

QLogic HCA and QLogic OFED Software Install Guide

QLogic OFED Version 1.4

Information furnished in this manual is believed to be accurate and reliable. However, QLogic Corporation assumes no responsibility for its use, nor for any infringements of patents or other rights of third parties which may result from its use. QLogic Corporation reserves the right to change product specifications at any time without notice. Applications described in this document for any of these products are for illustrative purposes only. QLogic Corporation makes no representation nor warranty that such applications are suitable for the specified use without further testing or modification. QLogic Corporation assumes no responsibility for any errors that may appear in this document.

No part of this document may be copied nor reproduced by any means, nor translated nor transmitted to any magnetic medium without the express written consent of QLogic Corporation. In accordance with the terms of their valid QLogic agreements, customers are permitted to make electronic and paper copies of this document for their own exclusive use.

The QHT7040, QHT7140, QLE7140, QLE7240, and QLE7280 QLogic Host Channel Adapters are covered by the following patent: 7308535.

Document Revision History	
Rev. 1.0, 8/20/2005	
Rev. 1.1, 11/15/05	
Rev. 1.2, 02/15/06	
Rev. 1.3, Beta 1, 4/15/06	
Rev 1.3, 6/15/06	
Rev. 2.0 Beta, 8/15/06, QLogic Version Number IB0056101-00 A	
Rev. 2.0 Beta 2 10/15/06, QLogic Version Number IB0056101-00 B	
Rev. 2.0 11/30/06, QLogic Version Number IB0056101-00 C	
Rev. 2.0 3/23/07, QLogic Version Number IB0056101-00 D	
Rev. 2.1 8/24/07, QLogic Version Number IB0056101-00 E	
Rev. 2.2 5/27/08, QLogic Version Number IB0056101-00 F	
Rev. 2.2 9/5/08, QLogic Version Number IB0056101-00 G	
Rev. QLogic OFED 1.4 4/7/09, QLogic Version Number IB0056101-00 G.02	
Changes	Sections Affected
Product name changed from <i>InfiniPath</i> to <i>QLogic OFED</i> . Version number is set to 1.4. Instances of <i>InfiniPath</i> changed where appropriate; some file-names and output messages keep old name.	All
Updated Contact Information.	"Contact Information" on page 1-5
Removed <i>InfiniPath</i> from 1st paragraph; replaced with <i>QLogic adapters</i> , etc.	"Hardware Installation" on page 4-1
Changed section title from What's New in this Release to Features.	"Features" on page 2-1

Replace EM64T with 64-bit Intel Xeon.	“Supported Linux Distributions” on page 5-1
Added more information on VNIC interface.	“OpenSM” on page 6-3
Split the Install section into three separate sections.	“Software Installation” on page 5-1 “Configuring Drivers and Services” on page 6-1 “Installation Verification and Additional Settings” on page 7-1
Modified installation checklist. Clarified which drivers are configured, which are optional. Added two methods of installation to the list.	“Software Installation” on page 3-2
Remove sections on configuring ipath_ether. ipath_ether now deprecated.	Was “Configuring the ipath_ether Network Interface” on page 5-12
Updated the supported distributions information.	Table 5-1 on page 5-2
Distribution identifiers are now RHEL4, RHEL5, and SLES10.	Table 5-2 on page 5-2
Delete reference to ipath_ether.	“Removing Software Packages” on page 5-24
Updated Lustre information; patches are no longer needed.	“Installing Lustre” on page 5-23
Updated compiler support information.	“Compiler Support” on page 5-3
Updated information on software components.	“Software Components” on page 2-4
Removed paragraph about this release adding support for QLE7240/7280, since this is outdated.	“Features” on page 2-1
Added new “Configuration Issues” to Troubleshooting.	“ibsrpdm Command Hangs When Two HCAs are Installed but Only Unit 1 is Connected to the Switch” on page A-5
Deleted Troubleshooting item: ifup on ipath_ether on SLES 10 Reports “unknown device”	Appendix A
Deleted entries pertaining to ipath_ether in this table.	Table C-1 on page C-1

<p>Sections rearranged, renamed, and expanded due to multiple install methods. Previous sections: Installing the InfiniPath and Open fabrics RPMs Downloading and Unpacking the InfiniPath and OpenFabrics Software</p>	<p>“Choose the Appropriate Download Files” on page 5-4 “Install QLogicIB-Basic with the Installer Tool” on page 5-7 “About rpm Installation” on page 5-13 “Install QLogic OFED Using Rocks” on page 5-20 “Install QLogic OFED Using a Platform OCS Kit” on page 5-22 “Using rpm to Install InfiniPath and OpenFabrics” on page 5-14</p>
<p>Added new section about installing using the Installer Tool.</p>	<p>“Install QLogicIB-Basic with the Installer Tool” on page 5-7</p>
<p>Moved introductory information about RPM installation before sections on installing using the two different methods.</p>	<p>“About rpm Installation” on page 5-13</p>
<p>Removed paragraph about kernel module support being part of InfiniPath RPMs; was relevant to previous release.</p>	<p>Was in old section “Installing the InfiniPath and Open fabrics RPMs”</p>
<p>Added note about need to use prefixed install if using QLogic MPI with mpi-selector.</p>	<p>“Using rpm to Install InfiniPath and OpenFabrics” on page 5-14</p>
<p>Modified module for installed layout. New location for ipath modules. Location of other OFED modules listed here.</p>	<p>“Installed Layout” on page 5-23</p>
<p>Modified module for driver overview.</p>	<p>“InfiniPath and OpenFabrics Driver Overview” on page 6-1</p>
<p>Modified module for configuring IPoIB. Only restart information has been changed.</p>	<p>“Configuring the IPoIB Network Interface” on page 6-2</p>
<p>Modified module for configuring VNIC interface. Changed infinipath start/stop commands, and added some introductory material.</p>	<p>“SRP stands for SCSI RDMA Protocol. It was originally intended to allow the SCSI protocol to run over InfiniBand for Storage Area Network (SAN) usage. SRP interfaces directly to the Linux file system through the SRP Upper Layer Protocol. SRP storage can be treated as another device.” on page 6-4</p>
<p>Modified module for configuring SRP.</p>	<p>“SRP” on page 6-4</p>
<p>Modified module for uDAPL configuration.</p>	<p>“MPI over uDAPL” on page 6-13</p>
<p>Deleted reference to MTRR BIOS setting.</p>	<p>“Configuring the BIOS” on page 4-4 “BIOS Settings” on page A-2</p>

New Appendix for Write Combining. MTRR settings information moved here.	“Write Combining” on page B-1
Minor wording change to introduction section.	“Software Installation” on page 5-1
Added new issue in Troubleshooting.	“openmpi_gcc Fails to Install Because of Dependency on gfortran (RHEL 4)” on page A-4
Added new issue in Troubleshooting.	“Outdated ipath_ether Configuration Setup Generates Error” on page A-5
Configuration file <code>qlogic_vnic.cfg</code> changed to <code>qlgc_vnic.cfg</code> .	Table C-1 on page C-1
Added new section.	“Uninstalling Software with Rocks or Platform OCS” on page 5-25
Added new section.	“Install QLogic OFED Using Rocks” on page 5-20
Change name of Appendix RPM Descriptions to Package Descriptions. Updated information.	“Package Descriptions” on page D-1
Removed phrase “and Transmission Control Protocol (TCP)” from introduction; related to now-obsolete <code>ipath_ether</code> .	“Interoperability” on page 1-3
Removed this issue from the Troubleshooting section.	Was “OpenFabrics Library Dependencies”
Combined What’s New in This Release and Features section.	“Feature Overview” on page 2-1
Deleted Note about Fedora Core 6 not supported in this release.	This was noted in InfiniPath 2.2.1 release notes.
Removed this line “Please check the QLogic web site for updated information on supported compilers.” This information is not on the web site.	“Software Components” on page 2-4
Added more details about setting the switch MTU default to 4K.	“Other Configuration: Changing the MTU Size” on page 6-14
Deleted the section “Installing QLogic MPI in an Alternate Location with rpm”; merged instructions into rpm install section.	“Using rpm to Install InfiniPath and OpenFabrics” on page 5-14
Minor text changes to this section. Modified title from Additional Installation Instructions to Install Additional Software. Moved this section to be with the rest of the installation instructions.	“Install Additional Software” on page 5-23

Changed title of Troubleshooting issue. Was "ipath_ether Configuration Setup Generates Error"	"Outdated ipath_ether Configuration Setup Generates Error" on page A-5
Change main configuration file from /etc/sysconfig/infinipath to /etc/infiniband/openib.conf. The infinipath file is still to enable or disable the ipath_mtrr script.	Table C-1 on page C-1 and "SRP" on page 6-4.
Updated note about setting datagram mode.	"Configuring the IPoB Network Interface" on page 6-2
Deleted Troubleshooting Issue OpenFabrics Dependencies.	"Software Installation Issues" on page A-2
Deleted Troubleshooting Issue Version Number Conflict with opensm-* on RHEL5 Systems	"Software Installation Issues" on page A-2
Two new tables showing OS package requirements. Replaces some bullet points in "Distribution Identifiers" on page 5-2	Table on page 5-4 and Table on page 5-4
Removed section "Check for Missing Files or RPMs"	Was "Check for Missing files or RPMs" on page 5-5
Merged the two sections Downloading the QLogic OFED Software and Choose the Appropriate Download Files	"Choose the Appropriate Download Files" on page 5-4
Removed Supported Compilers and Supported Linux Distributions from the 2 Feature Overview chapter. Information is still in Software Installation chapter.	2 Feature Overview
Moved table "InfiniPath and OpenFabrics RPMs to Use for Each Node in a Cluster" from Software Installation section to Appendix	"Package Descriptions" on page D-1
Consolidated Configuring the InfiniPath Drivers and InfiniPath and OpenFabrics Driver Overview	6 Configuring Drivers and Services
opensm is now off by default	Various
Merged "Install on an Unsupported Distribution with the rpm Command" with the Rebuilding...section.	Now "Rebuilding the kernel-ib Driver on an Unsupported Distribution or an Unsupported Distribution/Kernel Pair" on page 5-19
Removed \$ or # at beginning of lines indicating commands. The #, if cut and pasted from the document, may be interpreted as a shell command.	All. Also Table 1-1
Updated footnote b.	Table D-11

Adapter Settings section name changed. Also added new bullet point "Check the PCIe bus width"	"Adapter and Other Settings" on page 7-1
Updated tables of RPMs. Moved all documentation RPMs together. New table for OpenSM-Devel RPMs. Updated OtherMPiIs.	All tables in D Package Descriptions
Deleted table 5-7 RPMs to Install, as information is repeated in "Using rpm to Install InfiniPath and OpenFabrics" on page 5-14	"Using rpm to Install InfiniPath and OpenFabrics" on page 5-14
Moved RPM Organization section under RPM Installation.	"RPM Organization" on page 5-16
Merge two sections Uninstalling InfiniPath and OpenFabrics RPMs and Uninstalling OFED 1.3 Software.	"Uninstalling InfiniPath and OpenFabrics RPMs" on page 5-24
Reference to this location <code>/sys/bus/pci/drivers/ib_ipath/00/</code> changed to <code>/sys/class/infiniband/ipath*/device/</code>	"Form Factors" on page 4-2
Removed note: "OpenFabrics programs (32-bit) using the Verbs interfaces are not supported in this release, but may be supported in a future release."	"Software Components" on page 2-4

Notes

Draft

Table of Contents

1	Introduction	
	Who Should Read this Guide	1-1
	How this Guide is Organized	1-1
	Overview	1-2
	Interoperability	1-3
	Conventions Used in this Guide	1-3
	Documentation	1-4
	Contact Information	1-5
2	Feature Overview	
	Features	2-1
	Other Changes	2-2
	Continued Support	2-2
	Software Components	2-4
3	Step-by-Step Installation Checklist	
	Hardware Installation	3-1
	Software Installation	3-2
4	Hardware Installation	
	Hardware Installation Requirements	4-1
	Hardware	4-1
	Form Factors	4-2
	Cabling and Switches	4-3
	Optical Fibre Option	4-4
	Configuring the BIOS	4-4
	Safety with Electricity	4-5
	Unpacking Information	4-5
	Verify the Package Contents	4-5
	List of the Package Contents	4-5
	Unpacking the QLogic Adapter	4-8
	Hardware Installation	4-9

Hardware Installation for QLE7240, QLE7280, or QLE7140 with PCI Express Riser	4-9
Dual Adapter Installation.	4-9
Installation Steps.	4-9
Hardware Installation for QHT7140 with HTX Riser	4-12
Hardware Installation for QLE7240, QLE7280, and QLE7140 Without a PCI Express Riser.	4-15
Hardware Installation for the QHT7140 Without an HTX Riser	4-16
Switch Configuration and Monitoring	4-17
Cabling the Adapter to the InfiniBand Switch.	4-17
Completing the Installation	4-18

5 Software Installation

Cluster Setup	5-1
Types of Nodes in a Cluster Environment.	5-1
Supported Linux Distributions	5-1
Distribution Identifiers	5-2
Compiler Support	5-3
Setting Up Your Environment	5-3
Choose the Appropriate Download Files	5-4
Install QLogicIB-Basic with the Installer Tool	5-7
About <code>rpm</code> Installation	5-13
Using <code>rpm</code> to Install InfiniPath and OpenFabrics	5-14
RPM Organization	5-16
Install QLogic OFED User-level Software with the <code>rpm</code> Command	5-17
Rebuilding or Reinstalling the <code>kernel-ib</code> Driver with <code>rpm</code> After a Kernel Upgrade	5-19
Rebuilding the <code>kernel-ib</code> Driver on an Unsupported Distribution or an Unsupported Distribution/Kernel Pair.	5-19
Install QLogic OFED Using Rocks	5-20
Install Frontend and Compute Nodes	5-20
Rocks Installation on an Existing Frontend Node	5-21
Install QLogic OFED Using a Platform OCS Kit.	5-22
Install FastFabric Software CD/ISO Image	5-22
Install Additional Software	5-23
Installing Lustre	5-23
Installed Layout	5-23
Removing Software Packages	5-24
Uninstalling Using the Installer Tool.	5-24
Uninstalling InfiniPath and OpenFabrics RPMs	5-24

	Uninstalling Software with Rocks or Platform OCS	5-25
	Downgrading RPMs	5-25
6	Configuring Drivers and Services	
	InfiniPath and OpenFabrics Driver Overview	6-1
	OpenFabrics Drivers and Services Configuration and Startup	6-1
	Configuring the IPoIB Network Interface	6-2
	OpenSM	6-3
	SRP	6-4
	Using QLogic SRP	6-4
	Using OFED SRP	6-4
	Configuring and Administering the VNIC Interface	6-6
	MPI over uDAPL	6-13
	Other Configuration: Changing the MTU Size	6-14
	Managing the InfiniPath Driver	6-15
	Configure InfiniPath Driver State	6-16
	Start, Stop or Restart InfiniPath	6-16
	Unloading the Driver/Modules Manually	6-17
	Further Information on Configuring and Loading Drivers	6-17
7	Installation Verification and Additional Settings	
	LED Link and Data Indicators	7-1
	Adapter and Other Settings	7-1
	Customer Acceptance Utility	7-2
A	Installation Troubleshooting	
	Hardware Issues	A-1
	Node Spontaneously Reboots	A-1
	Some HTX Motherboards May Need Two or More CPUs in Use	A-1
	BIOS Settings	A-2
	Enable Advanced Configuration and Power Interface (ACPI)	A-2
	Issue with Supermicro® H8DCE-HTe and QHT7040	A-2
	Software Installation Issues	A-2
	Missing Kernel RPM Errors	A-2
	Resolving Conflicts	A-4
	openmpi_gcc Fails to Install Because of Dependency on gfortran (RHEL 4)	A-4
	mpirun Installation Requires 32-bit Support	A-4
	Lockable Memory Error on Initial Installation of InfiniPath	A-5
	Configuration Issues	A-5

<code>ibsrpdm</code> Command Hangs When Two HCAs are Installed but Only Unit 1 is Connected to the Switch	A-5
Outdated <code>ipath_ether</code> Configuration Setup Generates Error	A-5

B Write Combining

Introduction	B-1
Verify Write Combining is Working	B-1
PAT and Write Combining	B-2
MTRR Mapping and Write Combining	B-2
Edit BIOS Settings to Fix MTRR Issues	B-2
Use the <code>ipath_mtrr</code> Script to Fix MTRR Issues	B-3

C Configuration Files

D Package Descriptions

Package Names with the QLogicIB-Basic Download	D-1
Different Nodes May Use Different RPMs	D-1
InfiniPath RPM Version Numbers and Identifiers	D-2
OpenFabrics RPM Names	D-2
InfiniPath and OpenFabrics RPMs	D-2
Documentation RPMs	D-3
InfiniPath RPMs	D-3
OpenFabrics RPMs	D-4
Other HCAs	D-9
Other MPIs	D-10

List of Figures

Figure	Page
4-1 QLogic QLE7280 with IBA7220 ASIC	4-7
4-2 QLogic QLE7140 Card with Riser, Top View	4-7
4-3 QLogic QHT7040/QHT7140 Full and Low Profile Cards with Riser, Top View	4-8
4-4 PCIe Slot in a Typical Motherboard	4-10
4-5 QLogic PCIe HCA Assembly with Riser Card	4-11
4-6 Assembled PCIe HCA with Riser	4-12
4-7 HTX Slot	4-13
4-8 QLogic QHT7140 Adapter with Riser Card	4-14
4-9 Assembled QHT7140 with Riser	4-15
4-10 QHT7140 Without Riser Installed in a 3U Chassis	4-17

List of Tables

Table	Page
1-1 Typographical Conventions	1-3
2-1 QLogic Adapter Model Numbers	2-3
4-1 Adapter Models and Related Platforms	4-1
4-2 QLogic InfiniBand Cables	4-3
5-1 InfiniPath/OpenFabrics Supported Distributions and Kernels	5-2
5-2 Distribution Identifiers	5-2
5-3 Required OS Packages	5-4
5-4 Specific Component Requirements	5-4
5-5 Available Packages for QLogic OFED 1.4 Release	5-5
5-6 INSTALL Options	5-12
7-1 ipath_checkout Options	7-3
C-1 Configuration Files	C-1
D-1 Documentation/RPMs	D-3
D-2 InfiniPath/RPMs	D-3
D-3 InfiniPath-Devel/RPMs	D-4
D-4 InfiniPath-MPI/RPMs	D-4
D-5 OpenFabrics/RPMs	D-4
D-6 OpenFabrics-Devel/RPMs	D-7
D-7 OpenSM/RPM	D-8
D-8 OpenSM-Devel/RPM	D-9
D-9 Other HCAs/RPMs	D-9
D-10 Other HCAs-Devel/RPMs	D-9
D-11 OtherMPIs/RPMs	D-10

Notes

Draft

1 Introduction

This chapter describes the contents, intended audience, and organization of the *QLogic HCA and QLogic OFED Software Install Guide*.

The *QLogic HCA and QLogic OFED Software Install Guide* contains instructions for installing the QLogic Host Channel Adapters (HCAs) and the QLogic InfiniPath and OpenFabrics software. The following adapters are covered in this guide:

- QLE7140 PCI Express[®] (PCIe)
- QLE7240 PCI Express
- QLE7280 PCI Express
- QHT7040/QHT7140 HyperTransport Expansion (HTX[™])

Who Should Read this Guide

This installation guide is intended for cluster administrators responsible for installing the QLogic QLE7140, QLE7240, QLE7280 or QHT7040/QHT7140 adapter and QLogic InfiniPath software on their Linux[®] cluster. Additional detailed installation information and instructions for administering the QLogic cluster can be found in the *QLogic HCA and QLogic OFED Software Users Guide*.

The *QLogic HCA and QLogic OFED Software Install Guide* assumes that you are familiar with both cluster networking and the specific hardware that you plan to use. Before installing the HCA, you should have basic knowledge of your host and target operating systems, and working knowledge of message passing concepts.

This document does not contain all the information you need to use basic Linux commands or to perform all system administration tasks. For this information, see the software documentation you received with your system.

How this Guide is Organized

The *QLogic HCA and QLogic OFED Software Install Guide* is organized into these sections:

- [Section 1, Introduction](#), contains an overview of the HCAs and software, describes interoperability with other products, lists all related documentation, and provides QLogic contact information.
- [Section 2, Feature Overview](#), contains features for this release, the supported QLogic adapter models, supported distributions and kernels, and a list of the software components.

- [Section 3, Step-by-Step Installation Checklist](#), provides a high-level overview of the hardware and software installation procedures.
- [Section 4, Hardware Installation](#), includes instructions for installing the QLogic QLE7140, QLE7240, QLE7280, QHT7040, and QHT7140 HCAs.
- [Section 5, Software Installation](#), includes instructions for installing the QLogic InfiniPath and OpenFabrics software.
- [Section 6, Configuring Drivers and Services](#), includes instructions for configuring the QLogic InfiniPath and OpenFabrics drivers and services.
- [Section 7, Installation Verification and Additional Settings](#), describes tools for verifying the installation, and adapter settings for best performance.
- [Appendix A, Installation Troubleshooting](#) contains troubleshooting information about issues that may occur during installation.
- [Appendix B, Write Combining](#) contains information about settings that will ensure better performance.
- [Appendix C, Configuration Files](#) contains descriptions of the configuration used by the QLogic InfiniPath and OpenFabrics software.
- [Appendix D, Package Descriptions](#) contains RPM Descriptions.
- [Index](#) lists major subjects and concepts with page numbers for easy reference.

Overview

The material in this documentation pertains to a QLogic OFED *cluster*. A cluster is defined as a collection of nodes, each attached to an InfiniBand™-based fabric through the QLogic interconnect. The nodes are Linux-based computers, each having up to 16 processors.

The QLogic HCAs are InfiniBand 4X. The Double Data Rate (DDR) QLE7240 and QLE7280 adapters have a raw data rate of 20Gbps (data rate of 16Gbps). For the Single Data Rate (SDR) adapters, the QLE7140 and QHT7140, the raw data rate is 10Gbps (data rate of 8Gbps). The QLE7240 and QLE7280 can also run in SDR mode.

The QLogic adapters utilize standard, off-the-shelf InfiniBand 4X switches and cabling. The QLogic interconnect is designed to work with all InfiniBand-compliant switches.

NOTE:

If you are using the QLE7240 or QLE7280, and want to use DDR mode, then DDR-capable switches must be used.

QLogic OFED OpenFabrics software is interoperable with other vendors' InfiniBand Host Channel Adapters (HCAs) running compatible OpenFabrics releases. There are several options for subnet management in your cluster:

- Use the embedded Subnet Manager (SM) in one or more managed switches supplied by your InfiniBand switch vendor.
- Use a host-based Subnet Manager. QLogic provides one, HSM, as a part of the InfiniBand fabric Suite download.
- Use the Open source Subnet Manager (OpenSM) component of OpenFabrics.

Interoperability

QLogic InfiniPath participates in the standard InfiniBand subnet management protocols for configuration and monitoring. Note that:

- InfiniPath OpenFabrics (including Internet Protocol over InfiniBand (IPoIB)) is interoperable with other vendors' InfiniBand HCAs running compatible OpenFabrics releases.
- The QLogic MPI stack is not interoperable with other InfiniBand HCAs and Target Channel Adapters (TCAs). Instead, it uses an InfiniBand-compliant, vendor-specific protocol that is highly optimized for QLogic MPI and for MPI over verbs.

NOTE:

See the OpenFabrics web site at www.openfabrics.org for more information on the OpenFabrics Alliance.

Conventions Used in this Guide

This guide uses the typographical conventions listed in [Table 1-1](#).

Table 1-1. Typographical Conventions

Convention	Meaning
command	Fixed-space font is used for literal items such as commands, functions, programs, files and pathnames, and program output.
variable	Italic fixed-space font is used for variable names in programs and command lines.

