QLOGIC™

InfiniPath Install Guide

Version 2.0

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.

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 Beta2 10/15/06 QLogic Version Number IB0056101-00 B	
Rev. 2.0 11/30/06 QLogic Version Number IB0056101-00 C	
Changes	Document Sections Affected
Updated Introduction so that most information is contained in modules. The same info is contained in the User Guide Introduction.	section 1
Added cable management advisory	section 2.6
Updated some sections for clarity. Fixed spacing errors.	All
Added SLES 9 to list of distros that use openssh package.	section 3.2.2
Added SLES 9 to list of RPM identifiers	section 3.3, section 3.4
Updated tarball name	section 3.5.1
Added new install directory name for kernel modules	section 3.5.2
Updated install directory name for kernel modules	section 3.13
Added section on MPI over uDAPL	section 3.15

Table of Contents

Figures

Tables

Section 1

Introduction

This chapter describes the contents, intended audience, and organization of the InfiniPath Install Guide.

The InfiniPath Install Guide contains instructions for installing the QLogic InfiniPath Interconnect hardware adapters and the InfiniPath and OpenFabrics software. The adapters covered in this guide are the QLE7140, the QLogic InfiniPath PCI Express™ adapter, the the QMI7140, which runs on Power PC systems, particularly the IBM® BladeCenter H processor blades, and the QHT7040/QHT7140 HTX™ adapters.

For more information on system administration, using InfiniPath MPI, and troubleshooting adapter hardware and software, see the *InfiniPath Interconnect User Guide*.

1.1
Who Should Read this Guide

This installation guide is intended for cluster administrators responsible for installing the InfiniPath QMI7140, the QLE7140, or the InfiniPath QHT7040/QHT7140 product on your Linux cluster. More detailed installation information and instructions for administration of the InfiniPath cluster are provided in the *InfiniPath Interconnect User Guide*.

This guide assumes you are familiar with cluster networking and with the specific hardware to which your processors are connected. Prior to installing the InfiniPath Adapter, 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 in order to use basic Linux commands or to perform all system administration tasks. For this information, see the software documentation that you received with your system.

1.2
How this Guide is Organized

The *InfiniPath Install Guide* is organized into these sections:

- Section 1, "Introduction"

- Section 2, "Hardware Installation", instructions for installing the InfiniPath QLE7140, QMI7140, and QHT7040 or QHT7140 hardware

- Section 3, "Software Installation", instructions for installing the QLogic InfiniPath and OpenFabrics software

■ Appendix A, "RPM Descriptions"

1.3
Overview

The material in this documentation pertains to an InfiniPath cluster. This is defined as a collection of nodes, each attached to an InfiniBand™-based fabric through the InfiniPath Interconnect. The nodes are Linux-based computers, each having up to eight processors.

The InfiniPath interconnect is InfiniBand 4X, with a raw data rate of 10 Gb/s (data rate of 8Gb/s).

InfiniPath utilizes standard, off-the-shelf InfiniBand 4X switches and cabling. InfiniPath OpenFabrics software is interoperable with other vendors' InfiniBand HCAs running compatible OpenFabrics releases. There are two options for Subnet Management in your cluster:

■ Use the Subnet Manager on one or more managed switches supplied with your Infiniband switches.

■ Use the OpenSM component of OpenFabrics.

1.4
Switches

The InfiniPath interconnect is designed to work with all InfiniBand-compliant switches. Use of OpenSM as a subnet manager is now supported. OpenSM is part of the OpenFabrics component of this release.

1.5
Interoperability

InfiniPath participates in the standard InfiniBand Subnet Management protocols for configuration and monitoring. InfiniPath OpenFabrics (including IPoIB) is interoperable with other vendors' InfiniBand HCAs running compatible OpenFabrics releases. The InfiniPath MPI and Ethernet emulation stacks (`ipath_ether`) are not interoperable with other InfiniBand Host Channel Adapters (HCA) and Target Channel Adapters (TCA). Instead, InfiniPath uses an InfiniBand-compliant vendor-specific protocol that is highly optimized for MPI and TCP between InfiniPath-equipped hosts.

NOTE: OpenFabrics was known as OpenIB until March 2006. All relevant references to OpenIB in this documentation have been updated to reflect this change. See the OpenFabrics website at http://www.openfabrics.org for more information on the OpenFabrics Alliance.

1.6
What's New in this Release

QLogic Corp. acquired PathScale in April 2006. In this 2.0 release, product names, internal program and output message names now refer to QLogic rather than PathScale.

The new QLogic and former PathScale adapter model numbers are shown in the table below.

Table 1-1. PathScale-QLogic Adapter Model Numbers

Former PathScale Model Number	New QLogic Model Number	Description
HT-400	IBA6110	Single Port 10GBS InfiniBand to HTX ASIC ROHS
PE-800	IBA6120	Single Port 10GBS InfiniBand to x8 PCI Express ASIC ROHS
HT-460	QHT7040	Single Port 10GBS InfiniBand to HTX Adapter
HT-465	QHT7140	Single Port 10GBS InfiniBand to HTX Adapter
PE-880	QLE7140	Single Port 10GBS InfiniBand to x8 PCI Express Adapter
PE-850	QMI7140	Single Port 10GBS InfiniBand IBM Blade Center Adapter

This version of InfiniPath provides support for all QLogic's HCAs, including:

- InfiniPath QLE7140, which is supported on systems with PCIe x8 or x16 slots

- InfiniPath QMI7140, which runs on Power PC systems, particularly on the IBM® BladeCenter H processor blades

- InfiniPath QHT7040 and QHT7140, which leverage HTX™. The InfiniPath QHT7040 and QHT7140 are exclusively for motherboards that support HTXcards. The QHT7140 has a smaller form factor than the QHT7040, but is otherwise the same. Unless otherwise stated, QHT7140 will refer to both the QHT7040 and QHT7140 in this documentation.

Expanded MPI scalability enhancements for PCI Express have been added. The QHT7040 and QHT7140 can support 2 processes per context for a total of 16. The QLE7140 and QMI7140 also support 2 processes per context, for a total of 8.

Support for multiple versions of MPI has been added. You can use a different version of MPI and achieve the high-bandwidth and low-latency performance that is standard with InfiniPath MPI.

Also included is expanded operating system support, and support for the latest OpenFabrics software stack.

Multiple InfiniPath cards per node are supported. A single software installation works for all the cards.

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

http://www.qlogic.com

1.7
Supported Distributions and Kernels

The InfiniPath interconnect runs on AMD Opteron, Intel EM64T, and IBM Power Blade Center H) systems running Linux. The currently supported distributions and associated Linux kernel versions for InfiniPath and OpenFabrics are listed in the following table. The kernels are the ones that shipped with the distributions, unless otherwise noted.

Table 1-2. InfiniPath/OpenFabrics Supported Distributions and Kernels

Distribution	InfiniPath/OpenFabrics supported kernels
Fedora Core 3 (FC3)	2.6.12 (x86_64)
Fedora Core 4 (FC4)	2.6.16, 2.6.17 (x86_64)
Red Hat Enterprise Linux 4 (RHEL4)	2.6.9-22, 2.6.9-34, 2.6.9-42(U2/U3/U4) (x86_64)
CentOS 4.2-4.4 (Rocks 4.2-4.4)	2.6.9 (x86_64)
SUSE Linux 9.3 (SUSE 9.3)	2.6.11 (x86_64)
SUSE LInux Enterprise Server (SLES 9)	2.6.5 (x86_64)
SUSE LInux Enterprise Server (SLES 10)	2.6.16 (x86_64 and ppc64)

NOTE: IBM Power systems run only with the SLES 10 distribution.

The SUSE10 release series is no longer supported as of this InfiniPath 2.0 release. Fedora Core 4 kernels prior to 2.6.16 are also no longer supported.

1.8
Software Components

The software provided with the InfiniPath Interconnect product consists of:

■ InfiniPath driver (including OpenFabrics)

■ InfiniPath ethernet emulation

- InfiniPath libraries

- InfiniPath utilities, configuration, and support tools

- InfiniPath MPI

- InfiniPath MPI benchmarks

- OpenFabrics protocols, including Subnet Management Agent

- OpenFabrics libraries and utilities

OpenFabrics kernel module support is now built and installed as part of the InfiniPath RPM install. The InfiniPath release 2.0 runs on the same code base as OpenFabrics Enterprise Distribution (OFED) version 1.1. It also includes the OpenFabrics 1.1-based library and utility RPMs. InfiniBand protocols are interoperable between InfiniPath 2.0 and OFED 1.1.

This release provides support for the following protocols:

- IPoIB (TCP/IP networking)

- SDP (Sockets Direct Protocol)

- OpenSM

- UD (Unreliable Datagram)

- RC (Reliable Connection)

- UC (Unreliable Connection)

- SRQ (Shared Receive Queue)

- uDAPL (user Direct Access Provider Library)

This release includes a technology preview of:

- SRP (SCSI RDMA Protocol)

Future releases will provide support for:

- iSER (iSCSI Extensions for RDMA)

No support is provided for RD.

NOTE: 32 bit OpenFabrics programs using the verb interfaces are not supported in this InfiniPath release, but will be supported in a future release.

More details about the hardware and software are found in Section 2 and Section 3.

1.9
Conventions Used in this Guide

This Guide uses these typographical conventions:

Table 1-3. 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.
concept	Italic font is used for emphasis, concepts.
`user input`	Bold fixed-space font is used for literal items in commands or constructs that you type in.
$	Indicates a command line prompt.
#	Indicates a command line prompt as root when using bash or sh.
[ ]	Brackets enclose optional elements of a command or program construct.
...	Ellipses indicate that a preceding element can be repeated.
>	Right caret identifies the cascading path of menu commands used in a procedure.
2.0	The current version number of the software is included in the RPM names and within this documentation.
NOTE:	Indicates important information.

