration window, available from the **Resource Reservations** section of the Configurations page. If NDIS is disabled and SNP is removed, there is no access to re-enable the device.

Problem: A DCOM error message (event ID 10016) appears in the System Even Log during the installation of the Marvell adapter drivers.

Solution: This problem is a Microsoft issue. For more information, see Microsoft knowledge base KB913119 at http://support.microsoft.com/kb/913119.

Problem: Performance is degraded when multiple 57710 network adapters are used in a system.

Solution: Ensure that the system has at least 2GB of main memory when using up to four network adapters and 4GB of main memory when using four or more network adapters.

Problem: The network adapter has shut down and an error message appears indicating that the fan on the network adapter has failed.

Solution: The network adapter was shut down to prevent permanent damage. Contact Dell Support for assistance.

Problem: When using a 57840 four-port adapter in a blade server, ports 3 and 4 show no link.

Solution: The I/O (switch) module must support 32 internal ports. If it does not, ports 3 and 4 cannot establish a link.

A Revision History

Document Revision History	
Revision A, February 18, 2015	
Revision B, July 29, 2015	
Revision C, March 24, 2016	
Revision D, April 8, 2016	
Revision E, February 2, 2017	
Revision F, August 25, 2017	
Revision G, December 19, 2017	
Revision H, March 15, 2018	
Revision J, April 13, 2018	
Revision K, October 25, 2018	
Revision L, June 7, 2019	
Revision M, October 16, 2019	
Revision N, April 3, 2020	
Revision P, July 7, 2020	
Revision R, January 21, 2020	
Changes	**Sections Affected**
Added support for the following OSs: RHEL 7.9, 8.2, 8.3 SLES 15 SP2 Ubuntu 20.04 Azure Stack HCI Removed support for the following OSs: Windows 2012 (all versions) RHEL 7.6, 7.7, 8.0, 8.1 SLES 12 SP4, 15 VMware ESXi 6.5 U3 Citrix Hypervisor 7.1 CU	All
Removed **BCM** nomenclature in product names.	All
Removed information about QLASP, which is not supported on Windows Server 2016 and later.	All
Removed the NOTE about QCC GUI being the only GUI management tool across adapters.	"Preface" on page xxi

Removed “Downloading Documents” section.	Preface
Added bullet for VMDirectPath I/O.	“Features” on page 2
Revised the second paragraph to indicate that the `iface` file information is for all SLES versions.	“Bind iSCSI Target to Marvell iSCSI Transport Name” on page 59
Removed sub-section for bnx2x.	“Driver Defaults” on page 81
Removed sub-section for bnx2, bnx2x.	“Unloading and Removing Drivers” on page 83
Removed Driver Messages section (was for bnx2, bnx2x)	—
Changed section title and Steps 1 through 5.	“Booting from RHEL 7.x Installation Media With the FCoE Target Already Connected” on page 229
Indicated that Step 3 is not required for SLES 12 and later, or RHEL 7.*x* and later.	“Driver Upgrade on Linux Boot from SAN Systems” on page 236
Added TUV IEC 62368-1 2nd Edition and 3rd Edition	“Product Safety” on page 260
Added the Marvell model number associated with 957810A1006G.	“FCC, Class A” on page 262

MARVELL™

Marvell first revolutionized the digital storage industry by moving information at speeds never thought possible. Today, that same breakthrough innovation remains at the heart of the company's storage, networking and connectivity solutions. With leading intellectual property and deep system-level knowledge, Marvell semiconductor solutions continue to transform the enterprise, cloud, automotive, industrial, and consumer markets. For more information, visit www.marvell.com.

Doc No. BC0054508-00 Rev. R Revised: January 21, 2021