Table 1-1. Typographical Conventions (Continued)

Convention	Meaning
<i>concept</i>	Italic font is used for emphasis and concepts, as well as for documentation names/titles.
user input	Bold fixed-space font is used for literal items in commands or constructs that you type.
\$	Indicates a command line prompt.
#	Indicates a command line prompt as root.
[]	Brackets enclose optional elements of a command or program construct.
...	Ellipses indicate that a preceding element can be repeated.
>	A right caret identifies the cascading path of menu commands used in a procedure.
QLogic OFED 1.4	The current version number of the software is included within this documentation.
NOTE:	Indicates important information.

Documentation

The product documentation includes:

- The *QLogic HCA and QLogic OFED Software Install Guide*
- The *QLogic HCA and QLogic OFED Software Users Guide*
- The *QLogic InfiniBand Software Install Guide*
- The *QLogic ULP and Tools Reference Guide (OFED+ Users Guide)*
- Release Notes
- Quick Start Guide
- Readme file

For more information on system administration, using the QLogic Message-Passing Interface (MPI), and troubleshooting adapter hardware and software, see the *QLogic HCA and QLogic OFED Software Users Guide*.

Contact Information

Support Headquarters	QLogic Corporation 4601 Dean Lakes Blvd Shakopee, MN 55379 USA
QLogic Web Site	www.qlogic.com
Technical Support Web Site	support.qlogic.com
Technical Support Email	support@qlogic.com
Technical Training Email	tech.training@qlogic.com
Additional contact information is available from the Contact Support area of the Technical Support Web Site.	

Draft

Notes

Draft

2 Feature Overview

This section contains the features for this release, the supported QLogic adapter models, supported distributions and kernels, and a list of the software components.

Features

The QLogic OFED 1.4 software release contains the complete OFED 1.4, plus additional QLogic improvements, including an enhanced QLogic HCA driver. The InfiniPath 2.3 components (libraries, QLogic MPI/PSM, and utilities) are also included. QLogic also supplies MVAPICH and OpenMPI compiled with newer versions of each of four different compilers (GCC, PGI, Intel and PathScale).

The following features and enhancements are included in the QLogic OFED 1.4 release:

- Installation improvements. Provides a single software load for InfiniBand HCAs from QLogic and other vendors supported by OFED. the software is available packaged in the following ways:
 - Text User Interface (TUI) installer available (with the QLogicIB-Basic* download). TUI is used for install on smaller clusters. Software can be installed either standalone or via FastFabric (if the QLogic InfiniBand Fabric Suite is purchased).
 - Software packaged for use with `rpm` install method.
 - A subset of the software (the accelerated MPI stack, precompiled versions of MVAPICH and Open MPI, and other user-level tools) can be installed on top of stock OFED or on an IB-enabled distribution.
 - Software packaged for Rocks installation method.
 - Software packaged for Platform OCS installation method
- Write-combining (WC) mappings for the PIO buffers is now configured by default using the x86 Page Attribute Table (PAT) mechanism.

- MVAPICH and OpenMPI compiled with newer versions of each of four different compilers (GCC, PGI, Intel and PathScale) are available.
- The QLogic InfiniBand Fabric Suite (IFS) is available separately for purchase. It includes FastFabric, the QLogic Host Subnet Manager (HSM), and the Fabric Viewer, and the InfiniServ Host Software. The QLogic OFED 1.4 software is supported by IFS.
- Support for newer compiler versions (PathScale 3.x, PGI 7.x, PGI 8.x, Intel 10.x, Intel 11.x)
- Support for newer Linux distributions, including RHEL 4 U7
- Performance enhancements and bug fixes

Other Changes

- `ipath_ether` Ethernet emulation has been removed; IPoIB-CM may be used instead.
- The `/etc/init.d/infinipath` command to start the InfiniPath service has been replaced by the `/etc/init.d/openibd` command.
- The `infinipath-kernel` RPM no longer exists: it has been integrated into the `kernel-ib` RPM.

Continued Support

- Multiple high-performance native PSM Message Passing Interface (MPI) implementations. (PSM is QLogic's accelerated library for high performance MPIs). In addition to QLogic MPI, the currently supported MPI implementations are HP-MPI, Open MPI, MVAPICH, and Scali (Platform). Open MPI provides MPI-2 functionality, including one-sided operations and dynamic processes. These all offer the same high performance as QLogic MPI.
- Dual PCIe QLogic adapters per node.
- QLogic MPI supports running exclusively on a single node without the installation of the HCA hardware.
- 4K Maximum Transfer Unit (MTU) is supported and is on by default. To take advantage of 4KB MTU, use a switch that supports 4KB MTU. QLogic also supports 2KB switches, and 4KB MTU switches configured for 2KB MTU. QLogic switches with firmware version 4.2.x or later are recommended.

This version of the QLogic OFED software provides support for all of the QLogic HCAs in [Table 2-1](#).

Table 2-1. QLogic Adapter Model Numbers

QLogic Model Number	Description
QHT7040	Single port 10Gbps SDR 4X InfiniBand to HTX adapter. For systems with HTX expansion slots.
QHT7140 ^a	Single port 10Gbps SDR 4X InfiniBand to HTX adapter. For systems with HTX expansion slots.
QLE7140	Single port 10Gbps SDR 4X InfiniBand to PCI Express x8 adapter. Supported on systems with PCI Express (PCIe) x8 or x16 slots.
QLE7240	Single port 20Gbps DDR 4X InfiniBand to PCI Express x8 adapter. Supported on systems with PCI Express x8 or x16 slots.
QLE7280	Single port 20Gbps DDR 4X InfiniBand to PCI Express x16 adapter. Supported on systems with PCI Express x16 slots. The QLE7280 is backward compatible; it can also be used with PCIe adapters that connect to x8 slots.

Table Notes

PCIe is Gen 1

^a The QHT7140 has a smaller form factor than the QHT7040, but is otherwise the same. Throughout this document, the QHT7040 and QHT7140 will be collectively referred to as the *QHT7140* unless otherwise noted.

Additional up-to-date information can be found on the QLogic web site, specifically:

- The high performance computing page at www.qlogic.com/Products/HPC_products_landingpage.aspx
- The InfiniBand HCA page at www.qlogic.com/Products/HPC_products_infipathhcas.aspx

Software Components

This release includes all of OFED 1.4 with enhancements (QLogic OFED 1.4), including a new version of the VNIC tools and driver, and support for the QHT7xxx and QLE7xxx adapters. The software includes the QLogic InfiniPath HCA driver, libraries, QLogic MPI, Subnet Management Agent, and associated utilities. Included components are:

- InfiniPath driver
- InfiniPath libraries, InfiniPath utilities, configuration, and support tools, including `ipath_checkout`, `ipath_control`, `ipath_pkt_test`, and `ipathstats`
- QLogic MPI
- PSM support for accelerated MPI
- OpenMPI and MVAPICH (with PSM support) built with the GNU, PGI, PathScale, and Intel compilers, with corresponding `mpitests` and `mpi-selector`
- QLogic MPI benchmarks and utilities
- OpenFabrics protocols
- OpenFabrics libraries and utilities
- QLogic VNIC module
- FastFabric Enablement tools

This release provides support for the following protocols and transport services:

- IPoIB (TCP/IP networking in either Connected or Datagram mode)
- Sockets Direct Protocol (SDP)
- Open source Subnet Manager (OpenSM)
- Reliable Datagram Sockets (RDS)
- iSCSI Extensions for RDMA (iSER)

This release supports two versions of SCSI RDMA Protocol (SRP):

- OFED SRP
- QLogic SRP

No support is provided for Reliable Datagram (RD).

More details about the hardware and software can be found in [Section 4](#) and [Section 5](#).

3 Step-by-Step Installation Checklist

This section provides an overview of the hardware and software installation procedures. Detailed steps are found in [Section 4 “Hardware Installation”](#) and [Section 5 “Software Installation”](#).

Hardware Installation

The following steps summarize the basic hardware installation procedure:

1. Check that the adapter hardware is appropriate for your platform. See [Table 4-1](#).
2. Check to see that you have the appropriate cables and switches, as described in [“Cabling and Switches” on page 4-3](#).
3. Check to see that you are running a supported Linux distribution/kernel. See [Table 5-1](#).
4. Verify that the BIOS for your system is configured for use with the QLogic adapter. See [“Configuring the BIOS” on page 4-4](#).
5. Following the safety instructions in [“Safety with Electricity” on page 4-5](#). Unpack the adapter ([“Unpacking Information” on page 4-5](#)) and verify the package contents.
6. Install the adapter by following the instructions in [“Hardware Installation” on page 4-9](#).
7. Cable the adapter to the switch, as described in [“Cabling the Adapter to the InfiniBand Switch” on page 4-17](#). Check that all InfiniBand switches are configured.
8. Follow the steps in [“Completing the Installation” on page 4-18](#) to finish the installation.

Software Installation

The following steps summarize the basic QLogic OFED 1.4 software installation and startup. These steps must be performed on each node in the cluster:

1. Make sure that the HCA hardware installation has been completed according to the instructions in [“Hardware Installation” on page 4-1](#).
2. Verify that the Linux kernel software is installed on each node in the cluster. The required kernels and supported Linux distributions for both QLogic InfiniPath and OpenFabrics are defined in [Table 5-1](#).
3. Make sure that your environment has been set up as described in [“Setting Up Your Environment” on page 5-3](#).
4. Download your version of the QLogic InfiniPath/OpenFabrics software from the QLogic web site to a local server directory. See [“Choose the Appropriate Download Files” on page 5-4](#).
5. Install the selected packages on each cluster node using the corresponding method as described in one of the following: [“Install QLogicIB-Basic with the Installer Tool” on page 5-7](#), [“Using rpm to Install InfiniPath and OpenFabrics” on page 5-14](#), [“Install QLogic OFED User-level Software with the rpm Command” on page 5-17](#), [“Install QLogic OFED Using Rocks” on page 5-20](#), [“Install QLogic OFED Using a Platform OCS Kit” on page 5-22](#), [“Install FastFabric Software CD/ISO Image” on page 5-22](#).
6. The system can be rebooted after all the software has been installed.
7. The configuration file for the `ib_ipath` driver is set up correctly at installation and is loaded automatically during system boot once the RPMs have been installed. However, if you wish to change the configuration file, see [“Managing the InfiniPath Driver” on page 6-15](#).
8. If you want to configure the optional OpenFabrics driver `ipoib`, and you have not configured it yet with the Install tool, or if you have used the `rpm` install method, configure it as described in [“Configuring the IPoIB Network Interface” on page 6-2](#).
9. If you want to use the optional OpenFabrics services (`opensm`, `srp`, or VNIC), configure them as described in [“OpenSM” on page 6-3](#), [“SRP” on page 6-4](#), or [“SRP stands for SCSI RDMA Protocol. It was originally intended to allow the SCSI protocol to run over InfiniBand for Storage Area Network \(SAN\) usage. SRP interfaces directly to the Linux file system through the SRP Upper Layer Protocol. SRP storage can be treated as another device.” on page 6-4](#).
10. Check the system state by observing the LEDs. See [“LED Link and Data Indicators” on page 7-1](#).

11. You can optimize your system and adapter for the best performance. See [“Adapter and Other Settings” on page 7-1](#). Also see the Performance Settings and Management Tips section in the *QLogic HCA and QLogic OFED Software Users Guide*.
12. Perform the recommended health checks. See [“Customer Acceptance Utility” on page 7-2](#).
13. After installing the QLogic InfiniPath and OpenFabrics software, refer to the *QLogic HCA and QLogic OFED Software Users Guide* for more information about using QLogic InfiniPath, QLogic MPI, and OpenFabrics products. Refer to the *QLogic ULP and Tools Reference Guide* for more information about configuring and using QLogic SRP, QLogic VNIC, and the Enablement Tools. The *InfiniBand Software Installation Guide* also has information on installing the QLogic InfiniBand Fabric Suite.

Draft

Notes

Draft

4 Hardware Installation

This section lists the requirements and provides instructions for installing the QLogic InfiniPath Interconnect adapters. Instructions are included for the QLogic DDR PCI Express adapters, the QLE7240 and QLE7280; the QLogic PCIe adapter and PCIe riser card, QLE7140; and the QHT7040 or QHT7140 adapter hardware and HTX riser card. These components are collectively referred to as the *adapter* and the *riser card* in the remainder of this document.

The adapter is a low-latency, high-bandwidth, high message rate cluster interconnect for InfiniBand. The QLogic interconnect is InfiniBand 4X, with a raw data rate of 20Gbps (data rate of 16Gbps) for the QLE7240 and QLE7280; and 10Gbps (data rate of 8Gbps) for the QLE7140, QHT7040, and QHT7140.

OpenFabrics is interoperable with other vendors' InfiniBand Host Channel Adapters (HCAs) running compatible OpenFabrics releases.

Hardware Installation Requirements

This section lists hardware and software environment requirements for installing the QLogic QLE7240, QLE7280, QLE7140, QHT7040, or QHT7140.

Hardware

QLogic interconnect adapters are for use with UL listed computers. The following statement is true for all the adapters:

This device complies with part 15 of the FCC Rules. Operation 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 undesired operations.

Different adapter cards work on different platforms. [Table 4-1](#) shows the relationship between the adapter model and different types of motherboards.

Table 4-1. Adapter Models and Related Platforms

QLogic Model Number	Platform	Plugs Into
QLE7240	PCI Express systems	Standard PCI Express x8 or x16 slot
QLE7280	PCI Express systems	Standard PCI Express x16 slot

Table 4-1. Adapter Models and Related Platforms (Continued)

QLogic Model Number	Platform	Plugs Into
QLE7140	PCI Express systems	Standard PCI Express x8 or x16 slot
QHT7040	Motherboards with HTX connectors	HyperTransport HTX slot
QHT7140	Motherboards with HTX connectors	HyperTransport HTX slot

Installation of the QLE7240, QLE7280, QLE7140, QHT7040, or QHT7140 in a 1U or 2U chassis requires the use of a riser card. See [Figure 4-4](#) for an illustration of a PCI Express (PCIe) slot in a typical motherboard. See [Figure 4-7](#) for an illustration of an HTX slot for a typical Opteron motherboard.

The motherboard vendor is the optimal source for information on the layout and use of HyperTransport and PCI Express-enabled expansion slots on supported motherboards.

Form Factors

The **QLE7240**, **QLE7280**, and **QLE7140** are the model numbers for the adapters that ship in the standard PCI Express half-height, short-form factor. These adapters can be used with either full-height or low-profile face plates.

The **QHT7040** is the model number for the adapter that shipped in the HTX full-height factor. The HTX low-profile form factor is referred to as the **QHT7140**. It is the same as the QHT7040, except for its more compact size. In either case, the adapter is backward and forward compatible for the motherboards in which it is supported. The QHT7040 and QHT7140 HTX adapters are collectively referred to as the **QHT7140** unless otherwise stated.

When the QHT7040 or QHT7140 adapter is installed with the riser card, it may prevent some or all of the other PCI expansion slots from being used, depending on the form factor of the adapter and motherboard.

Run `ipath_control -i` to see information on which form adapter is installed. The file `/sys/class/infiniband/ipath0/device/boardversion` contains the same information. For more information, see the Useful Programs and Files appendix in the *QLogic HCA and QLogic OFED Software Users Guide*.

Cabling and Switches

The cable installation uses a standard InfiniBand (IB) 4X cable. Any InfiniBand cable that has been qualified by the vendor should work. For SDR, the longest passive copper IB cable that QLogic has currently qualified is 20 meters. For DDR-capable adapters and switches, the DDR-capable passive copper cables cannot be longer than 10 meters. Active cables can eliminate some of the cable length restrictions.

InfiniBand switches are available through QLogic.

NOTE:

If you are using the QLE7240 or QLE7280 and want to use DDR mode, then DDR-capable switches must be used.

The copper cables listed in [Table 4-2](#) are available from QLogic:

Table 4-2. QLogic InfiniBand Cables

Product Number	Description
7104-1M-Cable	4x-4x cable—1 meter
7104-2M-Cable	4x-4x cable—2 meters
7104-3M-Cable	4x-4x cable—3 meters
7104-4M-Cable	4x-4x cable—4 meters
7104-5M-Cable	4x-4x cable—5 meters
7104-6M-Cable	4x-4x cable—6 meters
7104-7M-Cable	4x-4x cable—7 meters
7104-8M-Cable	4x-4x cable—8 meters
7104-9M-Cable	4x-4x Cable—9 meters
7104-10M-Cable	4x-4x cable—10 meters
7104-12M-Cable	4x-4x cable—12 meters (SDR only)
7104-14M-Cable	4x-4x cable—14 meters (SDR only)
7104-16M-Cable	4x-4x cable—16 meters (SDR only)
7104-18M-Cable	4x-4x cable—18 meters (SDR only)

For cabling instructions, see [“Cabling the Adapter to the InfiniBand Switch” on page 4-17.](#)

Optical Fibre Option

The QLogic adapter also supports connection to the switch by means of optical fibres through optical media converters such as the EMCORE™ QT2400. Not all switches support these types of convertors. For more information on the EMCORE convertor, see www.emcore.com.

Intel® and Zarlink™ also offer optical cable solutions. See www.intel.com and www.zarlink.com for more information.

Configuring the BIOS

To achieve the best performance with QLogic adapters, you need to configure your BIOS with specific settings. The BIOS settings, which are stored in non-volatile memory, contain certain parameters characterizing the system. These parameters may include date and time, configuration settings, and information about the installed hardware.

This setting is required:

- Advanced Configuration and Power Interface (ACPI) BIOS option must be enabled.

For more information, see “[Enable Advanced Configuration and Power Interface \(ACPI\)](#)” on page A-2 and the Troubleshooting section of the *QLogic HCA and QLogic OFED Software Users Guide*.

Some other BIOS settings can be adjusted for better adapter performance. See “[Adapter and Other Settings](#)” on page 7-1.

For specific instructions about BIOS settings, follow the hardware documentation that came with your system.

NOTE:

The x86 Page Attribute Table (PAT) mechanism that allocates write-combining (WC) mappings for the PIO buffers has been added and is now the default. This was previously a BIOS setting. For more information, see “[Write Combining](#)” on page B-1.

Safety with Electricity

Observe these guidelines and safety precautions when working around computer hardware and electrical equipment:

- Locate the power source shutoff for the computer room or lab where you are working. This is where you will turn OFF the power in the event of an emergency or accident. Never assume that power has been disconnected for a circuit; always check first.
- Do not wear loose clothing. Fasten your tie or scarf, remove jewelry, and roll up your sleeves. Wear safety glasses when working under any conditions that might be hazardous to your eyes.
- Shut down and disconnect the system's power supply from AC service before you begin work, to insure that standby power is not active. Power off all attached devices such as monitors, printers, and external components. Note that many motherboards and power supplies maintain standby power at all times. Inserting or removing components while standby is active can damage them.
- Use normal precautions to prevent electrostatic discharge, which can damage integrated circuits.

Unpacking Information

This section provides instructions for safely unpacking and handling the QLogic adapter. To avoid damaging the adapter, always take normal precautions to avoid electrostatic discharge.

Verify the Package Contents

The QLogic adapter system should arrive in good condition. Before unpacking, check for any obvious damage to the packaging. If you find any obvious damage to the packaging or to the contents, please notify your reseller immediately.

List of the Package Contents

The package contents for the QLE7240 adapter are:

- QLogic QLE7240
- Additional short bracket
- Quick Start Guide

Standard PCIe risers can be used, typically supplied by your system or motherboard vendor.

The package contents for the QLE7280 adapter are:

- QLogic QLE7280
- Additional short bracket
- Quick Start Guide

Standard PCIe risers can be used, typically supplied by your system or motherboard vendor.

The package contents for the QLE7140 adapter are:

- QLogic QLE7140
- Quick Start Guide

Standard PCIe risers can be used, typically supplied by your system or motherboard vendor. The contents are illustrated in [Figure 4-2](#).

The package contents for the QHT7140 adapter are:

- QLogic QHT7140
- HTX riser card for use in 1U or 2U chassis
- Quick Start Guide

The contents are illustrated in [Figure 4-3](#).

The IBA6120, IBA6110, and IBA7220 are the QLogic ASICs, which are the central components of the interconnect. The location of the IBA7220 ASIC on the adapter is shown in [Figure 4-1](#). The location of the IBA6120 ASIC on the adapter is shown in [Figure 4-2](#). The location of the IBA6110 ASIC on the adapter is shown in [Figure 4-3](#).