1.10
Documentation and Technical Support

The InfiniPath product documentation includes:

■ The InfiniPath Install Guide

■ The InfiniPath Interconnect User Guide

■ Release Notes

■ Quick Start Guide

■ Readme file

The Troubleshooting Appendix for installation, InfiniPath and OpenFabrics administration, and MPI issues is located in the InfiniPath Interconnect User Guide.

Visit the QLogic support Web site for documentation and the latest software updates.

http://www.qlogic.com

Section 2
Hardware Installation

This chapter lists the requirements and provides instructions for installing the QLogic InfiniPath Interconnect adapters. Instructions are included for the QLogic InfiniPath PCI Express Adapter and PCIe riser card, QLE7140, the QMI7140 for IBM Blade Center H processor blades, and the InfiniPath QHT7040 or QHT7140 adapter hardware and HTX riser card. These components will also be referred to as the *adapter* and the *riser card* in the remainder of this Guide.

The adapter is a low-latency, high-bandwidth, high message rate cluster interconnect for InfiniBand. The InfiniPath interconnect is InfiniBand 4X, with a raw data rate of 10Gb/s (data rate of 8Gb/s).

OpenFabrics is interoperable with other vendors' InfiniBand HCAs running compatible OpenFabrics releases.

2.1
Hardware Installation Requirements

This section lists hardware and software environment requirements for installing the InfiniPath QLE7140, InfiniPath QMI7140, and the InfiniPath QHT7040 or QHT7140.

2.1.1
Hardware

QLogic InfiniPath 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. The following table shows the relationship between the adapter model and different types of motherboards.

Table 2-1. Adapter Models and Related Platforms

QLogic Model Number	Platform	Plugs Into
QMI7140	Blade Center H processor blades	CFFE Connector[a]
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

Table Notes

[a]See the IBM documentation for information on the location of the CFFE connector. The IBM web site is: http://www.ibm.com

Installation of the QLE7140, QHT7040 or QHT7140 in a 1U or 2U chassis requires the use of a riser card. See figure 2-3 for an illustration of a PCI Express slot in a typical motherboard. See figure 2-6 for an illustration of an HTX slot for a typical Opteron motherboard.

Note that the motherboard vendor is the optimal source for information on the layout and use of HyperTransport and PCIe-enabled expansion slots on supported motherboards. For the most up-to-date listing of InfiniPath PCIe and HTX Adapter model numbers and the motherboards in which they are supported, please go to our web site:

http://www.qlogic.com

2.1.1.1
Form Factors

The **QMI7140** is the model number for the adapter card that ships in the IBM CFFE form factor. It does not use a faceplate.

The **QLE7140** is the model number for the adapter card that ships in the PCIe low profile short form factor. This card can be used with either a full height or low profile faceplate.

The **QHT7040** is the model number for the adapter card that ships in the HTX full height short form 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. Unless otherwise stated, the HTX adapter is referred to as the QHT7140 in this documentation.

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. This depends upon the form factor of the adapter and motherboard.

Run `ipath_control -i` to see information on which form factor adapter is installed. The file `/sys/bus/pci/drivers/ib_ipath/00/boardversion` contains the same information. For more information, see section C.9 in the Troubleshooting Appendix in the *InfiniPath Interconnect User Guide.*

2.1.2
Cabling and Switches

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

A standard InfiniBand copper cable up to a length of twenty meters is required. See section 2.6.1 for more information. Optical fibre with optical adapters such as the Emcore is also supported. See section 2.6.2.

There are two options for Subnet Management in your cluster:

■ Use the Subnet Manager on one or more managed switches supplied with your Infiniband switches.

■ Use OpenSM component of OpenFabrics.

Cabling instructions are found in this Guide in section 2.6.

2.1.3
Software

Before installing the InfiniPath software, complete the hardware installation as described in this chapter.

Verify that each machine in the cluster is running Linux kernel software from a supported distribution.

See section 3.2 for information on supported distributions, associated kernels and other software requirements.

2.1.4
Configuring the BIOS

A properly configured BIOS is required. 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.

There are currently two issues concerning BIOS settings that you need to be aware of:

- ACPI needs to be enabled

- MTRR mapping needs to be set to "Discrete"

MTRR (Memory Type Range Registers) is used by the InfiniPath driver to enable write combining to the InfiniPath on-chip transmit buffers. This improves write bandwidth to the InfiniPath chip by writing multiple words in a single bus transaction (typically 64). This applies only to x86_64 systems.

However, some BIOSes don't have the MTRR mapping option. It may be referred to in a different way, dependent upon chipset, vendor, BIOS, or other factors. For example, it is sometimes referred to as "32 bit memory hole", which should be enabled.

If there is no setting for MTRR mapping or 32 bit memory hole, please contact your system or motherboard vendor and inquire as to how write combining may be enabled.

ACPI and MTRR mapping issues are discussed in greater detail in the Troubleshooting section of the *InfiniPath Interconnect User Guide*.

NOTE: BIOS settings on IBM Blade Center H (Power) systems do not need adjustment.

You can check and adjust these BIOS settings using the BIOS Setup Utility. For specific instructions on how to do this, follow the hardware documentation that came with your system.

2.2
Safety with Electricity

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

1. 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.

2. Don't 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.

3. 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.

4. Use normal precautions to prevent electrostatic discharge, which can damage integrated circuits.

2.3
Unpacking Information

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

2.3.1
Verify Package Contents

The InfiniPath 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.

2.3.2
List of the Package Contents

The package contents for the QMI7140 adapter are:

■   InfiniPath QMI7140

An illustration of the QMI7140 is not yet available.

The package contents for the QLE7140 adapter are:

■   InfiniPath QLE7140

Standard PCIe risers can be used, typically supplied by your system or motherboard vendor. The contents are illustrated in figure 2-1 below.

The package contents for the QHT7140 adapter are:

■   InfiniPath QHT7140

■   HTX riser card for use in 1U or 2U chassis

The contents are illustrated in figure 2-2 below.

The IBA6120 and the IBA6110 are the QLogic InfiniPath Interconnect ASICs, which are the central components of the interconnect. The IBA6120 is shown in figure 2-1, and the IBA6110 is shown in figure 2-2.
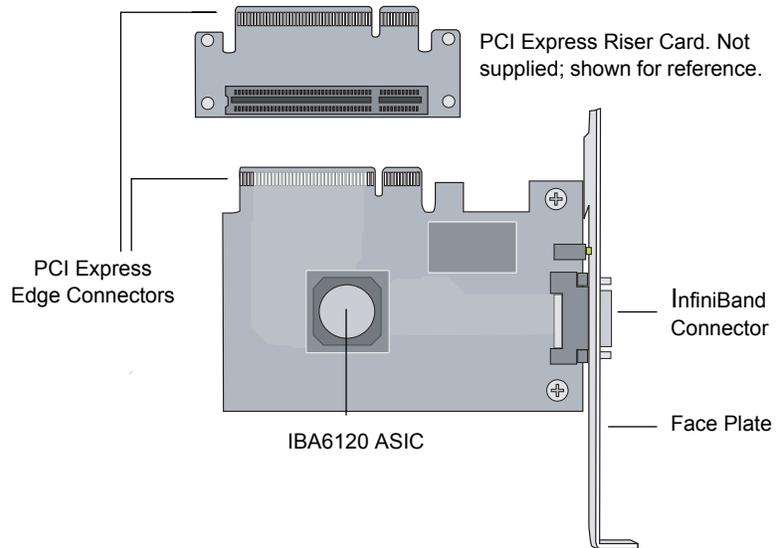
Figure 2-1. InfiniPath QLE7140 Card with Riser, Top View

QLOGIC
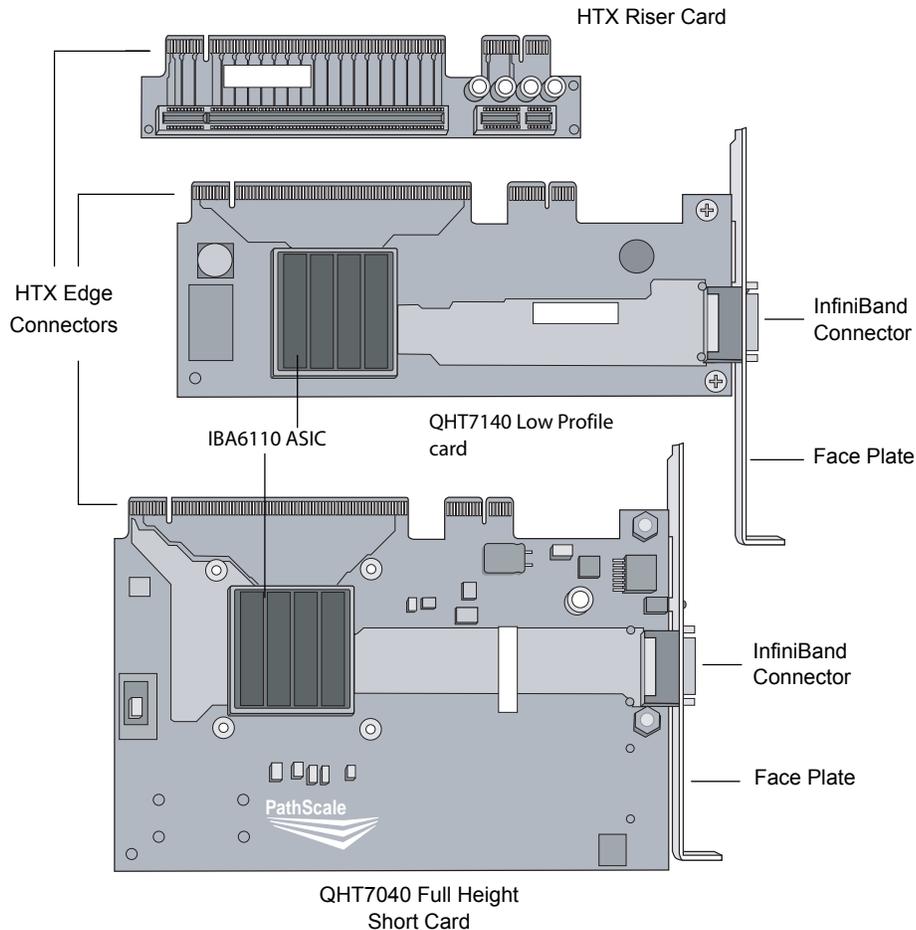
Figure 2-2. InfiniPath QHT7040/QHT7140 Full and Low Profile Cards with Riser, Top View

Unpacking the InfiniPath Adapter

When unpacking, ground yourself before removing the InfiniPath Adapter from the anti-static bag.