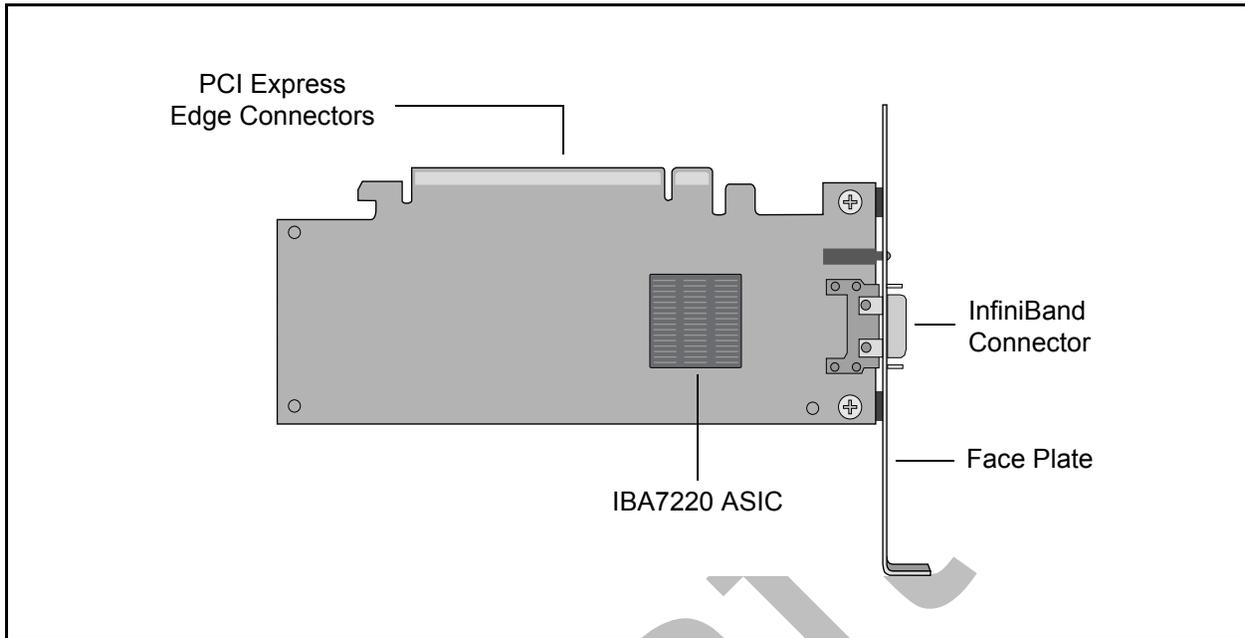


Figure 4-1. QLogic QLE7280 with IBA7220 ASIC

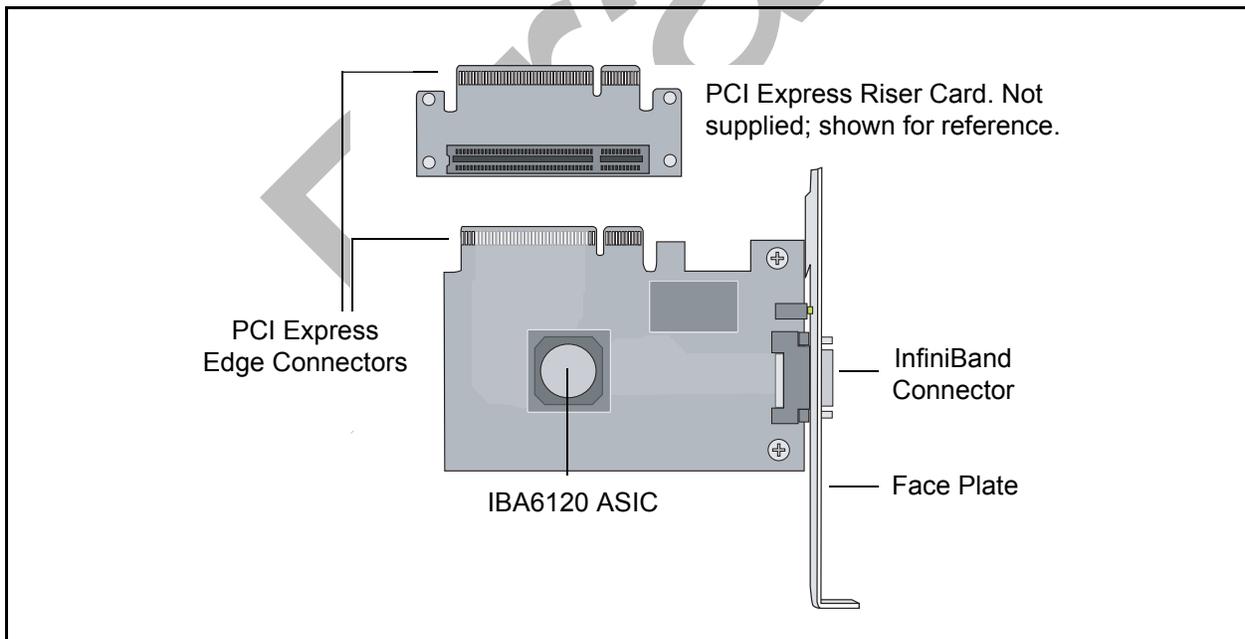


Figure 4-2. QLogic QLE7140 Card with Riser, Top View

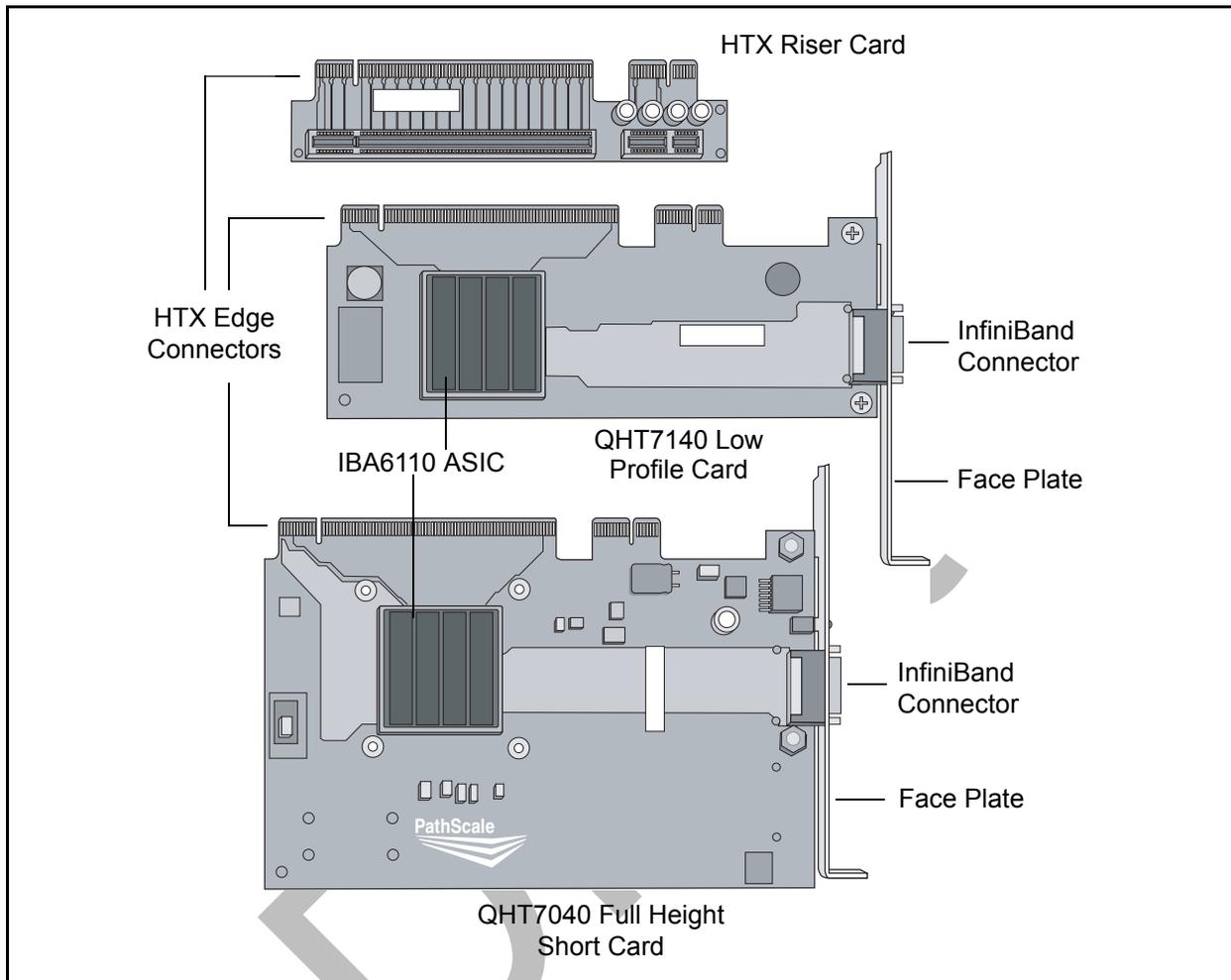


Figure 4-3. QLogic QHT7040/QHT7140 Full and Low Profile Cards with Riser, Top View

Unpacking the QLogic Adapter

Follow these steps when unpacking the QLogic adapter:

1. When unpacking, ground yourself before removing the QLogic adapter from the anti-static bag.
2. Grasping the QLogic adapter by its face plate, pull the adapter out of the anti-static bag. Handle the adapter only by its edges or the face plate. Do not allow the adapter or any of its components to touch any metal parts.
3. After checking for visual damage, store the adapter and the riser card in their anti-static bags until you are ready to install them.

Hardware Installation

This section contains hardware installation instructions for the QLE7240, QLE7280, QLE7140, QHT7040, and QHT7140.

Hardware Installation for QLE7240, QLE7280, or QLE7140 with PCI Express Riser

Installation for the QLE7240, QLE7280, and QLE7140 is similar. The following instructions are for the QLE7140, but can be used for any of these three adapters.

Most installations will be in 1U and 2U chassis, using a PCIe right angle riser card. This results in an installation of the adapter that is parallel to the motherboard. This type of installation is described first. Installation in a 3U chassis is described in [“Hardware Installation for the QHT7140 Without an HTX Riser” on page 4-16](#).

Installing the QLogic QLE7140 in a 1U or 2U chassis requires a PCIe right angle riser card.

A taller riser card can be used if necessary. The QLE7140 can connect to any of the standard compatible PCI Express riser cards.

Dual Adapter Installation

If you have a motherboard with dual PCIe slots, dual adapters can be installed. The adapters must match. For example, on a motherboard with two x16 slots, dual QLE7280 adapters can be installed, but not a QLE7240 adapter and a QLE7280 adapter. Check the design of your motherboard to see how riser cards can be used.

Follow the instructions in [“Installation Steps” on page 4-9](#).

See the Using MPI section in the *QLogic HCA and QLogic OFED Software Users Guide* for information on using the `IPATH_UNIT` environment variable to control which HCA to use.

Installation Steps

To install the QLogic adapter with a PCIe riser card:

1. The BIOS should already be configured properly by the motherboard manufacturer. However, if any additional BIOS configuration is required, it will usually need to be done before installing the QLogic adapter. See [“Configuring the BIOS” on page 4-4](#).
2. Shut down the power supply to the system into which you will install the QLogic adapter.
3. Take precautions to avoid electrostatic damage (ESD) to the cards by properly grounding yourself or touching the metal chassis to discharge static electricity before handling the cards.

4. Remove the cover screws and cover plate to expose the system's motherboard. For specific instructions on how to do this, follow the hardware documentation that came with your system.
5. Locate the PCIe slot on your motherboard. Note that the PCIe slot has two separate sections, with the smaller slot opening located towards the front (see [Figure 4-4](#)). These two sections correspond to the shorter and longer connector edges of the adapter and riser.

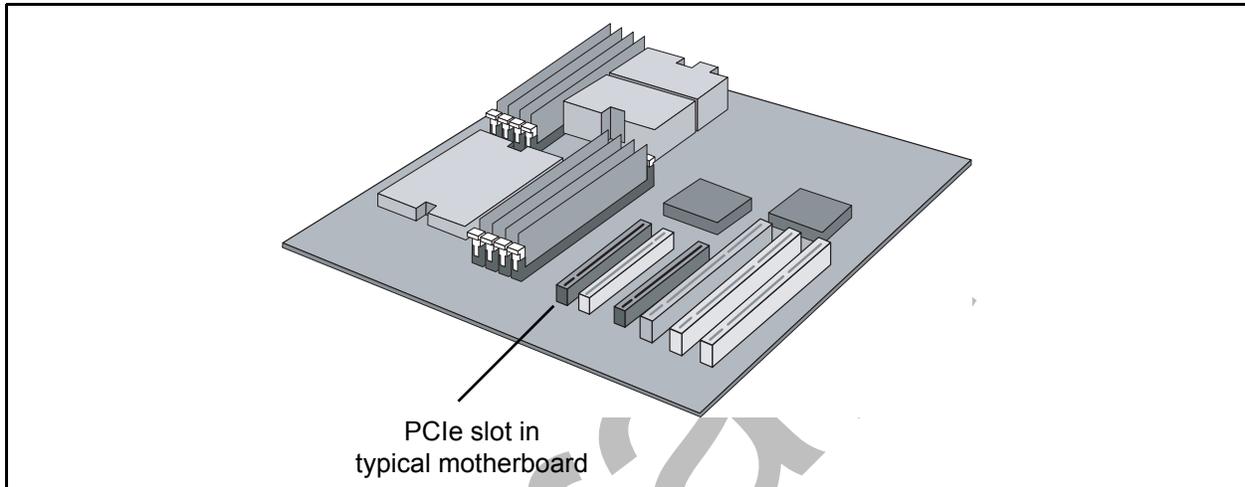


Figure 4-4. PCIe Slot in a Typical Motherboard

6. Determine if a blanking panel is installed in your chassis. If it is, remove it so that the InfiniBand connector will be accessible. Refer to your system vendor instructions for how to remove the blanking panel.
7. Remove the QLogic adapter from the anti-static bag.
8. Locate the face plate on the connector edge of the card.

9. Connect the QLogic adapter and PCIe riser card together, forming the assembly that you will insert into your motherboard. First, visually line up the adapter slot connector edge with the edge connector of the PCIe riser card (see [Figure 4-5](#)).

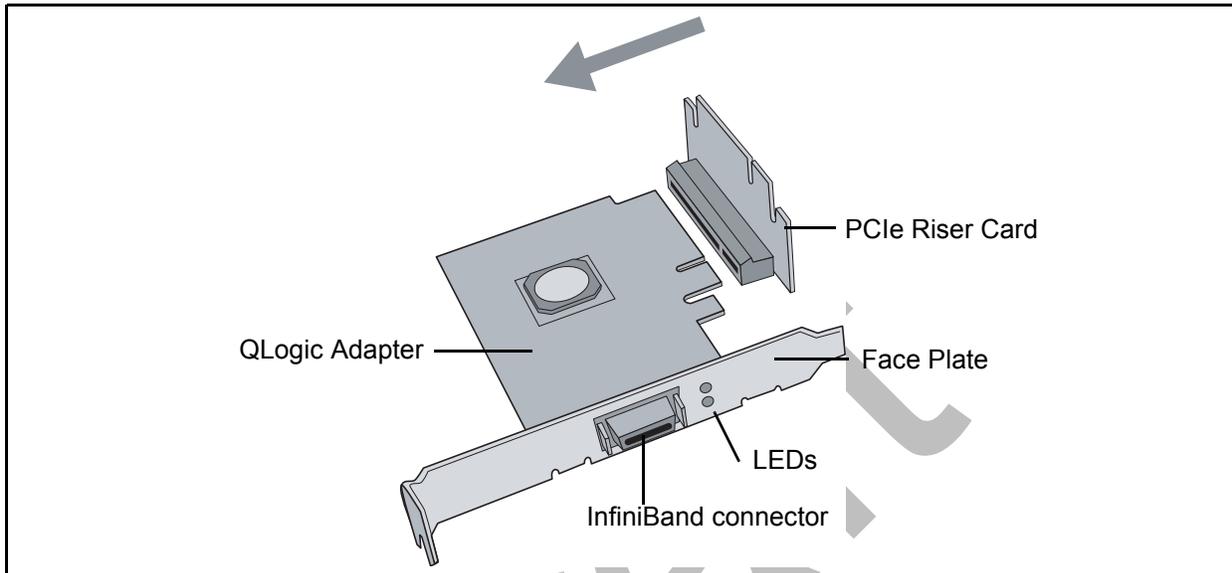


Figure 4-5. QLogic PCIe HCA Assembly with Riser Card

10. Holding the QLogic adapter by its edges, carefully insert the card slot connector into the PCIe riser card edge connector, as show in [Figure 4-5](#). The result is a combined L-shaped assembly of the PCIe riser card and QLogic adapter. This assembly is what you will insert into the PCIe slot on the motherboard in the next step.
11. Turn the assembly so that the riser card connector edge is facing the PCIe slot on the motherboard, and the face plate is toward the front of the chassis.
12. Holding this assembly above the motherboard at about a 45 degree angle, slowly lower it so that the connector on the face plate clears the blanking panel opening of the chassis from the inside. Slowly align the connector edge of the riser card with the motherboard's PCIe slot. The short section of the connector must align with the short section of the slot.

13. Insert the riser assembly into the motherboard's PCIe slot, ensuring good contact. The QLogic adapter should now be parallel to the motherboard and about one inch above it (see [Figure 4-6](#)).

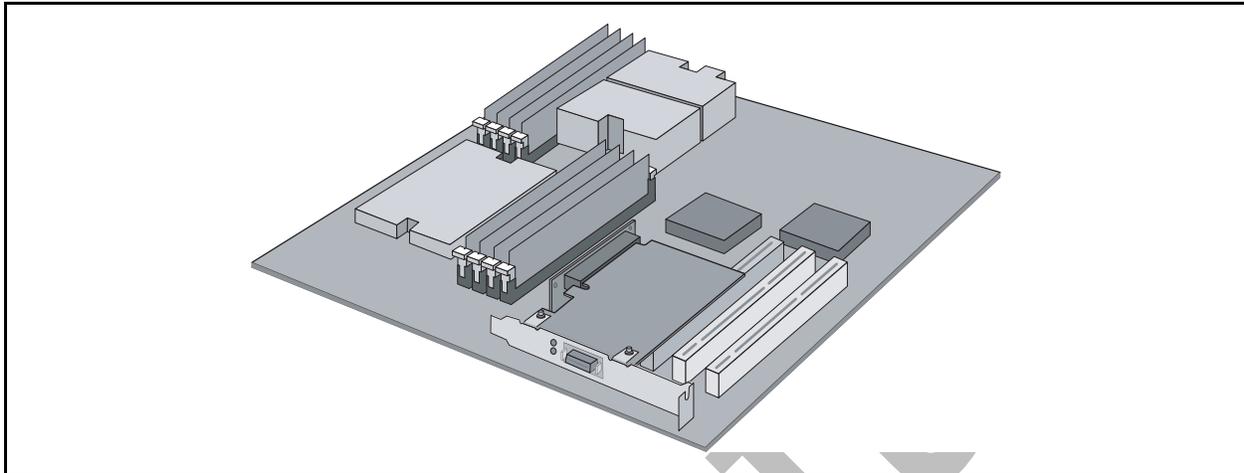


Figure 4-6. Assembled PCIe HCA with Riser

14. Secure the face plate to the chassis. The QLogic adapter has a screw hole on the side of the face plate that can be attached to the chassis with a retention screw. The securing method may vary depending on the chassis manufacturer. Refer to the system documentation for information about mounting details such as mounting holes, screws to secure the card, or other brackets.

The QLogic PCIe HCA with PCIe riser card is now installed. Next, install the cables as described in [“Cabling the Adapter to the InfiniBand Switch” on page 4-17](#). Then test your installation by powering up and verifying link status (see [“Completing the Installation” on page 4-18](#)).

Hardware Installation for QHT7140 with HTX Riser

Most installations will be in 1U and 2U chassis, using the HTX riser card. This results in a horizontal installation of the QHT7140. This type of installation is described in this section. Installation in a 3U chassis is described in [“Hardware Installation for the QHT7140 Without an HTX Riser” on page 4-16](#).

Installation of QLogic QHT7140 in a 1U or 2U chassis requires an HTX riser card.

NOTE:

The illustrations in this section are shown for the full height short form factor. Installation of the HTX low profile form factor follows the same steps.

To install the QLogic adapter with an HTX riser card:

1. The BIOS should be already be configured properly by the motherboard manufacturer. However, if any additional BIOS configuration is required, it will usually need to be done before installing the QLogic adapter. See [“Configuring the BIOS” on page 4-4](#).
2. Shut down the power supply to the system into which you will install the QLogic adapter.
3. Take precautions to avoid electrostatic discharge (ESD) damage to the cards by properly grounding yourself or touching the metal chassis to discharge static electricity before handling the cards.
4. Remove the cover screws and cover plate to expose the system’s motherboard. For specific instructions on how to do this, follow the hardware documentation that came with your system.
5. Locate the HTX slot on your motherboard. Note that the HTX slot has two separate connectors, corresponding to the connector edges of the adapter. See [Figure 4-7](#).

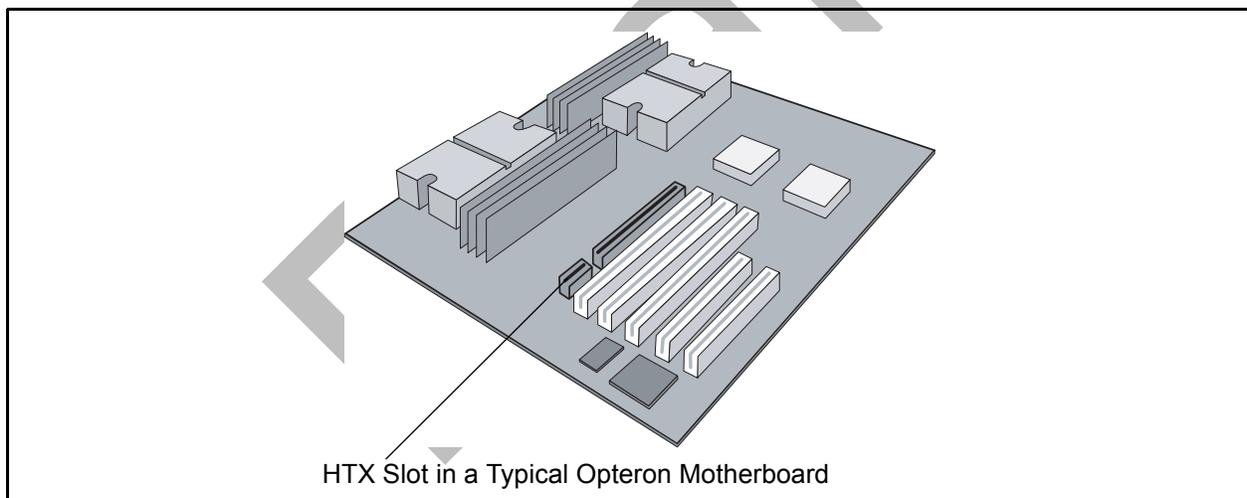


Figure 4-7. HTX Slot

6. Determine if a blanking panel is installed in your chassis. If it is, remove it so that the InfiniBand connector will be accessible. Refer to your system vendor instructions for how to remove the blanking panel.

7. Remove the QLogic QHT7140 from the anti-static bag.

NOTE:

Be careful not to touch any of the components on the printed circuit board during these steps. You can hold the adapter by its face plate or edges.

8. Locate the face plate on the connector edge of the card.
9. Connect the QLogic adapter and HTX riser card together, forming the assembly that you will insert into your motherboard. First, visually line up the adapter slot connector edge with the edge connector of the HTX riser card (see [Figure 4-8](#)).

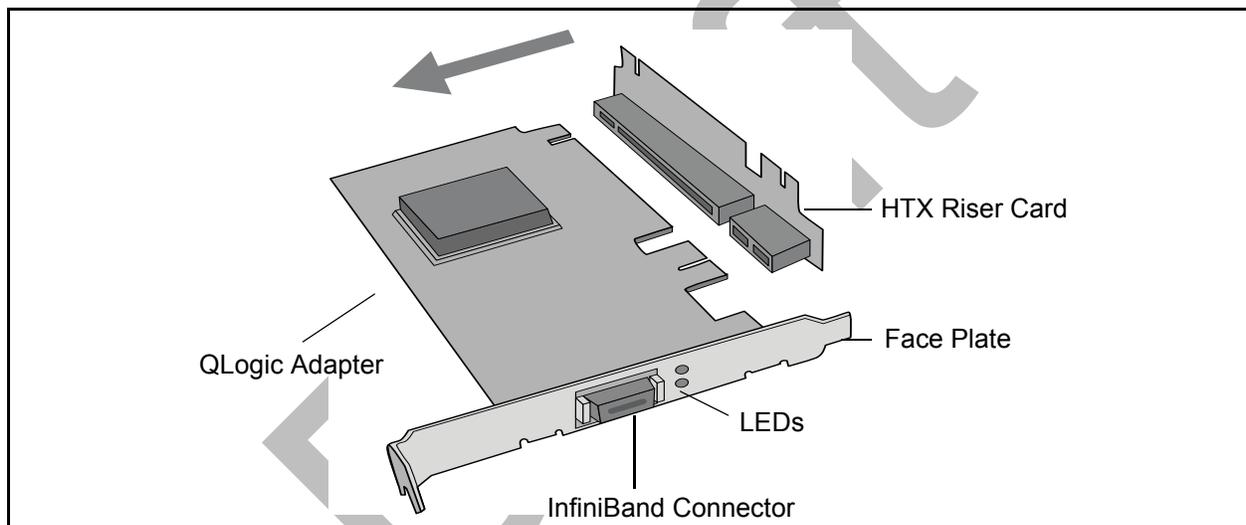


Figure 4-8. QLogic QHT7140 Adapter with Riser Card

10. Holding the QLogic adapter by its edges, carefully insert the card slot connector into the HTX riser card edge connector, as show in [Figure 4-8](#). The result is a combined L-shaped assembly of the HTX riser card and QLogic adapter. This assembly is what you will insert into the HTX slot on the motherboard in the next step.
11. Turn the assembly so that the riser card connector edge is facing the HTX slot on the motherboard, and the face plate is toward the front of the chassis.
12. Holding this assembly above the motherboard at about a 45 degree angle, slowly lower it so that the connector on the face plate clears the blanking panel opening of the chassis from the inside. Slowly align the connector edge of the HTX riser card with the motherboard's HTX slot. The HTX riser and HTX slot must line up perfectly.

13. Insert the HT riser assembly into the motherboard's HTX slot, ensuring good contact. The QLogic adapter should now be parallel to the motherboard and about one inch above it, as shown in [Figure 4-9](#).

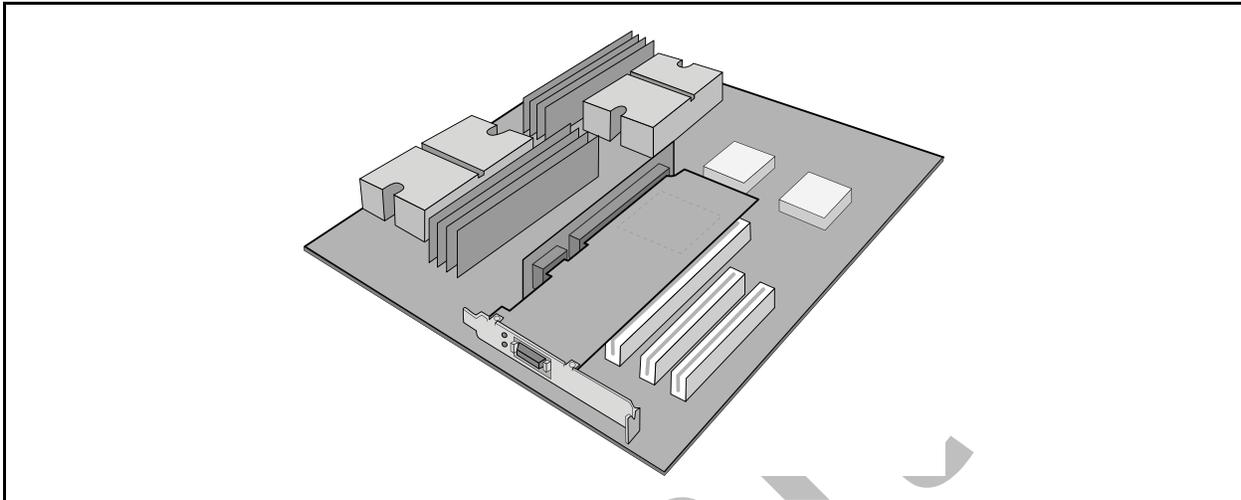


Figure 4-9. Assembled QHT7140 with Riser

14. Secure the face plate to the chassis. The QLogic adapter has a screw hole on the side of the face plate that can be attached to the chassis with a retention screw. The securing method may vary depending on the chassis manufacturer. Refer to the system documentation for information about mounting details such as mounting holes, screws to secure the card, or other brackets.

The QLogic QHT7140 with HTX riser card is now installed. Next, install the cables as described in [“Cabling the Adapter to the InfiniBand Switch”](#) on page 4-17. Then test your installation by powering up and verifying link status (see [“Completing the Installation”](#) on page 4-18).

Hardware Installation for QLE7240, QLE7280, and QLE7140 Without a PCI Express Riser

Installing the QLogic QLE7240, QLE7280, or QLE7140 without a PCI Express riser card requires a 3U or larger chassis.

Installation is similar to the QHT7140 HTX adapter, except that the card slot connectors on these adapters fit into the PCIe slot rather than the HTX slot. Follow the instructions in [“Hardware Installation for the QHT7140 Without an HTX Riser”](#) on page 4-16, substituting the PCIe slot for the HTX slot.

Hardware Installation for the QHT7140 Without an HTX Riser

Installing the QLogic QHT7140 without an HTX riser card requires a 3U or larger chassis. The card slot connectors on the QHT7140 fit into the HTX slot in a vertical installation.

To install the QLogic adapter without the HTX riser card:

1. The BIOS should already be configured properly by the motherboard manufacturer. However, if any additional BIOS configuration is required, it will usually need to be done before installing the QLogic adapter. See [“Configuring the BIOS” on page 4-4](#).
2. Shut down the power supply to the system into which you will install the QLogic adapter.
3. Take precautions to avoid electrostatic discharge (ESD) damage to the cards by properly grounding yourself or touching the metal chassis to discharge static electricity before handling the cards.
4. If you are installing the QLogic adapter into a covered system, remove the cover screws and cover plate to expose the system’s motherboard. For specific instructions on how to do this, follow the hardware documentation that came with your system.
5. Locate the HTX slot on your motherboard (see [Figure 4-7](#)).
6. Remove the QLogic adapter from the anti-static bag. Hold the card by the top horizontal section of the bracket, and the top rear corner of the card. Be careful not to touch any of the components on the printed circuit card.
7. Without fully inserting, gently align and rest the HTX card’s gold fingers on top of the motherboard’s HTX slot.

8. Insert the card by pressing firmly and evenly on the top of the horizontal bracket and the top rear corner of the card simultaneously. The card should insert evenly into the slot. Be careful not to push, grab, or put pressure on any other part of the card, and avoid touching any of the components. See [Figure 4-10](#).

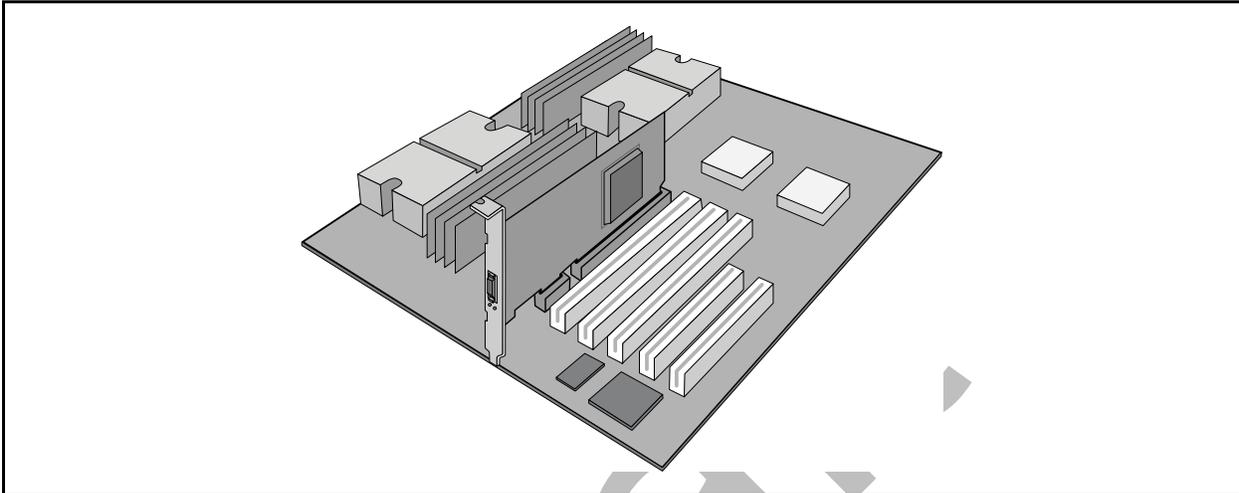


Figure 4-10. QHT7140 Without Riser Installed in a 3U Chassis

9. Secure the face plate to the chassis. The QLogic adapter has a screw hole on the side of the face plate that can be attached to the chassis with a retention screw. The securing method may vary depending on the chassis manufacturer. Refer to the system documentation for information about mounting details such as mounting holes, and screws to secure the card, or other brackets.

Next, install the cables, as described in [“Cabling the Adapter to the InfiniBand Switch” on page 4-17](#). Then test your installation by powering up the system (see [“Completing the Installation” on page 4-18](#)).

Switch Configuration and Monitoring

The QLogic interconnect is designed to work with all InfiniBand-compliant switches. Follow the vendor documentation for installing and configuring your switches.

Cabling the Adapter to the InfiniBand Switch

Follow the recommendations of your cable vendor for cable management and proper bend radius.

The QLE7240, QLE7280, QLE7140, QHT7040, and QHT7140 adapters are all cabled the same way.

To install the InfiniBand cables:

1. Check that you have removed the protector plugs from the cable connector ends.
2. Different vendor cables might have different latch mechanisms. Determine if your cable has a spring-loaded latch mechanism.
 - If your cable is spring-loaded, grasp the metal shell and pull on the plastic latch to release the cable. To insert, push and the cable snaps into place. You will hear a short “click” sound from the cable connector when it snaps in.
 - If your cable latch mechanism is not spring-loaded, push on the metal case, then push the plastic latch to lock the cable in place.
3. The InfiniBand cables are symmetric; either end can be plugged into the switch. Connect the InfiniBand cable to the connector on the QLogic QLE7240, QLE7280, QLE7140 or QHT7140. Depress the side latches of the cable when connecting. (On some cables, this latch is located at the top of the cable connector.) Make sure the lanyard handle on the cable connector is slid forward toward the card connector until fully engaged.
4. Connect the other end of the cable to the InfiniBand switch.

Completing the Installation

To complete the hardware installation:

1. Complete any other installation steps for other components.
2. Replace the cover plate and back panel.
3. Verify that the power cable is properly connected.
4. Turn on the power supply and boot the system normally.
5. Watch the LED indicators. The LEDs will flash only once, briefly, at power-up. The LEDs are functional only after the InfiniPath software has been installed, the driver has been loaded, and the system is connected to an InfiniBand switch. To use the LEDs to check the state of the adapter, see [“LED Link and Data Indicators” on page 7-1.](#)

5 Software Installation

This section provides instructions for installing QLogic OFED 1.4, which includes QLogic InfiniPath and the OpenFabrics software. The software includes drivers, protocol libraries, QLogic's implementation of the MPI message passing standard, associated utilities, and example programs, including benchmarks. A complete list of the provided software is in [“Software Components” on page 2-4](#).

Cluster Setup

Information on clusters, supported distributions and kernels, and environment setup is provided in [“Types of Nodes in a Cluster Environment” on page 5-1](#), [“Supported Linux Distributions” on page 5-1](#), and [“Distribution Identifiers” on page 5-2](#).

Types of Nodes in a Cluster Environment

In a cluster environment, different nodes can be used for different functions, such as launching jobs, developing software, or running jobs. The nodes are defined as follows:

- **Front end node.** This node launches jobs.
- **Compute node.** This node runs jobs.
- **Development or build node.** These are the machines on which examples or benchmarks can be compiled.

Any machine can serve any combination of these three purposes, but a typical cluster has many compute nodes and just a few (or only one) front end nodes. The number of nodes used for development will vary. These node names are used throughout this guide.

Supported Linux Distributions

The QLogic interconnect runs on AMD™ Opteron™ and 64-bit Intel Xeon systems running Linux®. The currently supported distributions and associated Linux kernel versions for InfiniPath and OpenFabrics are listed in [Table 5-1](#).

The kernels are the ones that shipped with the distributions. All are for the x86_64 architecture.

Table 5-1. InfiniPath/OpenFabrics Supported Distributions and Kernels

Distribution	InfiniPath/OpenFabrics Supported Kernels
Red Hat® Enterprise Linux® (RHEL) 4.5	2.6.9-55 (U5),
RHEL 4.6	2.6.9-67 (U6)
RHEL 4.7	2.6.9-78 (U7)
CentOS 4.5	2.6.9-55
CentOS 4.6	2.6.9-67
CentOS 4.7	2.6.9-78
Scientific Linux 4.5	2.6.9-55
Scientific Linux 4.6	2.6.9-67
Scientific Linux 4.7	2.6.9-78
Red Hat Enterprise Linux 5.1 (RHEL 5.1)	2.6.18-53, 2.6.18-92
RHEL 5.2	2.6.18-92
CentOS 5.1	2.6.18-53, 2.6.18-92
CentOS 5.2	2.6.18-92
Scientific Linux 5.1	2.6.18-53, 2.6.18-92
Scientific Linux 5.2	2.6.18-92
SUSE® Linux Enterprise Server 10 SP 1	2.6.16.46
SUSE® Linux Enterprise Server 10 SP 2	2.6.16.60

NOTE:

Support for RHEL4 U4 and SLES 10.0 has been removed.

Distribution Identifiers

Distribution identifiers for this release are listed in the table below. They are used in file naming conventions.

Table 5-2. Distribution Identifiers

Distribution Identifier	Used On
rhel4	Red Hat Enterprise Linux 4.5 (RHEL4.5), RHEL4.6, RHEL 4.7, CentOS 4.5-4.7, Scientific Linux 4.5-4.7 for x86_64 systems

Table 5-2. Distribution Identifiers (Continued)

Distribution Identifier	Used On
rhel5	Red Hat Enterprise Linux 5.1 (RHEL5.1), RHEL5.2, CentOS 5.1-5.2, Scientific Linux 5.1-5.2, for x86_64 systems
sles10	SLES 10 SP1-SP2 for x86_64 systems

Compiler Support

QLogic MPI supports use of a number of compilers. These include:

- GNU gcc 3.3.x, 3.4.x, 4.0, 4.1, 4.2.x, and 4.3.x compiler suites
- PathScale Compiler Suite 3.0, 3.1 and 3.2
- PGI 5.2, 6.0, 7.1, 7.2-4, and 8.0-3
- Intel 9.x, 10.1, and 11.0
- gfortran 4.1.x

The PathScale Compiler Suite Version 3.x is now supported on systems that have the GNU 4.0 and 4.1 compilers and compiler environment (header files and libraries).

Setting Up Your Environment

Keep the following in mind when setting up the environment:

- The runtime and build environments must be the same. Compatibility between executables built on different Linux distributions cannot be guaranteed.
- You will need Administrator privileges on your machine(s).
- If using the `rpm` install method, make sure that all previously existing (stock) OpenFabrics RPMs are uninstalled. See [“Uninstalling InfiniPath and OpenFabrics RPMs” on page 5-24](#) for more information on uninstalling. The Installer tool will uninstall previous RPMs automatically before upgrades.
- It is possible to have a cluster running with different kernel versions. However, QLogic recommends and supports clusters where all nodes run equivalent software.

Some operating system packages are required for OpenFabrics; they are listed in [Table 5-3](#).

Table 5-3. Required OS Packages

OS Distribution	Required Packages
All	gcc, glib, glibc
All (For development)	glib-devel, glibc-devel, glibc-devel-32bit (to build 32-bit libraries on x86_86 and ppc64), zlib-devel
Red Hat and Red Hat-derived kernels	kernel-devel, rpm-build
SLES 10	kernel-source, rpm

There are also OS package requirements for some specific components in QLogic OFED.

Table 5-4. Specific Component Requirements

QLogic OFED Component	Required OS Packages
QLogic infinipath* software (Listed as TrueScale HCA Libs in the Installer)	openssh and openssh-server. Note that in the SLES 10 distribution, openssh-server is a part of the openssh package. python, if the Multi-Purpose Daemon (MPD) job launcher or the ipath_mtrr script is to be used. These packages must be on every node.
MVAPICH	a Fortran Compiler
MVAPICH2	libstdc++-devel, sysfsutils (SLES), lib-sysfs-devel (RedHat 5.0)
Open MPI	libstdc++-devel
ibutils	tcl-8.4, tcl-devel-8.4, tk, lib-stdc++-devel
QLogic openmpi_gcc*	libgfortran (on RHEL4)

Choose the Appropriate Download Files

This section assumes that the correct Linux kernel, a supported distribution, and the required prerequisites (see [Table 5-3](#) and [Table 5-4](#)) have been installed on every node.

The components of the QLogic OFED 1.4 release are available packaged as separate downloads:

- QLogicIB-Basic with the Text User Interface Installer
- QLogic OFED 1.4 RPM Set
- QLogic OFED 1.4 User-level RPM Set
- Rocks Rolls
- Platform OCS Kits
- The InfiniBand Fabric Suite is available for purchase as a CD/ISO image

Check [Table 5-5](#) for the package contents available for each type of download.

All files are available from the QLogic web site: <http://www.qlogic.com>. Follow the Downloads tab to choose the appropriate download for your OS distribution, then follow the instructions for installing the QLogic OFED software in the following sections.

Table 5-5. Available Packages for QLogic OFED 1.4 Release

Package	Description	Installation and Documentation
QLogicIB-Basic	Includes: <ul style="list-style-type: none"> ■ QLogic OFED 1.4 ■ InfiniPath HCA driver ■ Optimized stack for MPI (PSM) ■ QLogic MPI ■ Other MPIs (MVAPICH and Open MPI compiled with GCC, PathScale, PGI, and Intel compilers) ■ User tools ■ QLogic SRP and VNIC ■ Text User Interface (TUI) Installer ■ FastFabric Enablement Tools 	Install using “ Install QLogicIB-Basic with the Installer Tool ” on page 5-7. Related Documentation: Readme and Release Notes <i>QLogic HCA and QLogic OFED Software Install Guide</i> <i>QLogic HCA and QLogic OFED Software Users Guide</i> . <i>QLogic ULP and Tools Reference Guide (OFED+ Users Guide)</i>

Table 5-5. Available Packages for QLogic OFED 1.4 Release (Continued)

Package	Description	Installation and Documentation
QLogic OFED 1.4 RPM Set	Includes: <ul style="list-style-type: none"> ■ QLogic OFED 1.4 ■ InfiniPath HCA driver ■ Optimized stack for MPI (PSM) ■ QLogic MPI ■ Other MPIs (MVAPICH and Open MPI compiled with GCC, PathScale, PGI, and Intel compilers) ■ User tools 	Install using “Using rpm to Install InfiniPath and OpenFabrics” on page 5-14 Related Documentation: Readme and Release Notes <i>QLogic HCA and QLogic OFED Software Install Guide</i> <i>QLogic HCA and QLogic OFED Software Users Guide.</i>
QLogic OFED 1.4 User-level Software RPM Set	Includes: <ul style="list-style-type: none"> ■ Optimized stack for MPI (PSM) ■ QLogic MPI ■ Other MPIs (MVAPICH and Open MPI compiled with GCC, PathScale, PGI, and Intel compilers) 	Install using “Install QLogic OFED User-level Software with the rpm Command” on page 5-17 Related Documentation: Readme and Release Notes <i>QLogic HCA and QLogic OFED Software Install Guide</i> <i>QLogic HCA and QLogic OFED Software Users Guide.</i> For installation over OFED 1.4 supplied from OpenFabrics or with Linux distribution.
QLogic OFED 1.4 Rocks Rolls	Includes: <ul style="list-style-type: none"> ■ QLogic OFED 1.4 ■ InfiniPath HCA driver ■ Optimized stack for MPI (PSM) ■ QLogic MPI ■ Other MPIs (MVAPICH and Open MPI compiled with GCC, PathScale, PGI, and Intel compilers) ■ User tools ■ QLogic SRP and VNIC ■ FastFabric Enablement Tools 	Install using “Install QLogic OFED Using Rocks” on page 5-20 Related Documentation: Readme and Release Notes <i>QLogic HCA and QLogic OFED Software Install Guide</i> <i>QLogic HCA and QLogic OFED Software Users Guide.</i>

Table 5-5. Available Packages for QLogic OFED 1.4 Release (Continued)

Package	Description	Installation and Documentation
QLogic OFED 1.4 Platform OCS Kits	Includes: <ul style="list-style-type: none"> ■ QLogic OFED 1.4 ■ InfiniPath HCA driver ■ Optimized stack for MPI (PSM) ■ QLogic MPI ■ Other MPIs (MVAPICH and Open MPI compiled with GCC, PathScale, PGI, and Intel compilers) ■ User tools ■ QLogic SRP and VNIC ■ FastFabric Enablement Tools 	Install using “ Install QLogic OFED Using a Platform OCS Kit ” on page 5-22 Related Documentation: Readme and Release Notes <i>QLogic HCA and QLogic OFED Software Install Guide</i> <i>QLogic HCA and QLogic OFED Software Users Guide.</i>
QLogic Infini-Band Fabric Suite	Includes: <ul style="list-style-type: none"> ■ QLogic FastFabric Toolset ■ QLogic Host Subnet Manager ■ QLogic Fabric Viewer ■ InfiniServ Host Software ■ QLogicIB-Basic 	Install using “ Install FastFabric Software CD/ISO Image ” on page 5-22 CD/ISO image may be purchased separately. Follow the links on the QLogic download page. Documentation is included.

Install QLogicIB-Basic with the Installer Tool

The Installer tool has a Text User Interface (TUI) that affords easy installation of the software. Use this method if you have downloaded the QLogic-IB-Basic package. This method is suitable for use on small clusters.

1. Use the Downloads tab on the QLogic web site to locate your adapter model.

<http://www.qlogic.com>

Download the QLogicIB-Basic tar file for your distribution to a directory that will not be deleted upon reboot. Then unpack the tar file:

```
$ tar zxvf <QLogicIB-Basic>.<version>.tgz
```

The `tar` command creates a directory based on the `tar` file name and places the RPMs and other files in this directory.

2. After unpacking the .tgz file, change directory to:

```
cd <QLogicIB-Basic>.<version>
```

3. Become root, then type:

```
# ./INSTALL
```

(Note that if you need 32-bit support on 64 bit OSs, invoke the installer with `./INSTALL --32bit`)

You will see a screen similar to this:

```
QLogic Inc. InfiniBand 4.4.1.0.8 Software
```

- 1) Install/Uninstall Software
 - 2) Reconfigure OFED IP over IB
 - 3) Reconfigure Driver Autostart
 - 4) Update HCA Firmware
 - 5) Generate Supporting Information for Problem Report
 - 6) Fast Fabric (Host/Chassis/Switch Setup/Admin)
- X) Exit

4. Type 1, which will display the screen for software installation. The next screen shows the packages to select for installation:

QLogic Inc. IB Install (4.4.1.0.8 release) Menu

Please Select Install Action (screen 1 of 3):

```

0) OFED IB Stack      [  Install  ][Available] 1.4.0.1.5
1) TrueScale HCA Libs [  Install  ][Available] 2.3.0.0.4237
2) QLogic IB Tools    [  Install  ][Available] 4.4.0.0.29
3) OFED IB Development [  Install  ][Available] 1.4.0.1.5
4) QLogic Fast Fabric [Don't Install][Not Avail]
5) QLogic SRP         [  Install  ][Available] 1.4.0.0.12
6) QLogic Virtual NIC [  Install  ][Available] 1.4.0.0.11
7) OFED IP over IB    [  Install  ][Available] 1.4.0.1.5
8) OFED SDP           [  Install  ][Available] 1.4.0.1.5
9) OFED uDAPL         [  Install  ][Available] 1.4.0.1.5
a) QLogic FM          [Don't Install][Not Avail]
b) MVAPICH (gcc)     [  Install  ][Available] 1.4.0.1.5
c) MVAPICH2 (gcc)    [  Install  ][Available] 1.4.0.1.5
d) OpenMPI (gcc)     [  Install  ][Available] 1.4.0.1.5

N) Next Screen
P) Perform the selected actions      I) Install All
R) Re-Install All                    U) Uninstall All
X) Return to Previous Menu (or ESC)

```

Pressing the keys corresponding to menu items (0-9, a-d in the example above) will toggle the selection for the given item.

The packages above are recommended for a new install. QLogic Fast Fabric (4) and QLogic FM (Fabric Manager) (a) are only available if you have purchased the InfiniBand Fabric Suite. MVAPICH2 (c) does not run over QLogic PSM; it runs over Open Fabrics Verbs only.

TrueScale HCA Libs (1) contains the enhanced InfiniPath HCA driver, optimized stack for MPI(PSM) and QLogic MPI, and user tools.

5. Then, type n to proceed to the next screen. You can cycle through the available screens by continuing to type n. (Typing x or pressing ESC will return you to the top level menu, and un-set all your current choices.)

The next screen contains these choices:

QLogic Inc. IB Install (4.4.1.0.8 release) Menu

Please Select Install Action (screen 2 of 3):

- 0) MVAPICH/PSM (gcc) [Install] [Available] 1.4.0.1.5
- 1) MVAPICH/PSM (PGI) [Install] [Available] 1.4.0.1.5
- 2) MVAPICH/PSM (PSc) [Install] [Available] 1.4.0.1.5
- 3) MVAPICH/PSM (Intel) [Install] [Available] 1.4.0.1.5
- 4) OpenMPI/PSM (gcc) [Install] [Available] 1.4.0.1.5
- 5) OpenMPI/PSM (PGI) [Install] [Available] 1.4.0.1.5
- 6) OpenMPI/PSM (PSc) [Install] [Available] 1.4.0.1.5
- 7) OpenMPI/PSM (Intel) [Install] [Available] 1.4.0.1.5
- 8) MPI Source [Install] [Available] 1.4.0.1.5
- 9) OFED RDS [Install] [Available] 1.4.0.1.5
- a) OFED SRP [Install] [Available] 1.4.0.1.5
- b) OFED SRP Target [Don't Install] [Available] 1.4.0.1.5
- c) OFED iSER [Don't Install] [Available] 1.4.0.1.5
- d) OFED iSER Target [Don't Install] [Available]

- N) Next Screen
- P) Perform the selected actions I) Install All
- R) Re-Install All U) Uninstall All
- X) Return to Previous Menu (or ESC)

Note that PSc is an acronym for the PathScale compiler. At this time, QLogic recommends choosing all items all items except for OFED SRP Target and the iSER options.

6. Then type n to proceed to the next screen:

QLogic Inc. IB Install (4.4.1.0.8 release) Menu

Please Select Install Action (screen 3 of 3):

- 0) OFED iWARP [Don't Install] [Available] 1.4.0.1.5
- 1) OFED Open SM [Don't Install] [Available] 1.4.0.1.5
- 2) OFED Debug Info [Don't Install] [Available] 1.4.0.1.5

- N) Next Screen
- P) Perform the selected actions I) Install All
- R) Re-Install All U) Uninstall All
- X) Return to Previous Menu (or ESC)

Open SM (1) should only be installed on one node in the cluster where it will be used. If desired, Type 1 for Open SM.