1. Grasping the InfiniPath Adapter by its face plate, pull the adapter out of the anti-static bag. Handle the adapter only by its edges or the IB connector. Do not allow the InfiniPath Adapter or any adapter card components to touch any metal parts.

2. After checking for visual damage, store the InfiniPath Adapter and the riser card in the anti-static bag until you are ready to install them.

2.4
Hardware Installation

This section contains hardware installation instructions for all of the adapters.

2.4.1
Hardware Installation for QMI7140

For installation of the QMI7140, see the IBM BladeCenter H Installation documentation. This should provide information on installing hardware in the CFFE slot. It connects to the blade and midplane of the IBM BladeCenter H. There is no faceplate with the QMI7140.

Cables do not need to be installed with the QMI7140. After installing the hardware according to the IBM BladeCenter H documentation, proceed to section 3 for instructions on installing the InfiniPath software.

2.4.2
Hardware Installation for QLE7140 with PCIe Riser

Most installations will be in 1U and 2U chassis, using the PCI Express riser card. This results in a horizontal installation of the QLE7140. This type of installation will be described first. Installation in a 3U chassis will be described in the next section.

Installation of InfiniPath QLE7140 in 1U or 2U chassis requires installation with a PCI Express riser card.

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

To install the InfiniPath adapter with a PCI Express 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 InfiniPath adapter. See section 2.1.4.

2. Shut down the power supply to the system into which you'll be installing the InfiniPath adapter.

3. Take precautions to avoid damage to the cards by grounding yourself or touching the metal chassis to discharge static electricity before handling them.

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 2-3, below. These two sections correspond to the shorter and longer connector edges of the adapter and riser.

PCIe slot in
typical motherboard
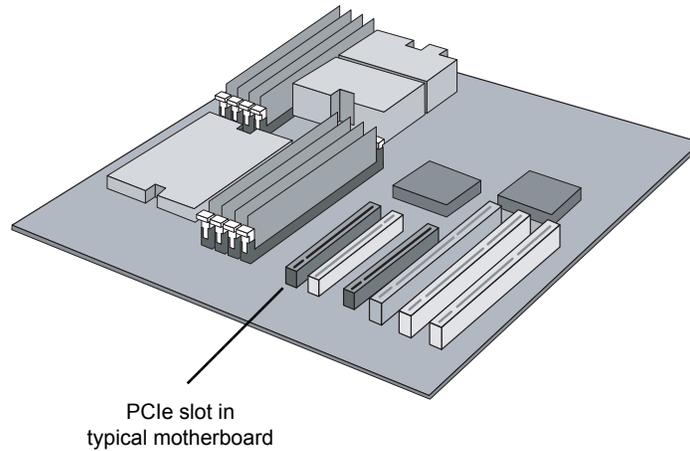
Figure 2-3. PCIe Slot in a Typical Motherboard

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 InfiniPath QLE7140 from the anti-static bag.

8. Locate the face plate on the connector edge of the card.

9. Connect the InfiniPath adapter and PCIe riser card together, forming the assembly that you'll insert into your motherboard. To do this, first visually line up the card slot connector edge with the edge connector of the PCIe riser card.

.
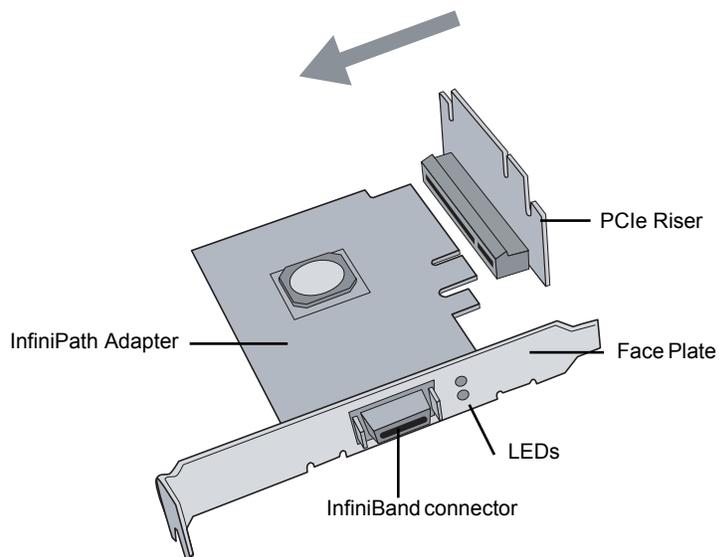
Figure 2-4. InfiniPath QLE7140 Assembly with Riser Card

10. Holding the InfiniPath adapter by its edges, carefully insert the card slot connector into the PCIe riser card edge connector, as show in figure 2-4 above. The result is a combined L-shaped assembly of the PCIe riser card and InfiniPath Adapter. This assembly is what you'll 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 mother board, 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 InfiniBand 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 InfiniPath Adapter should now be parallel to the motherboard and about one inch above it.
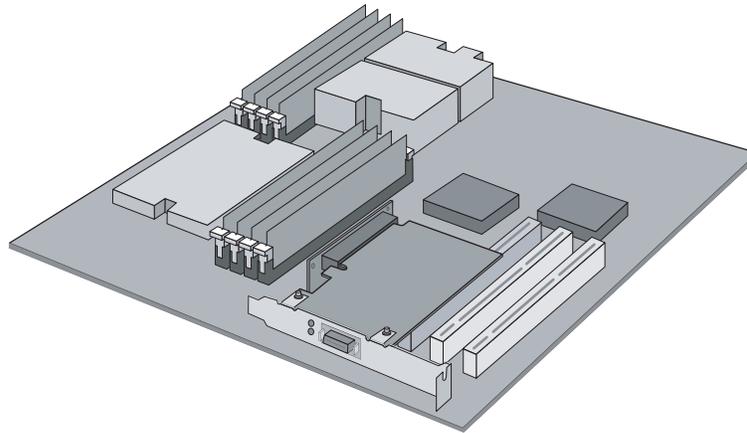
Figure 2-5. Assembled QLE7140 with Riser

14. Secure the face plate to the chassis. The InfiniPath adapter has a screw hole on the side of the face plate which can be attached to the chassis with a retention screw. The securing method may vary depending upon 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 InfiniPath QLE7140 with PCIe riser card is now installed. Next, install the cables as described in section 2.6 Then test your installation by powering up and verifying link status. See section 2.7.

2.4.3
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 will be described first. Installation in a 3U chassis will be described in the next section.

Installation of InfiniPath QHT7140 in 1U or 2U chassis requires installation with 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 will follow the same steps.

To install the InfiniPath 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 InfiniPath adapter. See section 2.1.4.

2. Shut down the power supply to the system into which you'll be installing the InfiniPath adapter.

3. Take precautions to avoid damage to the cards by grounding yourself or touching the metal chassis to discharge static electricity before handling them.

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 2-6.

HTX Slot in a Typical Opteron Motherboard
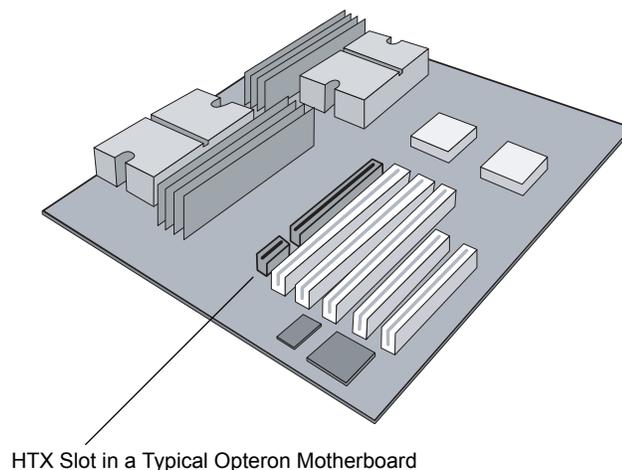
Figure 2-6. HTX Slot

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 InfiniPath QHT7140 from the anti-static bag.

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

8. Locate the face plate on the connector edge of the card.

9. Connect the InfiniPath adapter and HTX riser card together, forming the assembly that you'll insert into your motherboard. To do this, first visually line up the card slot connector edge with the edge connector of the HTX riser card.
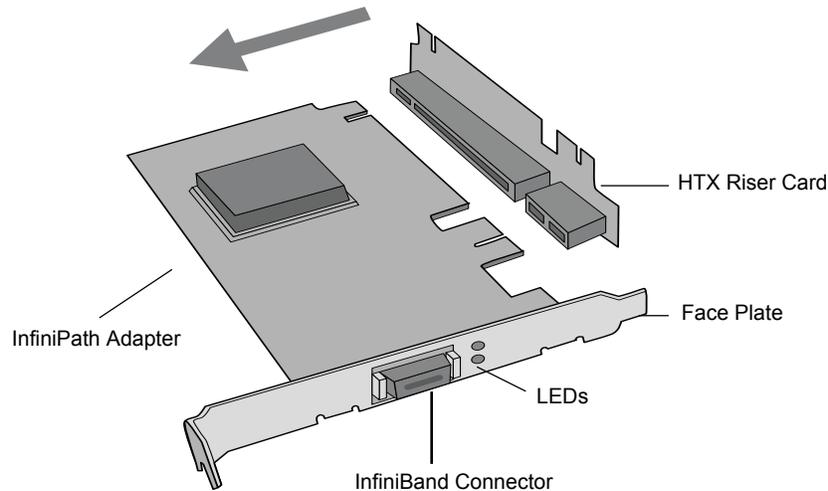
Figure 2-7. InfiniPath Adapter with Riser Card.

10. Holding the InfiniPath adapter by its edges, carefully insert the card slot connector into the HTX riser card edge connector, as show in Figure 3 above. The result is a combined L-shaped assembly of the HTX riser card and InfiniPath Adapter. This assembly is what you'll 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 mother board, and the face plate is toward the front of the chassis.

12. Holding this HT assembly above the motherboard at about a 45 degree angle, slowly lower it so that the InfiniBand 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 HT riser and HTX slot should line up perfectly at this point.

13. Insert the HT riser assembly into the motherboard's HTX slot, ensuring good contact. The InfiniPath Adapter should now be parallel to the motherboard and about one inch above it.
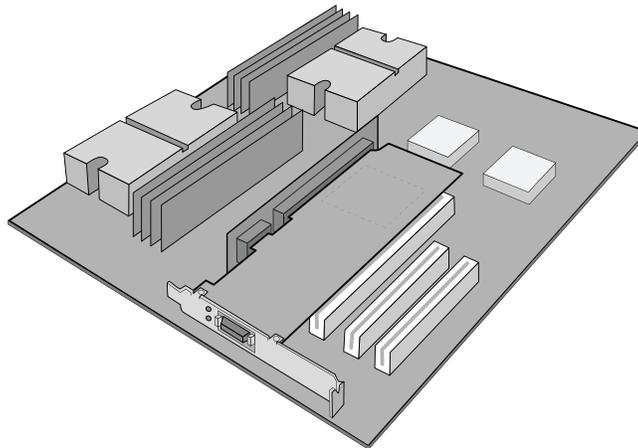
Figure 2-8. Assembled QHT7140 with Riser

14. Secure the face plate to the chassis. The InfiniPath adapter has a screw hole on the side of the face plate which can be attached to the chassis with a retention screw. The securing method may vary depending upon 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 InfiniPath QHT7140 with HTX riser card is now installed. Next, install the cables as described in section 2.6. Then test your installation by powering up and verifying link status. See section 2.7.

2.5
Hardware Installation for QHT7140 without HTX Riser

The installation of InfiniPath 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 InfiniPath Adapter without the 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 InfiniPath adapter. See section 2.1.4.

2. Shut down the power supply to the system into which you'll be installing the InfiniPath Adapter.

3. Take precautions to avoid damage to the cards by grounding yourself or touching the metal chassis to discharge static electricity before handling them.

4. If you are installing the InfiniPath Adapter into a covered system, you will first need to 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 2-6, above.

6. Remove the InfiniPath 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.
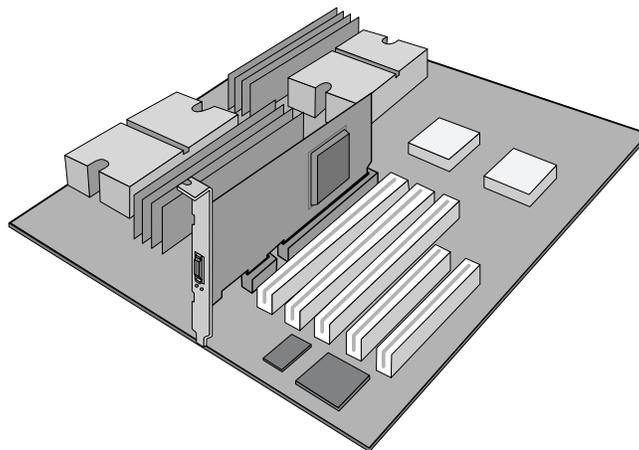
Figure 2-9. QHT7140 without Riser Installed in a 3U Chassis

9. Secure the face plate to the chassis. The InfiniPath adapter has a screw hole on the side of the face plate which can be attached to the chassis with a retention

screw. The securing method may vary depending upon 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 section 2.6. Then test your installation by powering up the system. See section 2.7.

2.6
Cabling the Adapter to the InfiniBand Switch

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

The QLE7140, QHT7040, and QHT7140 adapters are all cabled in the same way. If you have a QMI7140 adapter for Power systems, note that the cabling process is slightly different. See section 2.6.1.1 for more information.

2.6.1
Standard InfiniBand Copper Cabling

The cable installation uses a standard InfiniBand (IB) cable. Any InfiniBand cable that has been qualified by the vendor should work. The longest IB cable we have currently qualified is twenty meters.

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.

3. If your cable latch mechanism is not spring-loaded, simply push on the metal case, then push the plastic latch to lock the cable in place.

4. The InfiniBand cables are symmetric; either end can be plugged into the switch. Connect the InfiniBand cable to the connector on the InfiniPath 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.

5. Connect the other end of the cable to the InfiniBand switch.

2.6.1.1
Cabling for the QMI7140

Follow Steps 1-3 from the previous section. The cables for the QMI7140 plug into the Pass Through Switch Module in the back of the system, rather than into the adapter card. Please see the documentation for the IBM BladeCenter H for information on the Pass Through Switch Module.

2.6.2
Optical Fibre Option

The InfiniPath 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, contact www.emcore.com.

2.7
Completing the Installation

This section applies to the QLE7140, QHT7040, and QHT7140 adapters.

NOTE: If you have a QMI7140 adapter for Power systems, proceed to software installation, as described in section 3.

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-on. 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 state, see section 3.18.

Section 3
Software Installation

This chapter provides instructions for installing the InfiniPath and OpenFabrics software. The InfiniPath software includes drivers, protocol libraries, QLogic's implementation of the MPI message passing standard, and example programs, including benchmarks.

The software provided with the InfiniPath Interconnect product consists of:

- InfiniPath driver (including OpenFabrics)

- InfiniPath ethernet emulation

- InfiniPath libraries

- InfiniPath utilities, configuration, and support tools

- InfiniPath MPI

- InfiniPath MPI benchmarks

- OpenFabrics protocols, including Subnet Management Agent

- OpenFabrics libraries and utilities

OpenFabrics kernel module support is now built and installed as part of the InfiniPath RPM install. The InfiniPath release 2.0 runs on the same code base as OpenFabrics Enterprise Distribution (OFED) version 1.1. It also includes the OpenFabrics 1.1-based library and utility RPMs. InfiniBand protocols are interoperable between InfiniPath 2.0 and OFED 1.1.

This release provides support for the following protocols:

- IPoIB (TCP/IP networking)

- SDP (Sockets Direct Protocol)

- OpenSM

- UD (Unreliable Datagram)

- RC (Reliable Connection)

- UC (Unreliable Connection)

- SRQ (Shared Receive Queue)

- uDAPL (user Direct Access Provider Library)

This release includes a technology preview of:

- SRP (SCSI RDMA Protocol)

Future releases will provide support for:

■   iSER (iSCSI Extensions for RDMA)

No support is provided for RD.

NOTE: 32 bit OpenFabrics programs using the verb interfaces are not supported
in this InfiniPath release, but will be supported in a future release.

Further information, including the use of MPI, is found in the companion document,
the *InfiniPath Interconnect User Guide*. For convenience, some information appears
in both documents.

3.1
Software Installation Overview

InfiniPath and OpenFabrics software installation involves the following steps:

1. Make sure that the InfiniPath adapter installation has been completed according
    to the instructions in section 2.

2. Verify that the Linux kernel software is installed on each node in your cluster.
    The required kernels and supported Linux distributions for both InfiniPath and
    OpenFabrics are defined below. See section 3.2.

3. For each release, download the InfiniPath/OpenFabrics software from the
    QLogic web site to a local server directory, and from there install the appropriate
    packages on each cluster node as described under section 3.3, section 3.4,
    and section 3.5. Configure the InfiniPath and OpenFabrics drivers as described
    under section 3.8.

4. Check that all InfiniBand switches are configured. See section 3.17

5. Perform the recommended health checks. See section 3.19.

6. After installing the InfiniPath and OpenFabrics refer to the *InfiniPath
    Interconnect User Guide* for more information on using the InfiniPath and
    OpenFabrics products.

3.2
Software Installation Requirements

Information on supported distributions and kernels and on setting up your
environment are given below.

3.2.1
Linux Environment

The currently supported distributions and associated Linux kernel versions for InfiniPath and OpenFabrics are listed in the following table. The kernels are the ones that shipped with the distributions, unless otherwise noted.

Table 3-1. InfiniPath/OpenFabrics Supported Distributions and Kernels

Distribution	InfiniPath/OpenFabrics supported kernels
Fedora Core 3 (FC3)	2.6.12 (x86_64)
Fedora Core 4 (FC4)	2.6.16, 2.6.17 (x86_64)
Red Hat Enterprise Linux 4 (RHEL4)	2.6.9-22, 2.6.9-34, 2.6.9-42(U2/U3/U4) (x86_64)
CentOS 4.2-4.4 (Rocks 4.2-4.4)	2.6.9 (x86_64)
SUSE Linux 9.3 (SUSE 9.3)	2.6.11 (x86_64)
SUSE LInux Enterprise Server (SLES 9)	2.6.16 (x86_64)
SUSE LInux Enterprise Server (SLES 10)	2.6.16 (x86_64 and ppc64)

NOTE: IBM Power systems run only with the SLES 10 distribution.