7. Finally, type `p` to start the installation.

The installer will uninstall older OFED RPMs, and then will ask for input for a series of operations. You can accept the defaults, by pressing `<Enter>`. Take note of the following cases:

```
Install MPI with prefix compatible with mpi-selector
(/usr/mpi/qlogic) [y]: y
```

This default allows you to use the `mpi-selector` to choose between different MPI implementations.

If you type `y`, make sure, after installation, that the environment variable `$MPICH_ROOT` is set to the same prefix that is used here (`/usr/mpi/qlogic`). When set, the `$MPICH_ROOT` variable allows QLogic MPI to correctly locate header and library files for MPI compilation and running parallel jobs. Typing `n` will cause QLogic MPI to be installed in the default `/usr`.

The next case is:

```
Configure OFED IP over IB IPV4 addresses now? [n]:
```

Answer `y` if the IB IPV4 addresses and netmasks are available, and you want to enter them now. Answer `n` if the IB addresses are not available or you want to add them later. IPoIB can be configured manually by following the instructions for [“Configuring the IPoIB Network Interface” on page 6-2](#).

Finally, QLogic recommends you answer `n` to the following:

```
Enable QLogic SRP (qlgc_srp) to autostart? [y]: n
Enable OFED SRP (openibd) to autostart? [y]: n
```

Further instructions for using SRP are given in [“SRP” on page 6-4](#).

8. Once the install has completed, quit the installer by typing `x` until you have exited. After exiting, rebooting the machine is recommended.

NOTE:

If you want support for 32 bit programs, you can install the 32-bit libraries on a 64-bit system like this:

```
# ./INSTALL --32bit
```

The Installer can also be used as a command-line interface (CLI). There are numerous options for installation/upgrade/uninstallation/autostart of all the available components. Here are the available options:

```
./INSTALL [-r root] [-v|-vv] [-a|-n|-U|-F|-u|-s|-i comp|-e comp|-E
comp|-D comp] [-f] [--user_configure_options 'options']
```

./INSTALL -c lists all the available components, which include:

```
ib_stack, truescale, oftools, ib_stack_dev, fastfabric, qlgc_srp,
qlgc_vnic, ofed_ipoib, ofed_sdp, ofed_udapl, qlgc_fm, mvapich,
mvapich2, openmpi, mvapich_gcc_qlc, mvapich_pgi_qlc,
mvapich_pathscale_qlc, mvapich_intel_qlc, openmpi_gcc_qlc,
openmpi_pgi_qlc, openmpi_pathscale_qlc, openmpi_intel_qlc,
ofed_mpirsrc, ofed_rds, ofed_srp, ofed_srpt, ofed_iser, ofed_isert,
ofed_iwarp, opensm ofed_debug
```

Additional component names allowed for **-E** and **-D** options:

```
iba_mon, qlgc_fm_snmp
```

NOTE:

The component names here are not the same as the RPM names in the RPM downloads, even though they are RPM-based in most cases. See [“Package Descriptions” on page D-1](#) for more package description.

[Table 5-6](#) summarizes Installer command line options.

Table 5-6. INSTALL Options

Command	Meaning
-a	Install all ULPs and drivers with default options
-n	Install all ULPs and drivers with default options but with no change to autostart options
-U	Upgrade/re-install all presently installed ULPs and drivers with default options and no change to autostart options
-i comp	Install the given component with default options can appear more than once on command line
-f	Skip firmware upgrade during install
--user_configure_options 'options'	Specify additional OFED build options for user space srpms. Causes rebuild of all user srpms
--kernel_configure_options 'options'	Specify additional OFED build options for driver srpms. Causes rebuild of all driver srpms
--prefix dir	Specify alternate directory prefix for install default is /usr. Causes rebuild of needed srpms
--no32bit	Disable install of 32 bit libraries on 64 bit OSs

Table 5-6. INSTALL Options (Continued)

Command	Meaning
<code>--32bit</code>	Enable install of 32 bit libraries on 64 bit OSs default is no32bit
<code>--rebuild</code>	Force OFED rebuild
<code>--force</code>	Force install even if distros don't match Use of this option can result in undefined behaviors
<code>-F</code>	Upgrade HCA Firmware with default options
<code>-u</code>	Uninstall all ULPs and drivers with default options
<code>-s</code>	Enable autostart for all installed drivers
<code>-r</code>	Specify alternate root directory, default is /
<code>-e comp</code>	Uninstall the given component with default options can appear more than once on command line
<code>-E comp</code>	Enable autostart of given component can appear with -D or more than once on command line
<code>-D comp</code>	Disable autostart of given component can appear with -E or more than once on command line
<code>-v</code>	Verbose logging
<code>-vv</code>	Very verbose debug logging
<code>-C</code>	Output list of supported components
<code>--user_queries</code>	Permit non-root users to query the fabric (Default)
<code>--no_user_queries</code>	Non-root users cannot query the fabric. Default options retain existing configuration files.

About rpm Installation

Linux distributions of QLogic OFED (InfiniPath and OpenFabrics) software can be installed from binary RPMs. RPM is a Linux packaging and installation tool used by Red Hat, SUSE, and CentOS.

Instructions given below are for a single node. Parallel command starters can be used for installation on multiple nodes. The Yellowdog Updater, Modified (YUM) may also be used for installation. However, these subjects are beyond the scope of this document.

RPMs contain `config` files. Your current `config` files will not be overwritten when new RPMs are installed. New `config` files will contain the suffix `.rpmnew` and can be found in `/etc/sysconfig` and `/etc/infiniband`. Check the new files to see if there is anything you want to add to your standard config files

Please note:

- For convenience, QLogic recommends that the same set of RPMs are installed on all nodes (with the exception of OpenSM). Omitting the `*-Static/*` and `*-Debuginfo/*` RPMs is recommended. Use the `*/32bit/*` RPMs only if you need them. Some RPMs are optional depending on which type of node is being used. To see which RPMs are required or optional for each type of node, according to its function as a compute node, front end node, development machine, or Subnet Manager (SM), see [Appendix D “Package Descriptions”](#).
- Install the `OpenSM` RPM only if you do not plan to use a switch-based or host-based SM. The `OpenSM` RPM is normally installed on the node on which it will be used. If installed, it is *off* by default. This behavior can be modified. See [“OpenSM” on page 6-3](#) for more information.
- Programs that incorporate the user IB verbs interfaces, such as diagnostics, benchmarks, verbs-based MPIs (for example, Intel MPI), and SDP sockets must have the OpenFabrics RPMs installed.
- Install the `infinipath` RPM on all nodes where you install the `mpi-frontend` RPM.
- The `mpi-devel` and `infinipath-devel` RPMs will be installed when the `qlogic-mpi-register` RPM is installed, as there are dependencies.
- Check that all older stock OFED RPMs have been uninstalled ([“Uninstalling InfiniPath and OpenFabrics RPMs” on page 5-24](#))

Using `rpm` to Install InfiniPath and OpenFabrics

1. Use the Downloads tab on the QLogic web site to locate your adapter model.

<http://www.qlogic.com>

After downloading the appropriate tar file, type:

```
$ tar zxvf QLogicOFED<version>-<distribution>-x86_64.tgz
```

The `tar` command creates a directory based on the `tar` file name and places the RPMs and other files in this directory.

2. The RPMs need to be available on each node on which they will be used. You can copy the RPMs to a directory on each node that will need them.

Become root, then:

```
# cd QLogicOFED<version>-<distribution-x86_64
```

(Another way is to put the RPMs in a directory that is accessible (e.g., via Network File System (NFS)) to every node.)

NOTE:

If you wish to use the `mpi-selector` to switch between QLogic MPI and other MPI implementations, you need to install QLogic MPI in an alternate location, consistent with that of the other MPIs. Skip to Step 4.

3. To install InfiniPath, QLogic MPI, and OpenFabrics, run the command (as root):

```
# rpm -Uvh InfiniPath/*.rpm InfiniPath-MPI/*.rpm \
InfiniPath-MPI/32bit/mpi-frontend-*.rpm \
InfiniPath-Devel/*.rpm Documentation/*.rpm OtherMPIs/*.rpm \
OpenFabrics/*.rpm OpenFabrics-Devel/*.rpm \
OtherHCAs/*.rpm OtherHCAs-Devel/*.rpm
```

Note that you need to install the `InfiniPath-MPI/32bit/mpi-frontend-*.rpm` even if you do not plan to use any other 32-bit RPMs. However, you can add the other 32bit subdirectories to the `rpm` command, if you need 32-bit support.

Install the `OpenSM` RPM only if you do not plan to use a switch-based or host-based SM. The `OpenSM` RPM is normally installed on the node where it will be used. If installed, it is *off* by default. This behavior can be modified. See [“OpenSM” on page 6-3](#) for more information.

To add Open SM, add this command:

```
# rpm -Uvh OpenSM/*.rpm OpenSM-Devel/*.rpm
```

Add the desired 32bit subdirectories to the `rpm` commands, if you need 32-bit support.

Proceed to Step 5.

4. To install QLogic MPI in an alternate location, use these commands instead of those in Step 3:

```
# mkdir QLogic-MPI-prefixed
# mv InfiniPath-MPI/mpi-* \
InfiniPath-MPI/32bit/mpi-frontend-*.rpm \
InfiniPath-Devel/mpi-devel* \
OtherMPIs/qlogic-mpi-register* \
Documentation/mpi-doc* QLogic-MPI-prefixed/
```

Note that you need to install the `InfiniPath-MPI/32bit/mpi-frontend-*.rpm` even if you do not plan to use any other 32-bit RPMs. However, you can add the other 32-bit subdirectories to the `rpm` command, if you need 32-bit support.

Next, install all non-prefixed RPMs:

```
# rpm -Uvh InfiniPath/*.rpm \  
InfiniPath-Devel/infinipath-devel*.rpm \  
OpenFabrics/*.rpm OpenFabrics-Devel/*.rpm \  
OtherHCAs/*.rpm OtherHCAs-Devel/*.rpm \  
Documentation/infinipath-doc*.rpm \  
Documentation/ofed-doc*.rpm \  
OtherMPIs/mpi-selector*.rpm OtherMPIs/mpitests*.rpm \  
OtherMPIs/mvapich*.rpm OtherMPIs/openmpi*.rpm
```

Finally, install the prefixed QLogic-MPI RPMs in `/usr/mpi/qlogic`:

```
# rpm -Uvh --prefix /usr/mpi/qlogic QLogic-MPI-prefixed/*.rpm
```

The desired prefix should be made available in the `$MPICH_ROOT` environment variable, either by global shell configuration files or through third-party environment management utilities such as `mpi-selector` or the Environment Modules. This allows QLogic MPI to correctly locate header and library files for MPI compilation and running parallel jobs. See the *QLogic HCA and QLogic OFED Software Users Guide* for details.

5. Reboot.

RPM Organization

The complete RPM directories are organized as follows. Note that the suggested installation does not include all possible RPM directories. Install the files in the `*-Devel` directories if you are going use source code to do development work. The `*-Debuginfo` directories contain debug information and possibly source code, which may be useful for testing, debugging, and developing applications. The `*-Static` directories contain the static versions of the libraries, which can be used in place of the dynamic libraries when compiling and linking.

```
InfiniPath_license.txt,LEGAL.txt (top level)  
Documentation/  
InfiniPath/  
InfiniPath/32bit  
InfiniPath-Devel/  
InfiniPath-MPI/  
InfiniPath-MPI/32bit  
OpenFabrics/
```

```
OpenFabrics/32bit/  
OpenFabrics-Static/  
OpenFabrics-Static/32bit/  
OpenFabrics-Devel/  
OpenFabrics-Devel/32bit/  
OpenFabrics-Devel-Static/  
OpenFabrics-Debuginfo/  
OpenFabrics-Debuginfo/32bit/  
OpenSM/  
OpenSM/32bit/  
OpenSM-Debuginfo/  
OpenSM-Debuginfo/32bit/  
OpenSM-Devel/  
OpenSM-Devel/32bit/  
OpenSM-Static/  
OpenSM-Static/32bit/  
OtherHCAs/  
OtherHCAs-Debuginfo  
OtherHCAs-Debuginfo/32bit/  
OtherHCAs-Devel/  
OtherHCAs-Devel/32bit/  
OtherHCAs-Devel-Static/  
OtherHCAs-Devel-Static/32bit/  
OtherMPIs/
```

Install QLogic OFED User-level Software with the `rpm` Command

The QLogic User-level software is for installation over OFED 1.4 supplied from OpenFabrics or with Linux distribution.

For convenience, QLogic recommends that all RPMs are installed on all nodes. Please note that the `infinipath` RPM must be installed on all nodes where you install the `mpi-frontend` RPM.

1. Use the Downloads tab on the QLogic web site to locate your adapter model.

<http://www.qlogic.com>

After downloading the appropriate tar file, type:

```
$ tar zxvf  
InfiniPath<version>-<date>-<distribution>-x86_64.tgz
```

The `tar` command creates a directory based on the `tar` file name and places the RPMs and other files in this directory.

2. The RPMs need to be available on each node on which they will be used. You can copy the RPMs to a directory on each node that will need them.

Become root, then `cd` to:

```
# cd InfiniPath<version>-<date>-<distribution>-x86_64
```

(Another way is to put the RPMs in a directory that is accessible (e.g., via Network File System (NFS)) to every node.

NOTE:

If you wish to use the `mpi-selector` to switch between QLogic MPI and other MPI implementations, you need to install QLogic MPI in an alternate location, consistent with that of the other MPIs. Skip to Step 4.

3. To install, use this command (as root):

```
# rpm -Uvh InfiniPath/*.rpm InfiniPath-MPI/*.rpm \  
InfiniPath-Devel/*.rpm OtherMPIs/*.rpm \  
Documentation/*.rpm
```

Proceed to Step 5.

4. To install QLogic MPI in an alternate location, use these commands instead of those in Step 3:

```
# mv InfiniPath-MPI/*.rpm InfiniPath-Devel/mpi-devel* \  
Documentation/mpi-doc* QLogic-MPI-prefixed/
```

Next, install all non-prefixed RPMs:

```
# rpm -Uvh InfiniPath/*.rpm \  
InfiniPath-Devel/infinipath-devel*.rpm \  
Documentation/infinipath-doc*.rpm OtherMPIs/*.rpm
```

Finally, install the prefixed RPMs in `/usr/mpi/qlogic`:

```
# rpm -Uvh --prefix /usr/mpi/qlogic QLogic-MPI-prefixed/*.rpm
```

The desired prefix should be made available in the `$MPICH_ROOT` environment variable, either by global shell configuration files or through third-party environment management utilities such as `mpi-selector` or the Environment Modules. This allows QLogic MPI to correctly locate header and library files for MPI compilation and running parallel jobs. See the *QLogic HCA and QLogic OFED Software Users Guide* for details.

5. Reboot.

The complete RPM directories for this download are organized as follows:

```
InfiniPath_license.txt,LEGAL.txt (top level)
Documentation/
InfiniPath/
InfiniPath-Devel/
InfiniPath-MPI/
OtherMPIs/
```

Rebuilding or Reinstalling the `kernel-ib` Driver with `rpm` After a Kernel Upgrade

If you upgrade the kernel, then you must reboot and then rebuild or reinstall the InfiniPath kernel modules (drivers).

To rebuild the drivers, do the following (as root):

```
# cd /usr/src/qlogic_ib/kernel-ib-<version>
# ./make-install.sh
```

```
# /etc/init.d/openibd restart
```

An alternative method is to reinstall the InfiniPath kernel modules and then restart the InfiniPath service. Type (as root):

```
# rpm -U --replacepkgs kernel-ib-*
# /etc/init.d/openibd restart
```

Rebuilding the `kernel-ib` Driver on an Unsupported Distribution or an Unsupported Distribution/Kernel Pair

If the `rpm` install can't correctly determine the underlying distribution/kernel combination, you will see a warning message. This can occur if you try to install on an unsupported distribution or an unsupported distribution/kernel pair.

In this case you need to use `IPATH_DISTRO` to override the distribution version provided in either the `/etc/redhat-release` file or the `/etc/SuSE-release` file.

Here is an example. These commands must be done as root:

```
# export IPATH_DISTRO=2.6.18_EL5.1 KVER=2.6.18-53.1.14.el5
# cd /usr/src/qlogic_ib/kernel-ib-<version>
# ./make-install.sh
# /etc/init.d/openibd restart
```

NOTE:

Using the override may not result in a buildable or working driver if your distribution/kernel combination is not similar enough to a tested and supported distribution/kernel pair.

Install QLogic OFED Using Rocks

Rocks is a distribution designed for managing clusters from the San Diego Supercomputer center (SDSC).

Rocks is a way to manage the *kickstart* automated installation method created by Red Hat. By using the Rocks conventions, the installation process can be automated for clusters of any size. A *Roll* is an extension to the Rocks base distribution that supports different cluster types or provides extra functionality.

QLogic extends the normal Rocks compute node appliance xml file by adding two functions: one function installs the InfiniPath software, and the other function loads the drivers after kickstart reboots the machine.

Install Frontend and Compute Nodes

Rocks is based on a set of "kickstart graphs" (xml files) that tell the frontend node which pieces need to be installed on which type of compute nodes. The frontend node installs from local RPMs, and the compute nodes then collect the RPMs from the frontend. You will first have to install a Rocks frontend node, if you don't already have one.

To install a Rocks frontend node:

1. Use the Downloads tab on the QLogic web site to locate your adapter model.
<http://www.qlogic.com>
2. Download the InfiniPath roll (.iso image) for your distribution and burn the .iso image to a CD.
3. Download the required rolls from the Rocks web site:
<http://www.rocksclusters.org>
 - a. Follow the links to get the following .iso images that can be burned to a CD or DVD:
Kernel/Boot Roll
Core Roll
OS Roll (Disk 1 & 2)
-OR-

Boot, Core, OS Roll DVD

Note that you may also need updates; look for the latest files with the service-pack prefix. Make sure you downloaded the .iso images correctly; verify by checking the md5 checksum from the web site.

- b. Burn the .iso image(s) to a CD(s) or DVD.
4. Build the frontend node with the above .iso images (step 2) from the Rocks web site:
 - a. Insert the Kernel/Boot Roll CD into your frontend machine. After the frontend boots off the CD, follow the instructions on the screen. Insert the OS Rolls and any other of the Rocks rolls you need when prompted.
 - b. To install the InfiniPath roll, put the InfiniPath CD into the drive when prompted if you wish to install additional rolls. Follow the instructions. For more details, see the Rocks installation documentation on the Rocks web site:

<http://www.rocksclusters.org>

5. Install the compute nodes. Login to the frontend node as root, and run the command:

```
# insert-ethers
```

This runs a program which captures compute node DHCP requests and puts the information into the Rocks MySQL database. Follow the instructions on the Rocks web site:

<http://www.rocksclusters.org>

6. Once Rocks is up and running, test the rocks cluster according to your own testing procedures.

NOTE:

You can also get rolls directly from Platform OCS or ClusterCorp. The web sites are:

<http://my.platform.com/products/platform-ocs>

<http://www.clustercorp.com/>

Rocks Installation on an Existing Frontend Node

If the frontend node has already been installed, you can add the InfiniPath roll to the repository on the head node, update the master graph.xml, and re-install all the compute nodes. Use these instructions. You must be root.

If you have a burned a CD version of InfiniPath roll from the .iso image:

```
# mount /mnt/cdrom
# rocks-dist --install copyroll
# umount /mnt/cdrom
# cd /home/install
# rocks-dist dist
```

You can use this method if you download the .iso image without burning a CD:

```
# mount -o loop <package>.iso /mnt/cdrom
# rocks-dist --install copyroll
# umount /mnt/cdrom
# cd /home/install
# rocks-dist dist
```

Then use the following command for each node:

```
# shoot-node <compute_node_name>
```

or

Use the following command to rebuild the entire cluster:

```
# cluster-fork /boot/kickstart/cluster-kickstart
```

Install QLogic OFED Using a Platform OCS Kit

The Platform OCS Kit is an ISO image that installs the drivers in an automated way. Kits are a mechanism for packaging install scripts and applications for easy installation onto a Platform OCS cluster. To get started:

1. Use the Download tab on the QLogic web site to locate your adapter model.
<http://www.qlogic.com>
2. Download the InfiniPath Platform OCS Kit (.iso image) for your distribution.
3. Follow the instructions provided by Platform OCS for installing Kits. See:
<http://my.platform.com/products/platform-ocs>

Install FastFabric Software CD/ISO Image

The FastFabric CD/ISO image may be purchased separately.

1. Use the Download tab on the QLogic web site to locate your adapter model.
<http://www.qlogic.com>
2. Find the FastFabric link, and follow instructions for purchasing the software.
3. Instructions for installation are with the software.

See also the current version of the FastFabric Users Guide, available with the software or from the QLogic download page.

Install Additional Software

This section contains information about additional third-party software installation.

Installing Lustre

This InfiniPath release supports Lustre cluster filesystem Version 1.6.5.1. Lustre is a fast, scalable Linux cluster file system that interoperates with InfiniBand. For general instructions on downloading, installing, and using Lustre, go to:

<http://www.lustre.org>

Installed Layout

The default installed layout for the InfiniPath software and QLogic-supplied MPIs is described here.

The shared libraries are installed in:

```
/usr/lib for 32-bit applications  
/usr/lib64 for 64-bit applications
```

MPI include files are in:

```
/usr/include
```

MPI programming examples and the source for several MPI benchmarks are in:

```
/usr/share/mpich/examples
```

NOTE:

If QLogic MPI is installed in an alternate location, the argument passed to `--prefix (/usr/mpi/qlogic)` replaces the default `/usr` prefix. QLogic MPI binaries, documentation, and libraries are installed under that prefix. However, a few configuration files are installed in `/etc` regardless of the desired `--prefix`. The remaining InfiniPath libraries and tools stay their default installation location.

If you have installed the software into an alternate location, the `$MPICH_ROOT` environment variable needs to match `--prefix`.

InfiniPath utility programs, as well as MPI utilities and benchmarks, are installed in:

```
/usr/bin
```

Documentation is found in:

```
/usr/share/man  
/usr/share/doc/infinipath  
/usr/share/doc/mpich-infinipath
```

Note that license information only is found in `usr/share/doc/infinipath`. InfiniPath user documentation can be found on the QLogic web site on the software download page for your distribution.

Configuration files are found in:

```
/etc/sysconfig
```

Init scripts are found in:

```
/etc/init.d
```

The InfiniPath driver modules in this release are installed in:

```
/lib/modules/$(uname -r)/updates/kernel/drivers/infiniband/hw/ipath
```

Most of the other OFED modules are installed under the `infiniband` subdirectory, above. Some others are installed under:

```
/lib/modules/$(uname -r)/updates/kernel/drivers/net
```

The RDS modules are installed under:

```
/lib/modules/$(uname -r)/updates/kernel/net/rds
```

QLogic-supplied OpenMPI and MVAPICH RPMs with PSM support and compiled with GCC, PathScale, PGI, and the Intel compilers are now installed in directories using this pattern:

```
/usr/mpi/<compiler>/<mpi>-<mpi_version>-qlc
```

For example:

```
/usr/mpi/gcc/openmpi-1.2.8-qlc
```

Removing Software Packages

This section provides instructions for uninstalling or downgrading the InfiniPath and OpenFabrics software.

Uninstalling Using the Installer Tool

Software packages can be removed by using the Installer tool. Instructions will be similar to those given in [“Install QLogicIB-Basic with the Installer Tool” on page 5-7](#), except that you select the Uninstall option for the desired packages.

Uninstalling InfiniPath and OpenFabrics RPMs

QLogic recommends uninstalling the OFED software first, before uninstalling the InfiniPath software.

For both InfiniPath and OpenFabrics, QLogic recommends that you remove all the packages at the same time.

1. Use the script `ofed_uninstall.sh` to uninstall the OFED software. See the OFED Installation release notes that are part of the OFED documentation.
2. To uninstall the InfiniPath software packages on any node with the `rpm` command, type the following command (as root) using a `bash` shell:

```
# rpm -e --allmatches `rpm -qa | grep qlc`
```

The “`qlc`” is a part of all the InfiniPath package names.