The SUSE10 release series is no longer supported as of this InfiniPath 2.0 release. Fedora Core 4 kernels prior to 2.6.16 are also no longer supported.

3.2.2
Environment for InfiniPath and OpenFabrics

There are a few things to keep in mind when setting up the environment:

■ It is possible to have a cluster running with different kernel versions. However, QLogic recommends and supports clusters where all nodes run equivalent software.

■ Different distributions require different versions of the InfiniPath software distribution, as described below under section 3.3.

■ Among the many optional packages that each distribution offers, the InfiniPath software requires `openssh` and `openssh-server` and, if the MPD job launcher is to be used, `python`. These must be on every node. Note that in the SUSE, SLES 9, and SLES 10 distributions `openssh-server` is a part of the `openssh` package.

■ The runtime and build environments should be the same. Compatibility between executables built on different Linux distributions cannot be guaranteed.

■ Install `sysfsutils` for your distribution before installing the OpenFabrics RPMs, as there are dependencies. Check your distribution's documentation for

information about `sysfsutils`. This package is called `udev` on SUSE 9.3, SLES 9, and SLES 10.

3.2.3
Kernel Patches

Some kernels, such as some versions of Fedora Core 4 (2.6.16), have `CONFIG_PCI_MSI=n` as the default. If the InfiniPath driver is being compiled on a machine without `CONFIG_PCI_MSI=y` configured, you will get a compilation error. This default may also be introduced with updates to other Linux distributions or local configuration changes.

NOTE: IBM provides a pre-installed patched kernel For Power systems.

Systems that use the InfiniPath QLE7140, contain an AMD-8131 south bridge and use kernels between 2.6.12 - 2.6.16 require a patch for a problem that disabled the use of MSI interrupt handlers, which are used with the QLE7140.

NOTE: This does not apply to Power systems running SLES 10.

InfiniPath-specific kernels and patches for these problems are available for download on the QLogic web site. Please go to:

http://www.qlogic.com

Follow the links to the download pages. This topic is also covered in appendix C.4 in the *InfiniPath Interconnect User Guide*.

3.2.4
Types of Nodes in a Cluster Environment

In a cluster environment different nodes may be used for different functions, such as launching jobs, for software development, or running jobs. We define these nodes in the following way:

■ *Front end node*. This node will launch jobs. It is referred to as the *front end node* in this documentation.

■ *Compute node*. These are the nodes that are used to run jobs.

■ A *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 only one or just a few front end nodes. The number of nodes used for development will vary. The terms defined above will be used in the rest of the documentation.

3.3
InfiniPath Software RPMs

Linux distributions of InfiniPath software are installed from binary RPMs. RPM is a Linux packaging and installation tool used by Red Hat, SUSE, and CentOS. Each set of RPMs uses a build identifier xxx and a distribution identifier yyy. The RPM distribution identifiers are shown below with the associated operating systems.

Table 3-2. InfiniPath RPMs to Use for Each Node in a Cluster

Distribution identifier	Used on
fc3	Fedora Core 3, Fedora Core 4, Red Hat Enterprise Linux 4 (rhel4), CentOS 4.0-4.3 (Rocks 4.0-4.3) for x86_64 systems
suse9.3	SUSE 9.3, ,SLES 9, SLES 10 for x86_64 systems
sles10 (for ppc64)	SLES 10 for IBM Blade Center H (ppc64) systems only

The individual RPMs are listed and described in Appendix A "RPM Descriptions". To generate a list of InfiniPath software package contents on each RPM run:

```
$ rpm -qlp rpm_file_name
```

3.4
OpenFabrics RPMs

The OpenFabrics kernel module support is now part of the InfiniPath RPMs. However, in order to use OpenFabrics, you need to install these user libraries, utilities and other program RPMs.

Each set of the OpenFabrics RPMs uses a distribution identifier yyy. The identifiers for this release are `fc4`, `suse9.3`, and `sles10`. Note that these identifiers are different than those for the InfiniPath RPMs.

The RPM distribution identifiers are shown below with the associated operating systems.

Table 3-3. OpenFabrics RPMs to Use for Each Node in a Cluster

Distribution identifier	Used on
fc4	Fedora Core 4 x86_64 systems
suse9.3	SUSE 9.3, SLES 9, SLES 10 x86_64 systems
sles10	SLES 10 for IBM Blade Center H (ppc64) systems only

NOTE: The earlier OpenIB for PathScale RPMs (Versions 0.1.0 and 0.2.0) for the InfiniPath 1.1 and 1.2 releases are not compatible with this InfiniPath Version 2.0 release. If you have installed these RPMs, you must remove

or upgrade them prior to using the InfiniPath Version 2.0 release. The OpenFabrics RPMs as described in this documentation are compatible with the InfiniPath Version 2.0 release.

To generate a list of software package contents on each OpenFabrics RPM run:

```
$ rpm -qlp rpm_file_name
```

NOTE: There are currently some OpenFabrics limitations with IBM Power Systems. SRP and iSER do not yet work. This will be fixed in a future release.

3.5
Downloading and Installing InfiniPath and OpenFabrics Software

In this section we assume that the correct Linux kernel and a supported distribution have been installed on every node. Although QLogic recommends that all RPMs are installed on all nodes, some 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, see appendix A.

3.5.1
Downloading the `tar` File Containing the RPMs

There are multiple interdependent RPM packages that make up the InfiniPath and OpenFabrics software. All the RPMs are packaged into a single tar file for download.T he OpenFabrics RPMs are also contained in this tar file. The `tar` files can be downloaded from the QLogic web site:

http://www.qlogic.com

Follow the links to the desired download, and follow the instructions for downloading on the web page. Download it to any convenient directory. Then:

```
$ tar zxvf InfiniPath2.0-xxx-yyy.tgz
```

where *xxx* is the distribution identifier, `Fedora`, `SuSE9.3,` or `SLES10 and` *yyy* is the platform architecture, x86_64 or ppc64. The `tar` command will create a directory based on the tar file name and place the RPMs there.

NOTE: A specific download directory is not required. The install process will create and install in the correct directories. The locations of the install directories are given in the *InfiniPath Interconnect User Guide* in section 2.2.

The InfiniPath OFED 1.1 library and utility source `tar` file may also be downloaded from the web site.

3.5.2
Installing the Software

The new tarball layout makes it easier to identify the software. The RPMs are now organized in the following way:

```
InfiniPath_license, InfiniPath_release_notes, LEGAL (top level)
Documentation/
InfiniPath/
InfiniPath-Devel/
InfiniPath-MPI/
OpenFabrics/
OpenFabrics-Devel/
OpenSM/
OtherHCAs/
```

A complete list of RPMs is in appendix A.

Before installing, please take note of the following:

■ Make sure that all previously existing OpenFabrics RPMs (including the previous OpenIB for PathScale Beta 1 and 2 RPMs) are uninstalled. See section 3.20 for more information on uninstalling.

■ Installing all the RPMs on all nodes is easiest. However, if you do not plan to run OpenFabrics programs, you will only need to install the RPMs in the `InfiniPath` and `InfiniPath-MPI` directories. If you plan to run only OpenFabrics programs, install the RPMs in the `InfiniPath` and `OpenFabrics` directories. If you plan to run both InfiniPath and OpenFabrics, install all three of these RPMs.

■ You should only install the `OpenSM` RPM if you do not plan to use a switch-based subnet manager. It should normally be installed on just two nodes. If installed, it is enabled to be **on** by default. This behavior can be modified. See section 3.10 for more information.

■ The kernel modules in this release are installed in this directory:

  ```
  /lib/modules/`uname -r`/updates
  ```

  This avoids replacing kernel modules which may be provided by your Linux distribution, so they will be available if the InfiniPath software is removed. In this release the module `ipath_core.ko` has been renamed to `ib_ipath.ko`, and conflicts can arise if `ipath_core.ko` is also present. If it is found during installation of the `infinipath-kernel` RPM, `ipath_core.ko` is renamed to `ipath_core.ko.bak`.

The RPMs need to be available on each node on which they will be used. One way to do this is to copy the RPMs to a directory on each node that will need them. Another way is to put the RPMs in a directory accessible (e.g., via NFS) to every

node. After making sure the RPMs are on each node, login as root and, for the `InfiniPath` and `InfiniPath-MPI` RPMs, run the command:

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

For the OpenFabrics RPMs run the command:

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

Parallel command starters can be used for installation on multiple nodes, but this is beyond the scope of this documentation.

NOTE: Installation of the InfiniPath driver RPM (`infinipath-kernel-2.0-xxx-yyy`) builds kernel modules for the currently running kernel version. These InfiniPath modules will only work with that kernel. If a different kernel is booted, the InfiniPath driver RPM must be re-installed or rebuilt See section 3.14 for more information.

NOTE: 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`. Check the new file(s) to see if there is anything you want to add to your standard config file.

An alternate method to manage installation by using Rocks is given below in section 3.6.

3.5.3
Installed Layout

The default installed layout for the InfiniPath software is listed below.

The InfiniPath 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 source for several MPI benchmarks are in:

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

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

```
/usr/bin
```

The InfiniPath kernel modules are installed in the standard module locations in:

```
/lib/modules (version dependent)
```

Documentation can be found in:

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

3.5.4
Installing in an Alternate Location

The `--prefix` option can be used to specify an alternate installation location, replacing the `/usr` prefix as noted in section 3.5.3, above. Some files are installed in `/etc` regardless of whether or not `--prefix` is used.

If you install into an alternate location you should set the variable `INFINIPATH_ROOT` to the same value. See section 3.5.8.1 in the InfiniPath Interconnect User Guide for more information on setting this environment variable.

3.5.4.1
Installing for Your Distribution

You may be using a kernel which is compatible with one of the supported distributions, but which may not be picked during `infinipath-kernel` installation. It may also happen when using `make-install.sh` to manually recompile the drivers.

In this case, you can set your distribution with the `$IPATH_DISTRO` override. Run this command before installation, or before running `make-install.sh`. We use the RHEL4 Update 4 distribution as an example in this command for `bash` or `sh` users:

```
$ export IPATH_DISTRO=rhel4_U4
```

The distribution arguments that are currently understood are found in the file `build-guards.sh.make-install.sh` and `build-guards.sh` are both found in this directory:

```
/usr/src/infinipath/drivers
```

These are the arguments used for RHEL, CentOS(Rocks), and Scientific Linux.

```
rhel4_U2
rhel4_U3
rhel4_U4
```

These are the arguments used for SLES, SUSE, and Fedora:

```
sles10
suse9.3
fc3
fc4
```

3.5.4.2
Check for Missing Files or RPMs

Run the `rpm` command with the `--verify` option to check if there are files missing from the RPMs. For example:

```
$ rpm -a --verify 'InfiniPath-MPI/mpi*' 'InfiniPath/infinipath*'
```

In the case of the OpenFabrics RPMs, the verification command is slightly different. This is because these RPMs have many different prefixes:

```
$ rpm --verify rpm_name_pre
```

rpm_name_pre is the descriptive name that precedes the version and repository identifiers in an RPM. For example:

```
$ rpm --verify libibverbs
```

This would verify the contents of:

`libibverbs-2.0-`*xxx*`.1_0_4.`*yyy*`.x86_64.rpm`

The `rpm` command cannot be used to check for missing RPMs. Use `ipath_checkout` to flag missing RPMs. for more information.

3.6
Managing and Installing Software 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 that was 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.

We extend the normal Rocks compute node appliance .xml file by adding two functions: one to install the QLogic PathScale compilers, and an install script that loads the drivers after kickstart reboots the machine.

This section provides an overview on one way of building a Rocks cluster using the recommend rolls, and a sample xml file that describes the contents of a kickstart config file. By reading this and following the instructions on the Rocks web site, you can then install the InfiniPath RPMs on the required cluster nodes.

NOTE: There are many ways to use Rocks to manage clusters. Familiarizing oneself first with kickstart and then Rocks is advisable before using this method to install the InfiniPath RPMs.

3.6.1
Installing Rocks and InfiniPath RPMs

The following instructions are for building a Rocks 4.0.0 cluster with the 2.6.9 Linux kernel. It will install both the QLogic PathScale compilers and InfiniPath. These instructions are guidelines only; you will need to refer to material on the Rocks web site to complete an installation. If you want to use later versions of Rocks, these instructions will serve as general procedural steps.

The steps are summarized first, then more details are given in the corresponding numbered steps in the next section.

1. Download the required rolls from the Rocks web site:

   http://www.rocksclusters.org/

2. Build the front end node with these rolls.

3. Download and install the config and kickstart RPMs onto the front end node.

4. Download and install the InfiniPath RPMs onto the front end node.

5. Create a directory that contains the config, kickstart, and InfiniPath RPMs on the front end node.

6. Create the xml file /home/install/site-profiles/4.0.0/nodes/extend-compute.xml

7. Use the customized xml file to complete the installation. This file describes the contents of the kickstart config file to install the RPMs from the front end node onto all the other nodes. As this is a Rocks-specific process, you are referred to the Rocks web site for further step-by-step instructions.

3.6.2
Installing Rocks and InfiniPath RPMs-Details

This section fills in the installation details for the numbered steps listed above.

1. Download the required rolls from the Rocks web site. Go to:

   http://www.rocksclusters.org/ftp-site/

   Follow links to the following CDs:

   ```
   Boot Roll: Rocks 4.0.0 X86_64 Compute ISO
   (Area51+Base+Ganglia_HPC+Java+Kernel+SGE)
   OS Roll - Disk 1
   OS Roll - Disk 2
   ```

   You will also need:

   ```
   updates-4.0.0-0.x86_64.disk1.iso
   ```

2. Build the front end node with the above CDs. For more details see the Rocks installation documentation on the Rocks web site.

3. Download the config and kickstart RPMs onto the front end node. These can also be found at http://www.rocksclusters.org/ftp-site/

```
rocks-config-4.0.0-1.x86_64.rpm
rocks-config-frontend-4.0.0-1.x86_64.rpm
roll-kernel-kickstart-4.0.0-1.noarch.rpm
```

4. Install these RPMs listed above in step 3 on the front end node. For more details see the Rocks installation documentation on the Rocks web site. Then copy the RPMs to

```
/home/install/contrib/4.0.0/x86_64/RPMS
```

NOTE: This next step gives an example that presents solutions for two common issues Rocks users might encounter: installing both the InfiniPath RPMs and the QLogic PathScale compiler for a run-time environment. However, it is not a requirement to install the compiler.

5. Download and install the QLogic PathScale 2.5 compiler and InfiniPath 2.0 RPMs on the front end node; this example includes both for convenience. Then copy the RPMs to the same directory as in step 4.

The directory `/home/install/contrib/4.0.0/x86_64/RPMS` should now contain the files in this list. Note that just the prefix of the RPM names for infinipath, mpi, and the compiler are included; the full name will vary from release to release. See your current release for the complete names.

```
infinipath-
infinipath-kernel-
infinipath-doc-
infinipath-libs-
mpi-benchmark-
mpi-devel-
mpi-doc-
mpi-frontend-
mpi-libs-
infinipath-base-
infinipath-compilers-libs-
infinipath-pathdb-
infinipath-sub-client-
rocks-config-4.0.0-1.x86_64.rpm
rocks-config-frontend-4.0.0-1.x86_64.rpm
roll-kernel-kickstart-4.0.0-1.noarch.rpm
*kernel-smp-2.6.9-current_version.src.rpm
*kernel-smp-2.6.9-current_version.x86_64.rpm
*kernel-smp-devel-2.6.9-current_version.x86_64.rpm
```

*The kernel RPMs are built using the how-to instructions on the rocksclusters.org web site, and added into this file in the same manner as for the InfiniPath, config, and kickstart RPMs mentioned in the steps above. Check for the appropriate `current_version` of 2.6.9.

6. Create the file:

```
/home/install/site-profiles/4.0.0/nodes/extend-compute.xml
```

Use the following contents:

```
<?xml version="1.0" standalone="no"?>
<kickstart>
<description>
</description>
<changelog>
</changelog>
<-- QLogic PathScale Compilers -->
<package>infinipath-sub-client</package>
<package>infinipath-base</package>
<package>infinipath-compilers-libs</package>
<-- InfiniPath Drivers -->
<package>mpi-doc</package>
<package>mpi-frontend</package>
<package>mpi-libs</package>
<package>mpi-benchmark</package>
<package>mpi-devel</package>
<post>

<-- Here we fix up Grub boot behavior -->

<-- add serial console on tty1 -->
sed -e 's/kernel<.*>$/kernel\1 console=tty1
console=ttyS1,9600n8/' \
/boot/grub/grub-orig.conf > /tmp/grub-orig.conf

<-- set serial console speed -->
sed  -e  's/hiddenmenu/serial  --unit=0 --speed=9600 --word=8 \
--parity=no --stop=1\nterminal serial console/g' \
/tmp/grub-orig.conf > /tmp/grub-orig2.conf

<-- remove tsc -->

sed -e 's/tsc//g' /tmp/grub-orig2.conf > /tmp/grub-orig3.conf
```

```
<-- remove noacpi, needed to make InfiniPath happy -->
sed -e 's/pci=noacpi//g' /tmp/grub-orig3.conf >
/tmp/grub-orig4.conf
mv /tmp/grub-orig4.conf /boot/grub/grub-orig.conf

<!-- add getty for console to inittab at all runlevels -->
<file name="/etc/inittab" mode="append">
s1:2345:respawn:/sbin/agetty -L 9600 ttyS1 vt100
</file>

<-- Remove something for the kernel upgrade -->
rm -f /etc/rc.d/rocksconfig.d/pre-09-prep-kernel-source


<-- Reinstall our infinipath driver RPM -->
<!-- This adds the MPD disable-->
<file  name="/etc/rc.d/rocksconfig.d/post-97-infinipath"
mode="create perms="a+rx">
#!/bin/sh
rpm -ivh --force \
/home/install/contrib/4.0.0/x86_64/RPMS/infinipath*.rpm
/etc/init.d/infinipath start
rm /etc/rc.d/rocksconfig.d/post-97-infinipath
</file>

<-- Add InfiniPath compiler library path to ld.so -->
<file name="/etc/ld.so.conf" mode="append">
/opt/infinipath/lib/2.5/
</file>
</post>
</kickstart>
```

The important thing to note in this file is that the installation of the InfiniPath
drivers is done in the <post> section, as it is a "live" install. This file can be used
as a guideline: it may be cut and pasted, and then modified to suit your needs.

7. The completion of the installation is done using this xml file. Instructions for
   completing the install process are given under the Documentation link on the
   Rocks web site:

   http://www.rocksclusters.org/

3.6.3
Further Information on Rocks and kickstart

Checking the Rocks web site for updates is recommended. Extensive
documentation on installing Rocks and custom rolls is on the Rocks web site. For
more information on Rocks, visit:

http://www.rocksclusters.org/

To find more information on Red Hat Enterprise Linux 4, and on using kickstart, see:

http://www.redhat.com/

3.7
InfiniPath and OpenFabrics Driver Overview

The `ib_ipath` module provides low level InfiniPath hardware support, and is the base driver for both the InfiniPath and OpenFabrics software components. The `ib_ipath` module does hardware initialization, handles InfiniPath-specific memory management, and provides services to other InfiniPath and OpenFabrics modules. It provides the hardware and hardware management functions for InfiniPath MPI programs, the `ipath_ether` Ethernet emulation, and general OpenFabrics protocols such as IPoIB and SDP. IPoIB is the protocol that allows you to run Internet Protocol over an IB fabric. The module also contains a Subnet Management Agent.