Uninstalling Software with Rocks or Platform OCS

Follow the instructions for either Rocks or Platform OCS to uninstall software. See:

<http://www.rocksclusters.org>

<http://my.platform.com/products/platform-ocs>

Downgrading RPMs

If you want to downgrade, remove both the InfiniPath and OpenFabrics RPMs, then install the older bits. QLogic has determined that `rpm` flags like “`--oldpackage`” will not generate a correct downgrade.

NOTE:

Use the `rpm` method for downgrading rather than the Installer tool.

Notes

Draft

6 Configuring Drivers and Services

This section provides instructions for configuring and using the drivers and services available with QLogic OFED 1.4.

InfiniPath and OpenFabrics Driver Overview

The InfiniPath `ib_ipath` module provides low level QLogic hardware support, and is the base driver for both MPI/PSM programs, and general OpenFabrics protocols such as IPoIB and SDP. The driver also supplies the Subnet Management Agent (SMA) component.

Optional configurable OpenFabrics components and their default settings at startup are:

- IPoIB network interface. Required for TCP/IP networking for running Ethernet traffic over the InfiniPath link. It is not running until it is configured.
- VNIC. It is not running until it is configured.
- OpenSM. It is disabled on startup. You can either install it on only one node, or disable it on all nodes except where it will be used as an SM.
- SRP (OFED and QLogic modules). SRP is not running until the module is loaded and the SRP devices on the fabric have been discovered.
- MPI over uDAPL (can be used by Intel MPI or HP-MPI). IPoIB needs to be configured before MPI over uDAPL can be set up.

Other optional drivers can now be configured and enabled, as described in [“OpenFabrics Drivers and Services Configuration and Startup” on page 6-1](#).

Complete information about starting, stopping, and restarting the InfiniPath services are in [“Managing the InfiniPath Driver” on page 6-15](#).

OpenFabrics Drivers and Services Configuration and Startup

IPoIB, VNIC, OpenSM, SRP, and MPI over uDAPL configuration and startup is explained in more detail in the following sections.

Configuring the IPoIB Network Interface

The following instructions show you how to manually configure your OpenFabrics IPoIB network interface. This example assumes that you are using `sh` or `bash` as your shell, all required InfiniPath and OpenFabrics RPMs are installed, and your startup scripts have been run (either manually or at system boot).

For this example, the IPoIB network is 10.1.17.0 (one of the networks reserved for private use, and thus not routable on the Internet), with a /8 host portion, and therefore requires that the netmask be specified.

This example assumes that no hosts files exist, the host being configured has the IP address 10.1.17.3, and DHCP is not used.

NOTE:

Instructions are only for this static IP address case. Configuration methods for using DHCP will be supplied in a later release.

1. Type the following commands (as root):

```
# ifconfig ib0 10.1.17.3 netmask 0xffffffff00
```

2. To verify the configuration, type:

```
# ifconfig ib0
```

The output from this command will be similar to this:

```
ib0    Link encap:InfiniBand HWaddr  
00:00:00:02:FE:80:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00  
inet  addr:10.1.17.3  Bcast:10.1.17.255  Mask:255.255.255.0  
UP BROADCAST RUNNING MULTICAST  MTU:4096  Metric:1  
RX packets:0 errors:0 dropped:0 overruns:0 frame:0  
TX packets:0 errors:0 dropped:0 overruns:0 carrier:0  
collisions:0 txqueuelen:128  
RX bytes:0 (0.0 b)  TX bytes:0 (0.0 b)
```

3. Type:

```
# ping -c 2 -b 10.1.17.255
```

The output of the `ping` command will be similar to the following, with a line for each host already configured and connected:

```
WARNING: pinging broadcast address
PING 10.1.17.255 (10.1.17.255) 517(84) bytes of data.
174 bytes from 10.1.17.3: icmp_seq=0 ttl=174 time=0.022 ms
64 bytes from 10.1.17.1: icmp_seq=0 ttl=64 time=0.070 ms
(DUP!)
64 bytes from 10.1.17.7: icmp_seq=0 ttl=64 time=0.073 ms
(DUP!)
```

The IPoIB network interface is now configured.

4. Restart (as root) by typing:

```
# /etc/init.d/openibd restart
```

NOTE:

- The configuration must be repeated each time the system is rebooted.
- IPoIB-CM (Connected Mode) is enabled by default. The setting in `/etc/infiniband/openib.conf` is `SET_IPOIB_CM=yes`. To use datagram mode use `SET_IPOIB_CM=no`.

OpenSM

OpenSM is an optional component of the OpenFabrics project that provides a subnet manager for InfiniBand networks. This package can be installed on all machines, but only needs to be enabled on the machine in the cluster that will act as a subnet manager. You do not need to use OpenSM if any of your InfiniBand switches provide a subnet manager, or if you are running a host-based SM.

If you are using the Installer tool, you can set the OpenSM default behavior at the time of installation.

If you are using the `rpm` install method, note that after installing the `opensm` package, OpenSM is configured to be *off* after the next machine reboot. It only needs to be enabled on the node that acts as the subnet manager, so use the `chkconfig` command (as root) to enable it on the node where it is to run:

```
# chkconfig opensmd on
```

The command to disable it on reboot is:

```
# chkconfig opensmd off
```

You can start `opensmd` without rebooting your machine by typing:

```
# /etc/init.d/opensmd start
```

You can stop `opensmd` again like this:

```
# /etc/init.d/opensmd stop
```

If you want to pass any arguments to the OpenSM program, modify the following file, and add the arguments to the `OPTIONS` variable:

```
/etc/init.d/opensmd
```

For example:

```
# Use the UPDN algorithm instead of the Min Hop algorithm.  
OPTIONS="-R updn"
```

For more information on OpenSM, see the OpenSM `man` pages, or look on the OpenFabrics web site.

SRP

SRP stands for SCSI RDMA Protocol. It was originally intended to allow the SCSI protocol to run over InfiniBand for Storage Area Network (SAN) usage. SRP interfaces directly to the Linux file system through the SRP Upper Layer Protocol. SRP storage can be treated as another device.

In this release, two versions of SRP are available: QLogic SRP and OFED SRP. QLogic SRP is available as part of the QLogicIB-Basic*, Rocks Roll, and Platform OCS downloads.

It is not available as a part of the RPM downloads.

SRP has been tested on targets from Engenio™ (now LSI Logic®).

NOTE:

Before using SRP, the SRP targets must already be set up by your system administrator.

Using QLogic SRP

If you have installed QLogic SRP as part of the QLogicIB-Basic download, you will need to configure it according to the steps shown in the *QLogic ULP and Tools Reference Guide (OFED+ Users Guide)*.

Using OFED SRP

To use OFED SRP, follow these steps:

1. Add the line `SRP_LOAD=yes` to the module list in `/etc/infiniband/openib.conf` to have it automatically loaded.
2. Discover the SRP devices on your fabric by running this command (as root):

```
# ibsrpdm
```

In the output, look for lines similar to these:

```
GUID:      0002c90200402c04
ID:        LSI Storage Systems SRP Driver 200400a0b8114527
service entries: 1
service[ 0]: 200400a0b8114527 / SRP.T10:200400A0B8114527
```

```
GUID:      0002c90200402c0c
ID:        LSI Storage Systems SRP Driver 200500a0b8114527
           service entries: 1
service[ 0]: 200500a0b8114527 / SRP.T10:200500A0B8114527
```

```
GUID:      21000001ff040bf6
ID:        Data Direct Networks SRP Target System
           service entries: 1
service[ 0]: f60b04ff01000021 / SRP.T10:21000001ff040bf6
```

Note that not all the output is shown here; key elements are expected to show the match in [Step 3](#).

3. Choose the device you want to use, and run the command again with the `-c` option (as root):

```
# ibsrpdm -c
id_ext=200400A0B8114527,ioc_guid=0002c90200402c04,dgid=fe8000
00000000000002c90200402c05,pkey=ffff,service_id=200400a0b8114
527
id_ext=200500A0B8114527,ioc_guid=0002c90200402c0c,dgid=fe8000
00000000000002c90200402c0d,pkey=ffff,service_id=200500a0b8114
527
id_ext=21000001ff040bf6,ioc_guid=21000001ff040bf6,dgid=fe8000
000000000021000001ff040bf6,pkey=ffff,service_id=f60b04ff0100
021
```

4. Find the result that corresponds to the target you want, and echo it into the `add_target` file:

```
# echo
"id_ext=21000001ff040bf6,ioc_guid=21000001ff040bf6,dgid=fe800
0000000000021000001ff040bf6,pkey=ffff,service_id=f60b04ff0100
0021,initiator_ext=0000000000000001" >
/sys/class/infiniband_srp/srp-ipath0-1/add_target
```

5. You can look for the newly created devices in the `/proc/partitions` file. The file will look similar to this example (the partition names may vary):

```
# cat /proc/partitions
major minor #blocks name
8      64  142325760 sde
8      65  142319834 sde1
8      80   71162880 sdf
8      81   71159917 sdf1
8      96    20480 sdg
8      97    20479 sdg1
```

6. Create a mount point (as root) where you will mount the SRP device. For example:

```
# mkdir /mnt/targetname
# mount /dev/sde1 /mnt/targetname
```

NOTE:

Use `sde1` rather than `sde`. See the `mount(8)` man page for more information on creating mount points.

Configuring and Administering the VNIC Interface

The VirtualNIC (VNIC) Upper Layer Protocol (ULP) works in conjunction with firmware running on Virtual Input/Output (VIO) hardware such as the SilverStorm Ethernet Virtual I/O Controller (EVIC™) or the InfiniBand/Ethernet Bridge Module for IBM® BladeCenter®, providing virtual Ethernet connectivity.

The VNIC driver along with the QLogic EVIC's two 10 Gigabit ethernet ports, enables Infiniband clusters to connect to Ethernet networks. This driver also works with the earlier version of the I/O Controller, the VEx.

The QLogic VNIC driver creates virtual Ethernet interfaces and tunnels the Ethernet data to/from the EVIC over InfiniBand using an InfiniBand reliable connection.

The virtual Ethernet interface supports any Ethernet protocol. It operates normally like any other interface--`ping`, `ssh`, `scp`, `netperf` etc.

The VNIC interface must be configured before it can be used. To do so, perform the following steps:

1. Discover the EVIC/VEx Input/Output Controllers (IOCs) present on the fabric using `ib_qlge_vnic_query`. For writing the configuration file, you will need information about the EVIC/VEx IOCs present on the fabric, such as

their IOCGUID, IOCSTRING, etc. Use the `ib_qlgc_vnic_query` tool to get this information.

When `ib_qlgc_vnic_query` is executed without any options, it displays detailed information about all the EVIC/VEx IOCs present on the fabric. Run it as root. For example:

```
# ib_qlgc_vnic_query
HCA No = 0, HCA = mlx4_0, Port = 1, Port GUID = 0x0002c903000010f9,
State = Active
  IO Unit Info:
    port LID:      0009
    port GUID:     fe8000000000000000066a11de000070
    change ID:    0003
    max controllers: 0x02

  controller[ 1]
    GUID:      00066a01de000070
    vendor ID: 00066a
    device ID: 000030
    IO class : 2000
    ID:        EVIC in Chassis 0x00066a00db00001e, Slot 1, Ioc 1
    service entries: 2
      service[ 0]: 1000066a00000001 /
                   InfiniNIC.InfiniConSys.Control:01
      service[ 1]: 1000066a00000101 /
                   InfiniNIC.InfiniConSys.Data:01

  IO Unit Info:
    port LID:      000b
    port GUID:     fe8000000000000000066a21de000070
    change ID:    0003
    max controllers: 0x02

  controller[ 2]
    GUID:      00066a02de000070
    vendor ID: 00066a
    device ID: 000030
    IO class : 2000
    ID:        EVIC in Chassis 0x00066a00db00001e, Slot 1, Ioc 2
    service entries: 2
      service[ 0]: 1000066a00000002 /
                   InfiniNIC.InfiniConSys.Control:02
      service[ 1]: 1000066a00000102 /
                   InfiniNIC.InfiniConSys.Data:02
```

```
HCA No = 0, HCA = mlx4_0, Port = 2, Port GUID = 0x0002c903000010fa,
State = Active
IO Unit Info:
    port LID:          0009
    port GUID:         fe800000000000000000000066a11de000070
    change ID:         0003
    max controllers:  0x02

controller[ 1]
    GUID:              00066a01de000070
    vendor ID:         00066a
    device ID:         000030
    IO class :         2000
    ID:                EVIC in Chassis 0x00066a00db00001e, Slot 1, Ioc 1
    service entries:  2
        service[ 0]:  1000066a00000001 /
                       InfiniNIC.InfiniConSys.Control:01
        service[ 1]:  1000066a00000101 /
                       InfiniNIC.InfiniConSys.Data:01

IO Unit Info:
    port LID:          000b
    port GUID:         fe800000000000000000000066a21de000070
    change ID:         0003
    max controllers:  0x02

controller[ 2]
    GUID:              00066a02de000070
    vendor ID:         00066a
    device ID:         000030
    IO class :         2000
    ID:                EVIC in Chassis 0x00066a00db00001e, Slot 1, Ioc 2
    service entries:  2
        service[ 0]:  1000066a00000002 /
                       InfiniNIC.InfiniConSys.Control:02
        service[ 1]:  1000066a00000102 /
                       InfiniNIC.InfiniConSys.Data:02
```

NOTE:

A VIO hardware card can contain up to six IOCs (and therefore up to six IOCGUIDs); one for each Ethernet port on the VIO hardware card. Each VIO hardware card contains a unique set of IOCGUIDs: (e.g., IOC 1 maps to Ethernet Port 1, IOC 2 maps to Ethernet Port 2, IOC 3 maps to Ethernet Port 3, etc.).

2. Create the VNIC interfaces using the configuration file:

```
/etc/infiniband/qlgc_vnic.cfg.
```

Look at the `qlgc_vnic.cfg.sample` file to see how VNIC configuration files are written. It can be found with the OFED documentation, or in the `qlgc_vnic_tools` subdirectory of the QLogicIB_Basic download. You can use this configuration file as the basis for creating a configuration file by replacing the Destination Global Identifier (DGID), IOCGUID, and IOCSTRING values with those of the EVIC/VEx IOCs present on your fabric.

QLogic recommends using the DGID of the EVIC/VEx IOC, as it ensures the quickest startup of the VNIC service. When DGID is specified, the IOCGUID must also be specified. For more details, see the `qlgc_vnic.cfg` sample file.

3. Edit the VirtualNIC configuration file, `/etc/infiniband/qlgc_vnic.cfg`. For each IOC connection, add a CREATE block to the file using the following format:

```
{CREATE; NAME="eioc2";
PRIMARY={IOCGUID=0x66A0130000105; INSTANCE=0; PORT=1; }
SECONDARY={IOCGUID=0x66A013000010C; INSTANCE=0; PORT=2;}
}
```

NOTE:

The `qlgc_vnic.cfg` file is case and format sensitive.

- a. Format 1: Defining an IOC using the IOCGUID. Use the following format to allow the host to connect to a specific VIO hardware card, regardless of which chassis and/or slot the VIO hardware card resides:

```
{CREATE;
NAME="eioc1";
IOCGUID=0x66A0137FFFE7;}
The following is an example of VIO hardware failover:
{CREATE; NAME="eioc1";
PRIMARY={IOCGUID=0x66a01de000003; INSTANCE=1; PORT=1; }
SECONDARY={IOCGUID=0x66a02de000003; INSTANCE=1; PORT=1;}
}
```

NOTE:

Do not create EIOC names with similar character strings (e.g., eioc3 and eioc30). There is a limitation with certain Linux operating systems that cannot recognize the subtle differences. The result is that the user will be unable to ping across the network.

- b. Format 2: Defining an IOC using the IOCSTRING. Defining the IOC using the IOCSTRING allows VIO hardware to be hot-swapped in and out of a specific slot. The host attempts to connect to the specified IOC (1, 2, or 3) on the VIO hardware that currently resides in the specified slot of the specified chassis. Use the following format to allow the host to connect to a VIO hardware that resides in a specific slot of a specific chassis:

```
{CREATE;  
NAME="eioc1";  
IOCSTRING="Chassis 0x00066A0005000001, Slot 1, IOC 1";  
RX_CSUM=TRUE;  
HEARTBEAT=100; }
```

NOTE:

The IOCSTRING field is a literal, case-sensitive string. Its syntax must be exactly in the format shown in the previous example, including the placement of commas. To reduce the likelihood of syntax error, use the command `ib_qlgc_vnic_query -es`. Note that the chassis serial number must match the chassis Ox (Hex) value. The slot serial number is specific to the line card as well.

Each CREATE block must specify a unique NAME. The NAME represents the Ethernet interface name that will be registered with the Linux operating system.

5. Start the QLogic VNIC driver and the QLogic VNIC interfaces. Once you have created a configuration file, you can start the VNIC driver and create the VNIC interfaces specified in the configuration file by running the following command (as root):

```
# /etc/init.d/qlgc_vnic start
```

You can stop the VNIC driver and bring down the VNIC interfaces by running the following command:

```
# /etc/init.d/qlgc_vnic stop
```

To restart the QLogic VNIC driver, run the following command:

```
# /etc/init.d/qlgc_vnic restart
```

If you have not started the InfiniBand network stack (InfiniPath or OFED), then running the `/etc/init.d/qlgc_vnic start` command also starts the InfiniBand network stack, since the QLogic VNIC service requires the InfiniBand stack.

If you start the InfiniBand network stack separately, then the correct starting order is:

- Start the InfiniBand stack.
- Start QLogic VNIC service.

For example, if you use InfiniPath, the correct order of starting is:

```
# /etc/init.d/openibd start  
# /etc/init.d/qlgc_vnic start
```

Correct stopping order is:

- Stop QLogic VNIC service.
- Stop the InfiniBand stack.

For example, if you use InfiniPath, the correct stopping order is:

```
# /etc/init.d/qlgc_vnic stop  
# /etc/init.d/openibd stop
```

If you try to stop the InfiniBand stack when the QLogic VNIC service is running, an error message displays, indicating that some of the modules of the InfiniBand stack are in use by the QLogic VNIC service. Also, any QLogic VNIC interfaces that you created are removed (because stopping the InfiniBand network stack unloads the HCA driver, which is required for the VNIC interfaces to be present).

In this case, do the following:

- Stop the QLogic VNIC service with `/etc/init.d/qlgc_vnic stop`.
- Stop the InfiniBand stack again.

If you want to restart the QLogic VNIC interfaces, run the following command:

```
# /etc/init.d/qlgc_vnic restart
```

You can get information about the QLogic VNIC interfaces by using the following script (as root):

```
# ib_qlgc_vnic_info
```

This information is collected from the `/sys/class/infiniband_qlgc_vnic/interfaces/` directory, under which there is a separate directory corresponding to each VNIC interface.

VNIC interfaces can be deleted by writing the name of the interface to the `/sys/class/infiniband_qlgc_vnic/interfaces/delete_vnic` file. For example, to delete interface `veth0`, run the following command (as root):

```
# echo -n veth0 >
/sys/class/infiniband_qlgc_vnic/interfaces/delete_vnic
```

More information for configuration, starting and stopping the interface, and basic troubleshooting is found in the QLogic *OFED+ User Guide*.

MPI over uDAPL

Intel MPI can be run over uDAPL, which uses IB Verbs. uDAPL is the user mode version of the Direct Access Provider Library (DAPL), and is provided as a part of the OFED packages. You will also have to have IPoIB configured.

Setup for Intel MPI is described below.

1. Make sure that DAPL 1.2 (not version 2.0) is installed on every node. In this release they are called `compat-dapl`. (Both versions are supplied with the OpenFabrics RPMs.) They can be installed either with the Installer with the QLogicIB-Basic package or with `rpm` with the QLogic OFED 1.4 RPM set. They look something like this:

```
$ rpm -qa | grep compat-dapl
compat-dapl-1.2.12-1.x86_64.rpm
compat-dapl-debuginfo-1.2.12-1.x86_64.rpm
compat-dapl-devel-1.2.12-1.x86_64.rpm
compat-dapl-devel-static-1.2.12-1.x86_64.rpm
compat-dapl-utils-1.2.12-1.x86_64.rpm
```

2. Check that you have a `/etc/dat.conf` file. It should be installed by the `dapl-RPM`. The file `dat.conf` contains a list of interface adapters supported by uDAPL service providers. In particular, it must contain mapping entries for OpenIB-cma for `dapl 1.2.x`, in a form similar to this (all on one line):

```
OpenIB-cma u1.2 nonthreadsafe default libdaplcma.so.1 dapl.1.2  
"ib0 0" ""
```

3. On every node, type the following command (as root):

```
# modprobe rdma_ucm
```

To ensure that the module is loaded whenever the driver is loaded, add `RDMA_UCM_LOAD=yes` to the `/etc/infiniband/openib.conf` file. (Note that `rdma_cm` is also used, but it is loaded automatically.)

4. Bring up an IPoIB interface on every node, eg. `ib0`. See the instructions for configuring IPoIB for more details.

For more information on using Intel MPI, see the “Using Other MPIs” section in the *QLogic HCA and QLogic OFED Software Users Guide*.

Other Configuration: Changing the MTU Size

The Maximum Transfer Unit (MTU) is set to 4K and enabled in the driver by default. To change the driver default back to 2K MTU, add this line (as root) in `/etc/modprobe.conf` (or in `/etc/modprobe.conf.local` on SLES):

```
options ib_ipath mtu4096=0
```

Restart the driver as described in [“Managing the InfiniPath Driver” on page 6-15](#).

NOTE:

To use 4K MTU, set the switch to have the same 4K default. If you are using QLogic switches the following will apply:

For the Externally Managed 9024, use 4.2.2.0.3 firmware (9024DDR4KMTU_firmware.emfw) for the 9024 EM. This has the 4K MTU default, for use on fabrics where 4K MTU is required. If 4K MTU support is not required, then the 4.2.2.0.2 DDR *.emfw file should be used for DDR externally-managed switches. Use FastFabric to load the firmware on all the 9024s on the fabric.

For the 9000 Chassis, use the most recent 9000 code 4.2.4.0.1. The 4K MTU support is in 9000 Chassis version 4.2.1.0.2 and later. For the 9000 chassis, when the FastFabric 4.3 (or later) chassis setup tool is used, the user is asked what MTU they want. FastFabric (FF) can then set that MTU in all the 9000 Internally managed switches. The change will take effect on next reboot. Alternatively, for the Internally Managed 9000s, the `ismChassisSetMtu` command-line interface (CLI) command can be used. This should be executed on every switch and both hemispheres of the 9240s.

For reference, see the FastFabric Users Guide Version 4.3 and the SilverStorm 9000 CLI Reference Guide Version 4.2. Both are available from the QLogic web site.

For other switches, see the vendors' documentation.

Managing the InfiniPath Driver

The startup script for `ib_ipath` is installed automatically as part of the software installation, and normally does not need to be changed. It runs as a system service.

The primary configuration file for the InfiniPath driver `ib_ipath` and other modules and associated daemons, is:

```
/etc/infiniband/openib.conf
```

Normally, this configuration file is set up correctly at installation and the drivers are loaded automatically during system boot once the RPMs have been installed. However, the `ib_ipath` driver has several configuration variables that set reserved buffers for the software, define events to create trace records, and set the debug level.

If you are upgrading, your existing configuration files will not be overwritten.

The device files are:

```
/dev/ipath  
/dev/ipath0, /dev/ipath1, ...
```

The numbered device files allow access to a specific InfiniPath unit.

See the `ib_ipath` man page for more details.

Configure InfiniPath Driver State

Use the following commands to check or configure the state. These methods will not reboot the system.

To check the configuration state, use this command. You do not need to be root:

```
$ chkconfig --list openibd
```

To enable the driver, use the command (as root):

```
# chkconfig openibd on 2345
```

To disable the driver on the next system boot, use the command (as root):

```
# chkconfig openibd off
```

NOTE:

This command does not stop and unload the driver if the driver is already loaded.

Start, Stop or Restart InfiniPath

Restart the software if you install a new InfiniPath release, change driver options, or do manual testing.