This diagram shows the relationship between the InfiniPath and OpenFabrics software. Not all components are shown.
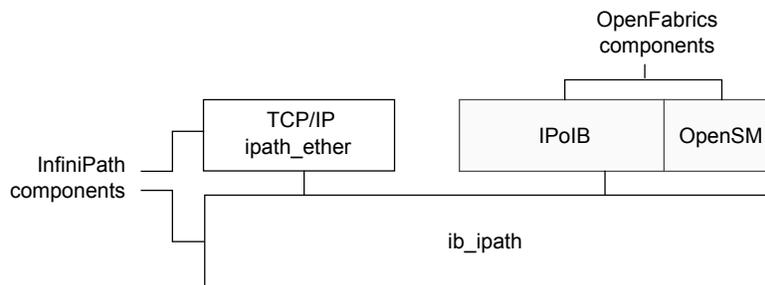
Figure 3-1. Relationship between InfiniPath and OpenFabrics Software

If you want to enable TCP-IP networking for running Ethernet traffic over the InfiniPath link, you can configure the optional `ipath_ether` network interface files.

NOTE: It is not necessary to configure the `ipath_ether` driver to run MPI jobs.

Optional configurable OpenFabrics components are:

■ IPoIB network interface

■ OpenSM

■ SRP. Provided as a technology demo in this release.

■ uDAPL. Can be used by Intel MPI

3.8
Configuring the InfiniPath Drivers

This section provides information on InfiniPath driver configuration.

3.8.1
Configuring the `ib_ipath` Driver

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

```
/etc/sysconfig/infinipath
```

Normally this configuration file is set up correctly at installation and the driver(s) are loaded automatically during system boot once the RPMs have been installed. Assuming that all the InfiniPath and OpenFabrics sofware has been installed, the default settings upon startup will be:

- InfiniPath `ib_ipath` is enabled
- InfiniPath `ipath_ether` is not running until configured
- OpenFabrics IPoIB is not running until configured
- OpenSM is enabled on startup. Disable it on all nodes except where it will be used as subnet manager.

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.

Specific additional instructions for enabling different drivers and modifying behaviors are presented in the following sections.

3.8.2
Configuring the `ipath_ether` Network Interface

These instructions are for enabling TCP-IP networking over the InfiniPath link. To enable IPoIB networking, see section 3.9.1.

You must create a network device configuration file for the layered Ethernet device on the InfiniPath adapter. This configuration file will resemble the configuration files for the other Ethernet devices on the nodes. Two slightly different procedures are given below for the ipath configuration; one for Fedora and one for SUSE or SLES 10.

Many of the entries that are used in the configuration directions below are explained in the file `sysconfig.txt`. To familiarize yourself with these, please see:

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

3.8.2.1
`ipath_ether` Configuration on Fedora and RHEL4

These configuration steps will cause the `ipath_ether` network interfaces to be automatically configured when you next reboot the system. These instructions are for the Fedora Core 3, Fedora Core 4 and Red Hat Enterprise Linux 4 distributions.

Typically on servers there are two Ethernet devices present, numbered as 0 (`eth0`) and 1 (`eth1`). This example assumes we create a third device, `eth2`.

NOTE: When multiple InfiniPath chips are present, the configuration for `eth3`, `eth4`, and so on follow the same format as for adding `eth2` in the examples below.

1. Check for the number of Ethernet drivers you currently have by either one of the two following commands :

   ```
   $ ifconfig -a
   $ ls /sys/class/net
   ```

   As mentioned above we assume that two Ethernet devices (numbered 0 and 1) are already present.

2. Edit the file `/etc/modprobe.conf` (as root) by adding the following line:

   ```
   alias eth2 ipath_ether
   ```

3. Create or edit the following file (as root).

   ```
   /etc/sysconfig/network-scripts/ifcfg-eth2
   ```

   If you are using DHCP (dynamic host configuration protocol), add the following lines to `ifcfg-eth2`:

   ```
   # QLogic Interconnect Ethernet
   DEVICE=eth2
   ONBOOT=yes
   BOOTPROTO=dhcp
   ```

   If you are using static IP addresses, use the following lines instead, substituting your own IP address for the sample one given here.The normal matching netmask is shown.

   ```
   # QLogic Interconnect Ethernet
   DEVICE=eth2
   BOOTPROTO=static
   ONBOOT=YES
   IPADDR=192.168.5.101 #Substitute your IP address here
   NETMASK="255.255.255.0"#Normal matching netmask
   TYPE=Ethernet
   ```

This will cause the ipath_ether Ethernet driver to be loaded and configured during system startup. To check your configuration, and make the ipath_ether Ethernet driver available immediately, use the command (as root):

```
# /sbin/ifup eth2
```

4. Check whether the Ethernet driver has been loaded with:

```
$ lsmod | grep ipath_ether
```

5. Verify that the driver is up with:

```
$ ifconfig -a
```

3.8.2.2
`ipath_ether` Configuration on SUSE 9.3, SLES 9, and SLES 10

These configuration steps will cause the `ipath_ether` network interfaces to be automatically configured when you next reboot the system. These instructions are for the SUSE 9.3, SLES 9 and SLES 10 distributions.

Typically on servers there are two Ethernet devices present, numbered as 0 (`eth0`) and 1 (`eth1`). This example assumes we create a third device, `eth2`.

NOTE: When multiple InfiniPath chips are present, the configuration for `eth3`, `eth4`, and so on follow the same format as for adding `eth2` in the examples below. Similarly , in step 2, add one to the unit number, so replace `.../00/guid` with `/01/guid` for the second InfiniPath interface, and so on.

Step 3 is applicable only to SLES 10; it is required because SLES 10 uses a newer version of the `udev` subsystem.

NOTE: The MAC address (media access control address) is a unique identifier attached to most forms of networking equipment. Step 2 below determines the MAC address to use, and will be referred to as $MAC in the subsequent steps. $MAC must be replaced in each case with the string printed in step 2.

The following steps must all be executed as the root user.

1. Be sure that the ipath_ether module is loaded:

```
# lsmod | grep -q ipath_ether || modprobe ipath_ether
```

2. Determine the MAC address that will be used:

```
# sed 's/^\(..:..:..\):..:../\1/' \
/sys/bus/pci/drivers/ib_ipath/00/guid
```

NOTE: Care should be taken when cutting and pasting commands such as the above from PDF documents, as quotes are special characters and may not be translated correctly.

The output should appear similar to this (6 hex digit pairs, separated by colons):

```
00:11:75:04:e0:11
```
The GUID can also be returned by running:

```
# ipath_control -i
$Id: QLogic Release2.0 $ $Date: 2006-10-15-04:16 $
00: Version: Driver 2.0, InfiniPath_QHT7140, InfiniPath1 3.2,
PCI 2, SW Compat 2
00: Status: 0xe1 Initted Present IB_link_up IB_configured
00: LID=0x30 MLID=0x0 GUID=00:11:75:00:00:04:e0:11 Serial:
1236070407
```

Note that removing the middle two `00:00` octets from the GUID in the above output will form the MAC address

If either step 1 or step 2 fails in some fashion, the problem must be found and corrected before continuing. Verify that the RPMs are installed correctly, and that `infinipath` has correctly been started. If problems continue, run `ipathbug-helper` and report the results to your reseller or InfiniPath support organization.

3. Skip to Step 4 if you are using SUSE 9.3 or SLES 9. This step is only done on SLES 10 systems. Edit the file:

```
/etc/udev/rules.d/30-net_persistent_names.rules
```

If this file does not exist, skip to Step 4.

Check each of the lines starting with `SUBSYSTEM=`, to find the highest numbered interface. (For standard motherboards, the highest numbered interface will typically be 1.)

Add a new line at the end of the file, incrementing the interface number by one. In this example, it becomes `eth2`. The new line will look like this:

```
SUBSYSTEM=="net", ACTION=="add", SYSFS{address}=="$MAC",
IMPORT="/sbin/ rename_netiface %k eth2"
```

This will appear as a single line in the file. $MAC is replaced by the string from step 2 above.

4. Create the network module file:

```
/etc/sysconfig/hardware/hwcfg-eth-id-$MAC
```
Add the following lines to the file:

```
MODULE=ipath_ether
STARTMODE=auto
```

This will cause the `ipath_ether` Ethernet driver to be loaded and configured during system startup.

5. Create the network configuration file:

```
/etc/sysconfig/network/ifcfg-eth2
```

If you are using DHCP (dynamically assigned IP addresses), add these lines to the file:

```
STARTMODE=onboot
BOOTPROTO=dhcp
NAME='InfiniPath Network Card'
_nm_name=eth-id-$MAC
```

Proceed to Step 6.

If you are you are using static IP addresses (not DHCP), add these lines to the file:

```
STARTMODE=onboot
BOOTPROTO=static
NAME='InfiniPath Network Card'
NETWORK=192.168.5.0
NETMASK=255.255.255.0
BROADCAST=192.168.5.255
IPADDR=192.168.5.211
_nm_name=eth-id-$MAC
```

Make sure that you substitute your own IP address for the sample IPADDR shown here. The BROADCAST, NETMASK, and NETWORK lines need to match for your network.

6. To verify that the configuration files are correct, you will normally now be able to run the commands:

```
# ifup eth2
# ifconfig eth2
```

Note that it may be necessary to reboot the system before the configuration changes will work.

3.9
OpenFabrics Configuration and Startup

In the prior InfiniPath 1.3 release the InfiniPath (ipath_core) and OpenFabrics (ib_ipath) modules were separate. In this release there is now one module, ib_ipath, which provides both low level InfiniPath support and management functions for OpenFabrics protocols. The startup script for ib_ipath is installed automatically as part of the software installation, and normally does not need to be changed.