You can start, stop, or restart (as root) the InfiniPath support with:

```
# /etc/init.d/openibd [start | stop | restart]
```

This method will not reboot the system. The following set of commands shows how to use this script. Note the following:

- If OpenSM is configured and running, it must be stopped before the `openibd stop` command, and must be started after the `openibd start` command. Omit the commands to start/stop `opensmd` if you are not running it on that node.

The sequence of commands to restart the driver are as follows.

```
# /etc/init.d/opensmd stop
# /etc/init.d/openibd stop
...
# /etc/init.d/openibd start
# /etc/init.d/opensmd start
```

The ... represents whatever activity you are engaged in after `infinipath` is stopped.

An equivalent way to restart the driver this is to use same sequence as above, except use the `restart` command instead of `start` and `stop`:

```
# /etc/init.d/opensmd stop
# /etc/init.d/openibd restart
# /etc/init.d/opensmd start
```

NOTE:

Stopping or restarting `openibd` terminates any QLogic MPI processes, as well as any OpenFabrics processes that are running at the time.

You can check to see if `opensmd` is running by using the following command (as root); if there is no output, `opensmd` is not configured to run:

```
# /sbin/chkconfig --list opensmd | grep -w on
```

When you need to determine which InfiniPath and OpenFabrics modules are running, use the following command. You do not need to be root:

```
$ lsmod | egrep 'ipath_|ib_|rdma_|findex'
```

Unloading the Driver/Modules Manually

You can also unload the driver/modules manually without using `/etc/init.d/openibd`. Use the following series of commands (as root):

```
# umount /ipathfs
# fuser -k /dev/ipath* /dev/infiniband/*
# lsmod | egrep '^ib_|^rdma_|^iw_' | xargs modprobe -r
```

Further Information on Configuring and Loading Drivers

See the `modprobe(8)`, `modprobe.conf(5)`, and `lsmod(8)` man pages for more information. Also see the file (on Red Hat systems):

```
/usr/share/doc/initscripts-*/sysconfig.txt
```

Notes

Draft

7 Installation Verification and Additional Settings

This section provides instructions for verifying that the software has been properly installed and that the drivers are loaded, and the fabric is active and ready to use. Information on adapter performance tuning is also provided.

LED Link and Data Indicators

The LEDs function as link and data indicators once the InfiniPath software has been installed, the driver has been loaded, and the fabric is being actively managed by a subnet manager.

Adapter and Other Settings

The following settings can be adjusted for better performance.

- **Use `taskset` to tune CPU affinity on Opteron systems with the QLE7240, QLE7280, and QLE7140.** Latency will be slightly lower for the Opteron socket that is closest to the PCI Express bridge. On some chipsets, bandwidth may be higher on this socket. See the *QLogic HCA and QLogic OFED Software Users Guide* for more information on using `taskset`. Also see the `taskset(1)` man page.
- **On the switch use an IB MTU of 4096 bytes instead of 2048 bytes, if available, with the QLE7240, QLE7280, and QLE7140.** 4K MTU is enabled in the InfiniPath driver by default. To change this setting for the driver, see [“Other Configuration: Changing the MTU Size” on page 6-14](#).
- **Use a PCIe Max Read Request size of at least 512 bytes with the QLE7240 and QLE7280.** QLE7240 and QLE7280 adapters can support sizes from 128 bytes to 4096 bytes in powers of two. This value is typically set in the BIOS.
- **Use PCIe Max Payload size of 256, where available, with the QLE7240 and QLE7280.** The QLE7240 and QLE7280 adapters can support 128, 256, or 512 bytes. This value is typically set by the BIOS as the minimum value supported both by the PCIe card and the PCIe root complex.

- **Make sure that write combining is enabled.** The x86 Page Attribute Table (PAT) mechanism that allocates write-combining (WC) mappings for the PIO buffers has been added and is now the default. If PAT is unavailable or PAT initialization fails for some reason, the code will generate a message in the log and fall back to the MTRR mechanism. See [B Write Combining](#) for more information.
- **Check the PCIe bus width.** If slots have a smaller electrical width than mechanical width, lower than expected performance may occur. Use this command to check PCIe Bus width:

```
$ ipath_control -iv
```

This will also show link speed.

Customer Acceptance Utility

`ipath_checkout` is a `bash` script that verifies that the installation is correct and that all the nodes of the network are functioning and mutually connected by the InfiniPath fabric. It must be run on a front end node, and requires specification of a `nodefile`. For example:

```
$ ipath_checkout [options] nodefile
```

The `nodefile` lists the hostnames of the nodes of the cluster, one hostname per line. The format of `nodefile` is as follows:

```
hostname1  
hostname2  
...
```

NOTE:

The hostnames in the `nodefile` are Ethernet hostnames, not IPv4 addresses.

NOTE:

To create a `nodefile` you can use the `ibhosts` program. It will generate a list of available nodes that are already connected to the switch. The `ibhosts` program is described in more detail in the *QLogic HCA and QLogic OFED Software Users Guide*.

`ipath_checkout` performs the following seven tests on the cluster:

1. Executes the `ping` command to all nodes to verify that they all are reachable from the front end.
2. Executes the `ssh` command to each node to verify correct configuration of `ssh`.

3. Gathers and analyzes system configuration from the nodes.
4. Gathers and analyzes RPMs installed on the nodes. Missing RPMs can be found this way.
5. Verifies QLogic hardware and software status and configuration. Includes tests for link speed, PIO bandwidth (incorrect MTRR settings), and MTU size.
6. Verifies the ability to `mpirun` jobs on the nodes.
7. Runs a bandwidth and latency test on every pair of nodes and analyzes the results.

The options available with `ipath_checkout` are shown in [Table 7-1](#).

Table 7-1. `ipath_checkout` Options

Command	Meaning
<code>-h, --help</code>	These options display help messages describing how a command is used.
<code>-v, --verbose</code> <code>-vv, --vverbose</code> <code>-vvv, --vvvverbose</code>	These options specify three successively higher levels of detail in reporting test results. There are four levels of detail in all, including the case where none of these options are given.
<code>-c, --continue</code>	When this option is not specified, the test terminates when any test fails. When specified, the tests continue after a failure, with failing nodes excluded from subsequent tests.
<code>-k, --keep</code>	This option keeps intermediate files that were created while performing tests and compiling reports. Results will be saved in a directory created by <code>mktemp</code> and named <code>infinipath_XXXXXX</code> or in the directory name given to <code>--workdir</code> .
<code>--workdir=DIR</code>	Use <code>DIR</code> to hold intermediate files created while running tests. <code>DIR</code> must not already exist.
<code>--run=LIST</code>	This option runs only the tests in <code>LIST</code> . See the seven tests listed previously. For example, <code>--run=123</code> will run only tests 1, 2, and 3.
<code>--skip=LIST</code>	This option skips the tests in <code>LIST</code> . See the seven tests listed previously. For example, <code>--skip=2457</code> will skip tests 2, 4, 5, and 7.
<code>-d, --debug</code>	This option turns on the <code>-x</code> and <code>-v</code> flags in <code>bash(1)</code> .

In most cases of failure, the script suggests recommended actions. Please see the `ipath_checkout` man page for more information and updates.

Also refer to the Troubleshooting appendix in the *QLogic HCA and QLogic OFED Software Users Guide*.

Draft

A Installation Troubleshooting

The following sections contain information about issues that may occur during installation. Some of this material is repeated in the Troubleshooting appendix of the *QLogic HCA and QLogic OFED Software User Guide*.

Many programs and files are available that gather information about the cluster, and can be helpful for debugging. See appendix D, Useful Programs and Files, in the *QLogic HCA and QLogic OFED Software User Guide*.

Hardware Issues

Some of the hardware issues that may occur during installation are described in the following sections. Use the LEDs, as described in “[LED Link and Data Indicators](#)” on page 7-1, to help diagnose problems.

Node Spontaneously Reboots

If a node repeatedly and spontaneously reboots when attempting to load the InfiniPath driver, it may be because the QLogic adapter is not installed correctly in the HTX or PCI Express slot.

Some HTX Motherboards May Need Two or More CPUs in Use

Some HTX motherboards may require that two or more CPUs be in use for the QLogic adapter to be recognized. This is most evident in four-socket motherboards.

BIOS Settings

This section covers issues related to BIOS settings. You can check and adjust BIOS settings using the BIOS Setup utility. For specific instructions, follow the hardware documentation that came with your system.

Enable Advanced Configuration and Power Interface (ACPI)

This setting must be enabled. If ACPI is disabled, it may cause initialization problems, as described in the Troubleshooting section of the *QLogic HCA and QLogic OFED Software User Guide*.

Issue with Supermicro® H8DCE-HTe and QHT7040

The QLogic adapter may not be recognized at startup when using the Supermicro H8DCE-HT-e and the QHT7040 adapter. To fix this problem, set the operating system selector option in the BIOS for Linux. The option will look like:

```
OS Installation [Linux]
```

Software Installation Issues

Some problems can be found by running `ipath_checkout`. Run `ipath_checkout` before contacting technical support.

Missing Kernel RPM Errors

Install the `kernel-source`, `kernel-devel`, and, if using an older release, `kernel-smp-devel` RPMs for your distribution before installing the InfiniPath RPMs, as there are dependencies. "" on page 5-4 lists all the required packages.

Use `uname -a` to find out which kernel is currently running, to make sure that you install the version with which it matches.

If these RPMs have not been installed, you will see error messages like this when installing InfiniPath:

```
Building and installing InfiniPath modules for 2.6.16_sles10
2.6.16.21-0.8-debug kernel
*** ERROR: /lib/modules/2.6.16.21-0.8-debug/build/.config is
missing.
***          Is the kernel-source rpm for 2.6.16.21-0.8-debug
installed?
=====

=====

Building and installing InfiniPath modules for 2.6.9_U4
2.6.9-42.ELsmp kernel
*** ERROR: /lib/modules/2.6.9-42.ELsmp/build/.config is missing.
***          Is the kernel-smp-devel rpm for 2.6.9-42.ELsmp
installed?
=====

.
.
.
```

Install the correct RPMs by using the `yum` or `yast` commands, for example:

```
# yum install kernel-devel
```

NOTE:

Check your distribution's documentation for more information on installing these RPMs, and for usage of `yum` or `yast`.

Next, the `kernel-ib` package must be re-installed, with the `--replacepks` option included. Then the `infinipath` service can be restarted. To do so, type the following (as root):

```
# rpm -U --replacepks kernel-ib*
# /etc/init.d/openibd restart
```

Resolving Conflicts

Occasionally, conflicts may arise when trying to install "on top of" an existing set of files that may come from a different set of RPMs. For example, if you install the QLogic MPI RPMs after having previously installed Local Area Multicomputer (LAM)/MPI, there will be conflicts, since both installations have versions of some of the same programs and documentation. You would see an error message similar to the following:

```
# rpm -Uvh Documentation/*rpm InfiniPath/*rpm
InfiniPath-Devel/*rpm InfiniPath-MPI/*rpm OpenFabrics/*rpm
OpenFabrics-Devel/*rpm OpenSM/*rpm
Preparing...
##### [100%]
file /usr/share/man/man3/MPIO_Request_c2f.3.gz from install of
mpi-doc-4321.776_rhel4_psc conflicts with file from package
lam-7.1.2-8.fc6
```

Use the following command to remove previously installed conflicting packages. This command will remove all the available LAM packages:

```
# rpm -e --allmatches lam lam-devel lam-libs
```

After the packages have been removed, continue with the InfiniPath installation.

You can also use the `--prefix` option with the `rpm` command to relocate the install directory of any packages that you need to move. See [“Using rpm to Install InfiniPath and OpenFabrics” on page 5-14](#) for more information.

openmpi_gcc Fails to Install Because of Dependency on gfortran (RHEL 4)

On RHEL 4 distributions, `libgfortran` must be installed before installing the QLogic `openmpi_gcc*` RPM, otherwise the installation will fail. The `libgfortran*` RPM is available as part of the RHEL 4 distribution.

mpirun Installation Requires 32-bit Support

On a 64-bit system, 32-bit `glibc` must be installed before installing the `mpi-frontend-*` RPM. `mpirun`, which is part of the `mpi-frontend-*` RPM, requires 32-bit support.

If 32-bit `glibc` is not installed on a 64-bit system, an error like this displays when installing `mpi-frontend`:

```
# rpm -Uv ~/tmp/mpi-frontend-2.3-14729.802_rhel4_qlc.i386.rpm
error: Failed dependencies:
/lib/libc.so.6 is needed by mpi-frontend-2.3
-14729.802_rhel4_qlc.i386
```

In older distributions, such as RHEL4, the 32-bit `glibc` is contained in the `libgcc` RPM. The RPM name will be similar to this:

```
libgcc-<version>.EL4.i386.rpm
```

In newer distributions, `glibc` is an RPM name. The 32-bit `glibc` is named similarly to:

```
glibc-<version>.i686.rpm OR  
glibc-<version>.i386.rpm
```

Check your distribution for the exact RPM name.

Lockable Memory Error on Initial Installation of InfiniPath

During the first installation of InfiniPath software, `/etc/initscript` is created or modified to increase the amount of lockable memory (up to 128 MB) for normal users. This change will not take effect until the system is rebooted, and jobs may fail with error messages about locking memory or failing `mmap`. This error is described in the QLogic MPI Troubleshooting section “Lock Enough Memory on Nodes When Using a Batch Queuing System” in the *QLogic HCA and QLogic OFED Software User Guide*.

This is not an issue when upgrading to a newer version of the InfiniPath software.

Configuration Issues

`ibsrpdm` Command Hangs When Two HCAs are Installed but Only Unit 1 is Connected to the Switch

If multiple HCAs (unit 0 and unit 1) are installed, and only unit 1 is connected to the switch, the `ibsrpdm` (to set up an SRP target) command can hang. If unit 0 is connected, and unit 1 is disconnected, the problem does not occur.

When only unit 1 is connected to the switch, use the `-d` option with `ibsrpdm`, then, using the output from the `ibsrpdm` command, echo the new target info into `/sys/class/infiniband_srp/srp-ipath1-1/add_target`.

For example:

```
# ibsrpdm -d /dev/infiniband/umad1 -c  
# echo \  
id_ext=21000001ff040bf6,ioc_guid=21000001ff040bf6,dgid=fe80000000  
0000021000001ff040bf6,pkey=ffff,service_id=f60b04ff01000021 >  
/sys/class/infiniband_srp/srp-ipath1-1/add_target
```

Outdated `ipath_ether` Configuration Setup Generates Error

Ethernet emulation (`ipath_ether`) has been removed in this release, and, as a result, an error may be seen if the user still has an alias set previously by `modprobe.conf` (for example, `alias eth2 ipath_ether`).

When `ifconfig` or `ifup` are run, the error will look similar to this (assuming `ipath_ether` was used for `eth2`):

```
eth2: error fetching interface information: Device not found
```

To prevent the error message, remove the following files (assuming `ipath_ether` was used for `eth2`):

```
/etc/sysconfig/network-scripts/ifcfg-eth2 (for RHEL)
```

```
/etc/sysconfig/network/ifcfg-eth-eth2 (for SLES)
```

QLogic recommends using the IP over InfiniBand protocol (IPoIB-CM), included in the standard OpenFabrics software releases, as a replacement for `ipath_ether`.

Draft

B Write Combining

Introduction

Write combining improves write bandwidth to the QLogic chip by writing multiple words in a single bus transaction (typically 64 bytes). Write combining applies only to x86_64 systems.

The x86 Page Attribute Table (PAT) mechanism that allocates write-combining (WC) mappings for the PIO buffers has been added and is now the default.

If PAT is unavailable or PAT initialization fails for some reason, the code will generate a message in the log and fall back to the Memory Type Range Registers (MTRR) mechanism.

If write combining is not working properly, lower than expected bandwidth may occur.

Instructions for checking write combining and for using PAT and MTRR are given below.

Verify Write Combining is Working

To see if write combining is working correctly and to check the bandwidth, run the following command:

```
$ ipath_pkt_test -B
```

With write combining enabled, the QLE7140 and QLE7240 report in the range of 1150–1500 MB/s; the QLE7280 reports in the range of 1950–3000 MB/s. The QHT7040/7140 adapters normally report in the range of 2300–2650 MB/s.

You can also use `ipath_checkout` (use option 5) to check bandwidth.

Although the PAT mechanism should work correctly by default, increased latency and low bandwidth may signal a problem. If so, the interconnect operates, but in a degraded performance mode, with latency increasing to several microseconds, and bandwidth decreasing to as little as 200 MB/s.

Upon driver startup, you may see these errors:

```
ib_ipath 0000:04:01.0: infinipath0: Performance problem: bandwidth  
to PIO buffers is only 273 MiB/sec
```

·
·

If you do not see any of these messages on your console, but suspect this problem, check the `/var/log/messages` file. Some systems suppress driver load messages but still output them to the log file.

Methods for enabling and disabling the two write combining mechanisms are given below. There are no conflicts between the two methods.

PAT and Write Combining

This is the default mechanism for allocating write-combining (WC) mappings for the PIO buffers. It is set as a parameter in `/etc/modprobe.conf` (on Red Hat systems) or `/etc/modprobe.conf.local` (on SLES systems). This is the default:

```
option ib_ipath wc_pat=1
```

If PAT is unavailable or PAT initialization fails for some reason, the code will generate a message in the log and fall back to the Memory Type Range Registers (MTRR) mechanism. To use MTRR, disable PAT by setting this module parameter to 0 (as root):

```
option ib_ipath wc_pat=0
```

Then, revert to using the MTRR-only behavior by following one of the two suggestions in [“MTRR Mapping and Write Combining” on page B-2](#), below.

The driver will need to be restarted after the changes have been made.

NOTE:

There will be no write-combining entry in `/proc/mtrr` when using PAT.

MTRR Mapping and Write Combining

Two suggestions for using MTRR mapping to fix this problem are described below.

See the Troubleshooting section of the *QLogic HCA and QLogic OFED Software User Guide* for more details on a related performance issue.

Edit BIOS Settings to Fix MTRR Issues

You can edit the BIOS setting for MTRR Mapping. The BIOS setting looks similar to:

```
MTRR Mapping          [Discrete]
```

For systems with very large amounts of memory (32GB or more), it may also be necessary to adjust the BIOS setting for the *PCI hole granularity* to 2GB. This setting allows the memory to be mapped with fewer MTRRs, so that there will be one or more unused MTRRs for the InfiniPath driver.

Some BIOS' do not have the MTRR mapping option. It may have a different name, depending on the chipset, vendor, BIOS, or other factors. For example, it is sometimes referred to as *32 bit memory hole*. This setting must be enabled.

If there is no setting for MTRR mapping or 32 bit memory hole, and you have problems with degraded performance, contact your system or motherboard vendor and ask how to enable write combining.

Use the `ipath_mtrr` Script to Fix MTRR Issues

QLogic also provides a script, `ipath_mtrr`, which sets the MTRR registers, enabling maximum performance from the InfiniPath driver. This Python script is available as a part of the InfiniPath software download, and is contained in the `infinipath*` RPM. It is installed in `/bin`.

To diagnose the machine, run it with no arguments (as root):

```
# ipath_mtrr
```

The test results will list any problems, if they exist, and provide suggestions on what to do.

To fix the MTRR registers, use:

```
# ipath_mtrr -w
```

Restart the driver after fixing the registers.

This script needs to be run after each system reboot. It can be set to run automatically upon restart by adding this line in `/etc/sysconfig/infinipath`:

```
IPATH_MTRR_ACTIVE=1
```

See the `ipath_mtrr(8)` man page for more information on other options.

Notes

Draft

C Configuration Files

Table C-1 contains descriptions of the configuration and configuration template files used by the InfiniPath and OpenFabrics software.

Table C-1. Configuration Files

Configuration File Name	Description
<code>/etc/infiniband/qlgc_vnic.cfg</code>	VirtualNIC configuration file. Create this file after running <code>ib_qlgc_vnic_query</code> to get the information you need. This file was named <code>/etc/infiniband/qlogic_vnic.cfg</code> or <code>/etc/sysconfig/ics_inic.cfg</code> in previous releases. See the sample file <code>qlgc_vnic.cfg.sample</code> (described below) to see how it should be set up.
<code>/etc/modprobe.conf</code>	Specifies options for modules when added or removed by the <code>modprobe</code> command. Also used for creating aliases. PAT write-combing option is set here. For Red Hat systems.
<code>/etc/modprobe.conf.local</code>	Specifies options for modules when added or removed by the <code>modprobe</code> command. Also used for creating aliases. PAT write-combing option is set here. For SLES systems.
<code>/etc/infiniband/openib.conf</code>	The primary configuration file for InfiniPath, OFED modules, and other modules and associated daemons. Automatically loads additional modules or changes IPOIB transport type.
<code>/etc/sysconfig/infinipath</code>	Contains settings, including the one that sets the <code>ipath_mtrr</code> script to run on reboot.

Table C-1. Configuration Files (Continued)

Configuration File Name	Description
<code>/etc/sysconfig/network/ifcfg- <NAME></code>	<p>Network configuration file for network interfaces</p> <p>When used for VNIC configuration, <NAME> is in the form <code>eiocX</code>, where <code>X</code> is the device number. There will be one interface configuration file for each interface defined in <code>/etc/infiniband/qlgc_vnic.cfg</code>.</p> <p>For SLES systems.</p>
<code>/etc/sysconfig/net- work-scripts/ifcfg-<NAME></code>	<p>Network configuration file for network interfaces</p> <p>When used for VNIC configuration, <NAME> is in the form <code>eiocX</code>, where <code>X</code> is the device number. There will be one interface configuration file for each interface defined in <code>/etc/infiniband/qlgc_vnic.cfg</code>.</p> <p>For Red Hat systems.</p>
Sample and Template Files	Description
<code>qlgc_vnic.cfg.sample</code>	<p>Sample VNIC config file. It can be found with the OFED documentation, or in the <code>qlgc_vnictools</code> subdirectory of the <code>QLogicIB_Basic</code> download. It is also installed in <code>/etc/infiniband</code>.</p>
<code>/usr/share/doc/initscripts-*/ sysconfig.txt</code>	<p>File that explains many of the entries in the configuration files</p> <p>For Red hat systems</p>

D Package Descriptions

The following sections contain detailed descriptions of the packages for the InfiniPath and OpenFabrics software. In this release, software may be installed in several ways. With the QLogicIB-Basic download, which uses the Installer tool, the package groupings are a bit different than those in the RPM-based downloads.

Package Names with the QLogicIB-Basic Download

The QLogic IB-Basic download has been packaged differently than the RPM download. Some but not all of the underlying packages are RPMs. The Installer tool groups the components in this way:

```
ib_stack, truescale, mvapich_gcc_qlc, mvapich_pgi_qlc,  
mvapich_pathscale_qlc, mvapich_intel_qlc, openmpi_gcc_qlc,  
openmpi_pgi_qlc, openmpi_pathscale_qlc, openmpi_intel_qlc, oftools,  
ib_stack_dev, fastfabric, qlgc_srp, qlgc_vnic, qlgc_fm,  
ofed_ipoib, ofed_sdp, ofed_udapl, mvapich, mvapich2, openmpi,  
ofed_mpirsrc, ofed_rds, ofed_srp, ofed_srpt, ofed_iser, ofed_isert,  
ofed_iwarp, opensm, ofed_debug
```

Once the Installer tool has begun the installation, for some packages, the underlying RPM names are displayed on the screen.

Different Nodes May Use Different RPMs

In a cluster environment, different nodes may be used for different functions, such as launching jobs, software development, or running jobs. These nodes are defined as follows:

- **Front end node.** This node launches jobs. It is referred to as the *front end node* throughout this document.
- **Compute node.** These nodes run jobs.
- **Development or build node.** These are the machines on which examples or benchmarks can be compiled.

Any machine can serve any combination of these three purposes, but a typical cluster has many compute nodes and just a few (or only one) front end nodes. The number of nodes used for development will vary.

Although QLogic recommends installing all RPMs on all nodes, not all InfiniPath software is required on all nodes. See [Table D-2](#), [Table D-3](#), or [Table D-4](#) for information on installation of software RPMs on specific types of nodes.

InfiniPath RPM Version Numbers and Identifiers

The InfiniPath RPMs that are shipped have the InfiniPath release number, build identifiers, and distribution contained in the RPM name. The architecture is designated by `x86_64`, `noarch`, or `i386`, and is dependent upon the distribution. For example:

```
infinipath-<release>-<build_identifier>_<distro>_qlc.x86_64.rpm
```

Note that `_qlc` is always added after the distribution identifier.

In the InfiniPath tables in this appendix, the release and build identifiers are contained in `xxx` and the distribution identifier plus `_qlc` is contained in `yyy`. Using this convention, the previous RPM would be listed as:

```
infinipath-xxx_yyy.x86_64.rpm
```

OpenFabrics RPM Names

Non-InfiniPath components have their own version numbering, which can vary dependent upon the source of the RPM. For example

```
mvapich_gcc-1.1.0-3143.x86_64.rpm
```

`1_1_0` is the 1.1.0 build for `mvapich`. In the following tables the version number is replaced by `xxx`.

InfiniPath and OpenFabrics RPMs

QLogic recommends installing all RPMs (with the exception of the `*-Static/*`, `*-Debuginfo/*`, and `*32bit/*`) on all nodes, if using the `rpm` method.

Some RPMs are optional. Since cluster nodes can be used for different functions, it is possible to selectively install RPMs. For example, you can install the `opensm` package for use on the node that will act as a subnet manager. If you want to selectively install the RPMs, see the following tables for a comparison of required and optional packages.

To generate a list of the InfiniPath software package contents on each RPM, type:

```
$ rpm -qlp rpm_file_name
```

Documentation RPMs

Table D-1. Documentation/RPMs

RPM Name	Front End	Compute	Development
infinipath-doc-xxx_yyy.noarch.rpm InfiniPath man pages	Optional	Optional	Optional
mpi-doc-xxx_yyy.noarch.rpm Man pages for MPI functions and other MPI documents	Optional	Optional	Optional
ofed-docs-xxx.x86_64.rpm OpenFabrics documentation	Optional	Optional	Optional

InfiniPath RPMs

Table D-2. InfiniPath/RPMs

RPM Name	Front End	Compute	Development
infinipath-xxx_yyy.x86_64.rpm Utilities and InfiniPath configuration files Contains ipath_checkout and ipathbug-helper ^a	Optional	Required	Optional
kernel-ib-xxx_yyy.x86_64.rpm InfiniPath Drivers and OpenFabrics kernel modules.	Optional	Required	Optional
infini-path-libs-xxx_yyy.x86_64.rpm infini-path-libs-xxx_yyy.i386.rpm InfiniPath protocol shared libraries for 32-bit and 64-bit systems	Optional	Required	Optional

Table Notes

^a If you want to use ipath_checkout and ipathbug-helper, install this RPM wherever you install mpi-frontend.

Table D-3. InfiniPath-Devel/RPMs

RPM Name	Front End	Compute	Development
infini-path-devel-xxx_yyy.noarch.rpm Development files for InfiniPath	Optional	Optional	Optional
mpi-devel-xxx_yyy.noarch.rpm Source code for the MPI development environment, including headers and libs, MPI examples and benchmarks. Use to build the examples or rebuild the benchmarks.	Optional	Optional	Required

Table D-4. InfiniPath-MPI/RPMs

RPM Name	Front End	Compute	Development
mpi-bench-mark-xxx_yyy.x86_64.rpm MPI benchmark binaries	Optional	Required	Optional
mpi-frontend-xxx_yyy.i386.rpm MPI job launch scripts and binaries, including mpirun and MPD	Required	Required	Optional
mpi-libs-xxx_yyy.i386.rpm mpi-libs-xxx_yyy.x86_64.rpm Shared libraries for MPI	Optional	Required	Required

OpenFabrics RPMs

For ease of installation, QLogic recommends that all of the OpenFabrics RPMs listed in [Table D-5](#) be installed on all nodes. Version numbers are indicated by *xxx*.

Table D-5. OpenFabrics/RPMs

RPM Name	Comments
compat-dapl-xxx.x86_64.rpm uDAPL 1.2.12 support	Optional for OpenFabrics
compat-dapl-utils-xxx.x86_64.rpm uDAPL 1.2.12 support	Optional for OpenFabrics

Table D-5. OpenFabrics/RPMs (Continued)

RPM Name	Comments
dapl-xxx.x86_64.rpm uDAPL 2.0.15 support	Optional for OpenFabrics
dapl-utils-xxx.x86_64.rpm uDAPL support	Optional for OpenFabrics
ibsim-xxx.x86_64.rpm InfiniBand Fabric Simulator	Optional for OpenFabrics
ibutils-xxx.x86_64.rpm ibutils provides InfiniBand (IB) network and path diagnostics.	Optional for OpenFabrics
ibvexdmtools-xxx.x86_64.rpm Discover and use QLogic Virtual NIC devices via VNIC protocol over InfiniBand	Optional for OpenFabrics
infiniband-diags-xxx.x86_64.rpm Diagnostic tools	Optional for OpenFabrics
iscsi-initiator-utils-xxx.x86_64.rpm ^a Server daemon and utility programs for iSCSI. Also iSER support For Red Hat systems	Optional for OpenFabrics
libibcm-xxx.x86_64.rpm Along with the OpenFabrics kernel drivers, libibcm provides a user-space IB connection management API.	Optional for OpenFabrics
libibcommon-xxx.x86_64.rpm Common utility functions for IB diagnostic and management tools	Required for OpenSM
libibmad-xxx.x86_64.rpm Low-layer IB functions for use by the IB diagnostic and management programs. These include management datagrams (MADs), SA, SMP, and other basic IB functions.	Required for OpenSM
libibumad-xxx.x86_64.rpm Provides the user MAD library functions that sit on top of the user MAD modules in the kernel. These functions are used by IB diagnostic and management tools, including OpenSM.	Required for OpenSM

Table D-5. OpenFabrics/RPMs (Continued)

RPM Name	Comments
<p>libibverbs-xxx.x86_64.rpm</p> <p>Library that allows userspace processes to use InfiniBand <i>verbs</i> as described in the <i>InfiniBand Architecture Specification</i>. This library includes direct hardware access for fast path operations. For this library to be useful, a device-specific plug-in module must also be installed.</p>	<p>Required for OpenFabrics</p>
<p>libibverbs-utils-xxx.x86_64.rpm</p> <p>Useful libibverbs example programs such as <code>ibv_devinfo</code>, which displays information about IB devices</p>	<p>Required for OpenFabrics</p>
<p>libipathverbs-xxx.x86_64.rpm</p> <p>Provides device-specific userspace driver for QLogic HCAs</p>	<p>Required for OpenFabrics</p>
<p>librdmacm-xxx.x86_64.rpm</p> <p>Support for the new connection manager</p>	<p>Optional for OpenFabrics</p>
<p>librdmacm-utils-xxx.x86_64.rpm</p> <p>Utilities for the new connection manager</p>	<p>Optional for OpenFabrics</p>
<p>libsdp-xxx.x86_64.rpm</p> <p>Can be LD_PRELOAD-ed to have a sockets application use IB Sockets Direct Protocol (SDP) instead of TCP, transparently and without recompiling the application</p>	<p>Required for OpenFabrics</p>
<p>ofed-scripts-xxx.x86_64.rpm</p> <p>OpenFabrics scripts</p>	<p>Optional for OpenFabrics</p>
<p>openib-diags-xxx.x86_64.rpm</p> <p>Useful programs for troubleshooting and checking the state of the adapter, IB fabric, and its components</p>	<p>Optional for OpenFabrics</p>
<p>open-iscsi-xxx.x86_64.rpm^a</p> <p>Transport independent, multi-platform implementation of RFC3720 iSCSI with iSER support</p> <p>For SLES systems</p>	<p>Optional for OpenFabrics</p>
<p>opensm-libs-xxx.x86_64.rpm</p> <p>Provides the library for OpenSM</p>	<p>Required for OpenSM</p>
<p>perftest-xxx.x86_64.rpm</p> <p>IB performance tests</p>	<p>Optional for OpenFabrics</p>

Table D-5. OpenFabrics/RPMs (Continued)

RPM Name	Comments
qlgc_vnic_daemon-xxx.x86_64.rpm Used with VNIC ULP service	Optional for OpenFabrics
qlvnictools-xxx.x86_64.rpm Startup script, sample config file, and utilities	Optional for OpenFabrics
qperf-xxx.x86_64.rpm IB performance tests	Optional for OpenFabrics
rds-tools-xxx.x86_64.rpm Supports RDS	Optional for OpenFabrics
scsi-target-utils-xxx.x86_64.rpm Contains the daemon and tools to setup SCSI (SRP) targets	Optional for OpenFabrics
sdpNetstat-xxx.x86_64.rpm Provides network statistics for SDP	Optional for OpenFabrics
srptools-xxx.x86_64.rpm Support for SRP	Optional for OpenFabrics
tgt-xxx.x86_64.rpm Used for setting up SCSI (SRP) targets	Optional for OpenFabrics

Table Notes

There are two versions of the `dapl*` packages: version 1_2_5 and version 2_0_7. QLogic recommends installing the 1_2_5 version for compatibility with most `dapl` applications.

^a `iscsi-initiator-utils-` and `open-iscsi-` are essentially the same, except that the former is for Red Hat and the latter is for SLES.

The development RPMs in [Table D-6](#) are only needed on the nodes where OFED programs are compiled.

Table D-6. OpenFabrics-Devel/RPMs

RPM Name	Comments
compat-dapl-devel-xxx.x86_64.rpm Development files for uDAPL support	Optional for OpenFabrics
dapl-devel-xxx.x86_64.rpm Development files for uDAPL support	Optional for OpenFabrics

Table D-6. OpenFabrics-Devel/RPMs (Continued)

RPM Name	Comments
libibcm-devel-xxx.x86_64.rpm Development files for the libibcm library	Optional for OpenFabric
libibcommon-devel-xxx.x86_64.rpm Development files for the libibcommon library	Optional for OpenFabrics
libibmad-devel-xxx.x86_64.rpm Development files for the libibmad library	Optional for OpenFabrics
libibumad-devel-xxx.x86_64.rpm Development files for the libibumad library	Optional for OpenFabrics
libibverbs-devel-xxx.x86_64.rpm Libraries and header files for the libibverbs verbs library	Optional for OpenFabrics
libipathverbs-devel-xxx.x86_64.rpm Libraries and header files for the libibverbs verbs library	Optional for OpenFabrics
librdmacm-devel-2.1-xxx.x86_64.rpm Development files for the new connection manager	Optional for OpenFabrics
libsdp-devel-xxx.x86_64.rpm Can be LD_PRELOAD-ed to have a sockets application use Sockets Direct Protocol (SDP) instead of TCP	Optional for OpenFabrics

The `opensm` package in [Table D-7](#) should be installed only on the node that will be used as a subnet manager.

Table D-7. OpenSM/RPM

RPM Name	Comments
opensm-xxx.x86_64.rpm OpenSM provides an implementation of an InfiniBand subnet manager and administrator. At least one per each InfiniBand subnet is required to initialize the InfiniBand hardware.	Required for OpenSM

Table D-8. OpenSM-Devel/RPM

RPM Name	Comments
opensm-devel-xxx.x86_64.rpm Development files for OpenSM	Optional for OpenSM

Other HCAs

The packages in [Table D-9](#) should be installed only if other HCAs are used.

Table D-9. Other HCAs/RPMs

RPM Name	Comments
libcxgb3-xxx.x86_64.rpm Support for the Chelsio 10GbE HCA	Optional for OpenFabrics
libmlx4-xxx.x86_64.rpm Userspace driver for Mellanox® ConnectX™ InfiniBand HCAs	Optional for OpenFabrics
libmthca-xxx.x86_64.rpm Provides a device-specific userspace driver for Mellanox HCAs for use with the libibverbs library	Optional for OpenFabrics
libnes-xxx.x86_64.rpm Provides a userspace driver for NetEffect RNICs for use with the libibverbs library	Optional for OpenFabrics
mstflint-xxx.1x86_64.rpm Firmware update tool for other HCAs	Optional for OpenFabrics
tvflash-xxx.x86_64.rpm Query and update the firmware flash memory for other HCAs	Optional for OpenFabrics

The development RPMs in [Table D-10](#) are only needed on the nodes where programs are compiled

Table D-10. Other HCAs-Devel/RPMs

RPM Name	Comments
libcxgb3-devel-xxx.x86_64.rpm Development files for the Chelsio 10GbE HCA	Optional for OpenFabrics

Table D-10. Other HCAs-Devel/RPMs (Continued)

RPM Name	Comments
libmlx4-devel-xxx.x86_64.rpm Development files for Mellanox® ConnectX™ InfiniBand HCAs	Optional for OpenFabrics

Other MPIs

The packages in [Table D-11](#) should be installed if other MPI implementations are desired. RPMs that include *yyy* in the name are supplied by QLogic; others are from OFED.

Table D-11. Other MPIs/RPMs

RPM Name	Comments
mpi-selector-xxx.x86_64.rpm Tool to select MPI compiled with different compilers	Optional
mpitests_mvapich_gcc-xxx.x86_64.rpm MVAPICH MPI tests compiled with GNU	Optional
mpitests_mvapich2_gcc-xxx.x86_64.rpm MVAPICH2 MPI tests compiled with GNU	Optional
mpitests_openmpi_gcc-xxx.x86_64.rpm Open MPI MPI tests compiled with GNU	Optional
mvapich_gcc-xxx.x86_64.rpm MVAPICH compiled with GNU	Optional
mvapich2_gcc-xxx.x86_64.rpm MVAPICH2 compiled with GNU	Optional
mvapich_gcc_qlc-xxx.yyy.x86_64.rpm MVAPICH compiled with GNU for PSM support	Optional
mvapich_intel_qlc-xxx.yyy.x86_64.rpm ^a MVAPICH compiled with Intel for PSM support	Optional
mvapich_pathscale_qlc-xxx.yyy.x86_64.rpm MVAPICH compiled with PathScale for PSM support	Optional
mvapich_pgi_qlc-xxx.yyy.x86_64.rpm MVAPICH compiled with PGI for PSM support	Optional

Table D-11. OtherMPIs/RPMs (Continued)

RPM Name	Comments
openmpi_gcc-xxx.x86_64.rpm Open MPI compiled with GNU	Optional
openmpi_gcc_qlc-xxx.yyy.x86_64.rpm Open MPI compiled with GNU for PSM support	Optional
openmpi_intel_qlc-xxx.yyy.x86_64.rpm ^a Open MPI compiled with Intel for PSM support	Optional
openmpi_pathscale_qlc-xxx.yyy.x86_64.rpm Open MPI compiled with PathScale for PSM support	Optional
openmpi_pgi_qlc-xxx.yyy.x86_64.rpm Open MPI compiled with PGI for PSM support	Optional
qlogic-mpi-register-xxx.yyy.noarch.rpm ^a Helps QLogic MPI interoperate with other MPIs through the mpi-selector utility	Optional

Table Notes

Filenames with `_qlc` after the compiler name denote QLogic-built versions which enable PSM. The compiler versions used are: GNU 4.1, PathScale 3.0, Intel icc 11.0 (Version 11.0, Build 20081105, Package ID: `l_cproc_p_11.0.074`), and PGI 7.2-5.

^a The `mpi-devel` and `infinipath-devel` RPMs will be installed when the `qlogic-mpi-register` RPM is installed, as there are dependencies

Notes

Draft

Index

A

ACPI 4-4, A-2
Adapter, *see* HCA

B

BIOS
 configuring 4-4
 settings A-2
 settings to fix MTRR issues B-2

C

-c 7-3
Cables supported 4-3
Compiler support 5-3
Configuration
 files C-1
 ib_ipath 6-15
 OpenSM 6-3
 VNIC 6-6
--continue 7-3
CPUs, HTX motherboards may require two or more CPUs A-1

D

-d 7-3
--debug 7-3
Distributions supported 5-1
Document conventions 1-3
Documentation for InfiniPath 1-4
Downloading software files 5-4

Driver

 starting, stopping, and restarting 6-16
 unloading manually 6-17

Drivers

 configuration and loading information 6-17
 InfiniPath and OpenFabrics overview 6-1
 IPoB configuration 6-1, 6-2
 rebuilding on an unsupported distribution/kernel pair 5-19
 rebuilding or reinstalling after a kernel upgrade 5-19

E

Error

 /lib/modules/2.6.16.21-0.8-debug/build/.config is missing error message A-3
 Failed dependencies error message A-4

F

 file/usr/share/man/man3/MPIO_Request_c2f.3.gz from install of mpi-doc-2.1-4321.776_rhel4_psc conflicts with file from package lam-7.1.2-8.fc6 error message A-4

Form factors for HCAs 4-2

H

-h 7-3
Hardware installation overview 3-1

Hardware requirements 4-1

HCA

- form factors 4-2
- other HCAs-Devel/RPMs D-9
- model numbers 2-1
- not recognized A-1
- other HCAs/RPMs D-9
- QHT7140 installation with HTX riser 4-12
- QHT7140 installation without an HTX riser 4-16
- QLE7140/7240/7280 installation with PCI Express riser 4-9
- QLE7140/7240/7280 installation without a PCI Express riser 4-15
- supported 1-1
- tuning for performance 7-1
- help 7-3
- HTX motherboards may required two or more CPUs A-1

I

ib_ipath

- module 6-1
- 0000:04:01.0:infinipath0:Performance problem: error message B-1
- configuration 6-15
- IBA6110 4-8
- IBA6120 4-7
- IBA7220 4-7

InfiniPath

- see *also* InfiniPath software
- cluster 1-2
- devel/RPMs D-4
- documentation 1-4
- documentation and InfiniPath RPMs D-3
- downgrading RPMs 5-25
- environment 5-3
- interconnect overview 1-2
- MPI/RPMs D-4
- OpenFabrics interoperability 1-3
- restarting the service 6-16
- scripts, using to start, stop, or restart drivers 6-16

starting the service 6-1

- tar file, unpacking 5-7, 5-14, 5-17
- uninstalling RPMs 5-24
- using rpm to install 5-14
- using scripts to start, stop, or restart drivers 6-16

InfiniPath software

- list of 2-4
- starting, stopping, and restarting 6-16

Installation

- hardware overview 3-1
- hardware requirements for 4-1–4-2
- Lustre 5-23
- QHT7140 with HTX riser 4-12
- QHT7140 without an HTX riser 4-16
- QLE7140/7240/7280 with PCI Express riser 4-9
- QLE7140/7240/7280 without a PCI Express riser 4-15
- QLogic MPI in an alternate location 5-15, 5-18
- using Rocks for 5-20–5-22
- software overview 3-2
- verification of 7-2
- Interoperability, InfiniPath OpenFabrics 1-3
- ipath
 - checkout 7-2, A-2
 - checkout options 7-3
 - mtrr script to fix MTRR issues B-3
- IPoIB driver configuration 6-1, 6-2

K

- k 7-3
- keep 7-3
- Kernel
 - supported 5-1
- Kernel, missing kernel RPM errors A-2

L

- LEDs, blink patterns 7-1
- Linux, supported distributions 5-1

Lockable memory error A-5
Lustre, installing 5-23

M

Model numbers for HCAs 2-1
MPI, other MPIs/RPMs D-10
MPI over uDAPL 6-13
mpirun, installation requires 32-bit support
 A-4
mpi-selector 5-15, 5-18
MTRR
 editing BIOS settings to fix B-2
 mapping and write combining B-2
 using ipath_mtrr script to fix issues B-3
MTU, changing the size 6-14

N

Node repeatedly reboots error A-1
Node types 5-1

O

OFED software, uninstalling 5-24, 5-25
OFED SRP 6-4
OpenFabrics
 devel/RPMs D-7
 configuration 6-1
 downgrading RPMs 5-25
 environment 5-3
 interoperability with InfiniPath 1-3
 RPMs D-4
 RPMs, uninstalling 5-24
 using rpm to install 5-14
OpenSM 6-3
OpenSM/RPMs D-8, D-9
openssh 5-3
openssh-server 5-3
Optical media converters 4-4

P

Package contents 4-5
Performance tuning 7-1
Protocols supported 2-4
Protocols, InfiniBand subnet management 1-3

Q

QHT7040 4-8
QHT7140 4-8
 installation with HTX riser 4-12
 installation without an HTX riser 4-16
QLE7140 4-7
QLE7140/7240/7280
 installation with PCI Express riser 4-9
 installation without a PCI Express riser 4-15
QLE7280 4-7
QLogic MPI, installing in an alternate location
 5-15, 5-18
QLogic OFED software
 installed layout 5-23

R

RD 2-4
Rocks, managing and installing software with
 5-20
RPMs
 choosing which to install 5-14, 5-17
 different sets on the same system A-4
 downgrading 5-25
 other HCAs D-9
 other HCAs-Devel D-9
 InfiniPath-devel D-4
 InfiniPath-MPI D-4
 list of InfiniPath 5-16
 missing 7-3
 missing kernel RPM errors A-2
 other MPIs D-10
 OpenFabric D-4
 OpenFabrics-devel D-7
 OpenSM D-8, D-9

using to install InfiniPath and OpenFabrics
5-14
version numbers and identifiers D-2
--run=LIST 7-3

S

Safety precautions 4-5
--skip=LIST 7-3

Software

files to download 5-4
InfiniPath 2-4
installation overview 3-2
package requirements 5-3
QLogic OFED layout 5-23
SRP 6-4
SRP, OFED SRP 6-4
Subnet management 1-3
Subnet Management Agent (SMA) 6-1
Supermicro H8DCE-HTe, problems with
QHT7040 A-2
Support, technical 1-5
Switch configuration and monitoring 4-17
Switches supported 1-2, 4-3

T

tar file, unpacking 5-7, 5-14, 5-17
Technical support 1-5
Terminology 1-3
Transport services supported 2-4

U

uDAPL, with MPI 6-13
Uninstalling InfiniPath and OpenFabrics RPMs
5-24
Uninstalling OFED software 5-24, 5-25

V

-v 7-3
--verbose 7-3
VNIC, configuration 6-6
-vv 7-3
--vverbose 7-3
-vvv 7-3
--vvverbose 7-3

W

--workdir=DIR 7-3

Draft

Draft



QLOGIC[®]

Corporate Headquarters QLogic Corporation 26650 Aliso Viejo Parkway Aliso Viejo, CA 92656 949.389.6000 www.qlogic.com

The Ultimate in Performance Europe Headquarters QLogic (UK) LTD. Quatro House Lyon Way, Frimley Camberley Surrey, GU16 7ER UK +44 (0) 1276 804 670

© 2005-2009 QLogic Corporation. Specifications are subject to change without notice. All rights reserved worldwide. QLogic and the QLogic logo are registered trademarks of QLogic Corporation. QLA, SANsurfer, InfiniPath, and SilverStorm are trademarks or registered trademarks of QLogic Corporation. AMD Opteron is a trademark of Advanced Microdevices Inc. BladeCenter and IBM are registered trademarks of International Business Machines Corporation. DataDirect Networks is a trademark of DataDirect Networks, Inc. EMCORE is a trademark of EMCORE Corporation. HTX is a trademark of the HyperTransport Technology Consortium. IBM and BladeCenter are registered trademarks of International Business Machines Corporation. InfiniBand is a trademark and service mark of the InfiniBand Trade Association. Intel is a registered trademark of Intel Corporation. Linux is a registered trademark of Linus Torvalds. LSI Logic and Engenio are trademarks or registered trademarks of LSI Logic Corporation. Lustre is a registered trademark of Cluster File Systems, Inc. Mellanox is a registered trademark and ConnectX is a trademark of Mellanox Technologies, Inc. PathScale is a trademark of PathScale LLC. PCI Express is a registered trademark of PCI-SIG Corporation. Red Hat and Enterprise Linux are registered trademarks of Red Hat, Inc. Supermicro is a registered trademark of Super Micro Computer Inc. SUSE is a registered trademark of Novell Inc. Zarlink is a trademark of Zarlink Semiconductor Inc. All other brand and product names are trademarks or registered trademarks of their respective owners. Information supplied by QLogic Corporation is believed to be accurate and reliable. QLogic Corporation assumes no responsibility for any errors in this brochure. QLogic Corporation reserves the right, without notice, to make changes in product design or specifications.