The IPoIB network interface and OpenSM components of OpenFabrics can be configured to be on or off. IPoIB is off by default; OpenSM is on by default

IPoIB and OpenSM configuration is explained in greater detail in the following sections.

NOTE: The following instructions work for FC4, SUSE 9.3, SLES 9, and SLES 10.

3.9.1
Configuring the IPoIB Network Interface

Instructions are given here to manually configure your OpenFabrics IPoIB network interface. This example assumes that you are using `sh` or `bash` as your shell, and that all required InfiniPath and OpenFabrics RPMs are installed, and your startup scripts have been run, either manually or at system boot.

For this example, we assume that your 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, and that the host being configured has the IP address 10.1.17.3, and that DHCP is not being used.

NOTE: We supply instructions only for this static IP address case. Configuration methods for using DHCP will be supplied in a later release.

Type the following commands (as root):

ifconfig ib0 10.1.17.3 netmask 0xffffff00

To verify the configuration, type:

ifconfig ib0

The output from this command should 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
inet addr:10.1.17.3  Bcast:10.1.17.255 Mask:255.255.255.0
UP BROADCAST RUNNING MULTICAST  MTU:2044  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)
```

```
Next, type:
```

ping -c 2 -b 10.1.17.255

The output of the `ping` command should be similar to that below, 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.

NOTE: The configuration must be repeated each time the system is rebooted.

3.10
OpenSM

OpenSM is an optional component of the OpenFabrics project that provides a subnet manager for InfiniBand networks. You do not need to use OpenSM if any of your InfiniBand switches provide a subnet manager.

Note that after installing the `opensm` package, OpenSM is configured to be **on** on the next machine reboot. Install it on the machine that will act as a subnet manager in your cluster, or if it has been installed on more than one machine, use the `chkconfig` command (as root) to disable it on all the nodes *except* for the one on which it will be used. Use this method:

```
# chkconfig opensmd off
```

The command to enable it on reboot is:

```
# chkconfig opensmd on
```

You can start `opensmd` without rebooting your machine as follows:

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

and you can stop it again like this:

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

If you wish to pass any arguments to the OpenSM program, modify the file:

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

and add the arguments to the "OPTIONS" variable. Here is an example:

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

3.11
SRP

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

In this release SRP is provided as a technology preview. Add `ib_srp` to the module list in `/etc/sysconfig/infinipath` to have it automatically loaded.

NOTE: SRP does not yet work with IBM Power Systems. This will be fixed in a future release.

3.12
Starting and Stopping the InfiniPath Software

The InfiniPath driver software runs as a system service, normally started at system startup. Normally you will not need to restart the software, but you may wish to do so after installing a new InfiniPath release, or after changing driver options, or if doing manual testing.

The following commands can be used to check or configure state. These methods will not reboot the system.

To check the configuration state, use the command:

```
$ chkconfig --list infinipath
```

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

```
# chkconfig infinipath on 2345
```

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

```
# chkconfig infinipath off
```

NOTE: This does not stop and unload the driver, if it is already loaded.

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

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

This method will not reboot the system. The following set of commands shows how this script can be used. Please take note of the following:

■ You should omit the commands to start/stop `opensmd` if you are not running it on that node.

■ You should omit the `ifdown` and `ifup` step if you are not using `ipath_ether` on that node.

The sequence of commands to restart `infinipath` are given below. Note that this next example assumes that `ipath_ether` is configured as `eth2`.

```
# /etc/init.d/opensmd stop
# ifdown eth2
# /etc/init.d/infinipath stop
...

# /etc/init.d/infinipath start
# ifup eth2
# /etc/init.d/opensmd start
```

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

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

```
# /etc/init.d/opensmd stop
# ifdown eth2
# /etc/init.d/infinipath restart
# ifup eth2
# /etc/init.d/opensmd start
```

NOTE: Restarting InfiniPath will terminate any InfiniPath MPI processes, as well as any OpenFabrics processes that are running at the time. Processes using networking over ipath_ether will return errors.

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

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

You can check to see if ipath_ether is running by using the following command. If it prints no output, it is not running.

```
$ /sbin/lsmod | grep ipath_ether
```

If there is output, you should look at the output from this command to determine if it is configured:

```
$ /sbin/ifconfig -a
```

Finally, if you need to find which InfiniPath and OpenFabrics modules are running, try the following command:

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

3.13
Rebuilding or Reinstalling the Drivers

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/infinipath/drivers
# ./make-install.sh
# /etc/init.d/infinipath restart
```

To reinstall the InfiniPath kernel modules, do the following (as root):

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

3.14
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:

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

These topics are also covered in more detail in the *InfiniPath Interconnect User Guide*, in Section 2.

3.15
MPI Over uDAPL

Some MPI implementations can be run over uDAPL. uDAPL is the user mode version of the Direct Access Provider Library (DAPL). Examples of such MPI implementations are Intel MPI and one option on OpenMPI.

If you are running such an MPI implementation, the rdma_cm and rdma_ucm modules will need to be loaded. To test these modules, use these commands (as root):

```
# modprobe rdma_cm
# modprobe rdma_ucm
```

To ensure that the modules are loaded whenever the driver is loaded, add rdma_cm and rdma_ucm to the OPENFABRICS_MODULES assignment in /etc/sysconfig/infinipath.

3.16
Compiler Support

Presently, in addition to the QLogic PathScale Compiler Suite, InfiniPath MPI supports a number of other compilers. These include PGI 5.2 and 6.0, Intel 9.x, the GNU gcc 3.3.x, 3.4.x, and 4.x compiler suites and gfortran. The IBM XL family of compilers are also supported on ppc64 (Power) systems.

The QLogic PathScale Compiler Suite is not yet supported on Fedora Core 4, Fedora Core 5, or SLES 10, all systems that have the GNU 4.x compilers and compiler environment (header files and libraries).

However, it is possible to cross-compile on a distribution that supports the QLogic PathScale compiler, such as FC3 or SUSE 9.3, and then run on one of these other distributions that uses the GNU 4.x compiler and environment. For more information on cross-compilation, see Section 3.5.4 in the *InfiniPath Interconnect User Guide*.

gfortran is not currently supported on Fedora Core 3.

Please check the QLogic website for updated information on supported compilers.

3.17
Switch Configuration and Monitoring

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

3.18
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. The following table describes the two states of the LEDs. The green LED will normally illuminate first. The normal state is **Green On**, **Amber On**.

Table 3-4. LED Link and Data Indicators

LED	Color	Status	
Power	Green	ON Signal detected. Ready to talk to an SM to bring link fully up.	OFF Switch not powered up. Software not installed or started. Loss of signal. Check cabling.
Link	Amber	ON Link configured. Properly connected and ready to receive data and link packets.	OFF SM may be missing. Link may not be configured. Check the connection.

3.19
Customer Acceptance Utility

`ipath_checkout` is a `bash` script used to verify that the installation is correct and that all the nodes of the network are functioning and mutually connected by the InfiniPath fabric. It is to be run on a front end node, and requires specification of a hosts file:

$ **ipath_checkout [options]** *hostsfile*

where `hostsfile` designates a file listing the hostnames of the nodes of the cluster, one hostname per line. The format of `hostsfile` is as follows:

```
hostname1

hostname2

...
```

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

1. `ping` all nodes to verify all are reachable from the frontend.

2. `ssh` to each node to verify correct configuration of `ssh`.

3. Gather and analyze system configuration from nodes.

4. Gather and analyze RPMs installed on nodes.

5. Verify InfiniPath hardware and software status and configuration.

6. Verify ability to `mpirun` jobs on nodes.

7. Run bandwidth and latency test on every pair of nodes and analyze results.

The possible options to `ipath_checkout` are:

`-h, --help`
Displays help messages giving defined usage.

`-v, --verbose`
`-vv, --vverbose`
`-vvv, --vvverbose`
These specify three successively higher levels of detail in reporting results of tests. So, there are four levels of detail in all, including the case of where none these options are given.

`-c, --continue`
When not specified, the test terminates when any test fails. When specified, the tests continue after a failure, with failing nodes excluded from subsequent tests.

`--workdir=DIR`
Use DIR to hold intermediate files created while running tests. DIR must not already exist.

`-k, --keep`
Keep intermediate files that were created while performing tests and compiling reports. Results will be saved in a directory created by `mktemp` and named `infinipath_XXXXXX` or in the directory name given to `--workdir`.

`--skip=LIST`
Skip the tests in `LIST` (e.g. --skip=2,4,5,7 will skip tests 2, 4, 5, and 7)

`-d, --debug`
Turn on `-x` and `-v` flags in bash.

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

Also refer to appendix C (The Troubleshooting Appendix) in the InfiniPath Interconnect User Guide.

3.20

Removing Software Packages

Instructions for uninstalling or downgrading InfiniPath and OpenFabrics software is given below.

To uninstall the InfiniPath software packages on any node, using a `bash` shell, type the command (as root):

```
# rpm -e $(rpm -qa 'InfiniPath-MPI/mpi*' 'InfiniPath/infinipath*')
```

This will uninstall the InfiniPath and MPI software RPMs on that node.

To uninstall the OpenFabrics software packages on any node, using a `bash` shell, type the command (as root):

```
# rpm -e rpm_name_pre
```

`rpm_name_pre` is the descriptive name that precedes the version and repository identifiers in an RPM. For example:

```
# rpm -e libibverbs
```

This would uninstall `libibverbs-2.0-`*xxx*`.1_0_4.`*yyy*`.x86_64.rpm` on that node.

For both InfiniPath and OpenFabrics it is recommended that you remove all the packages at the same time.

3.20.1

Downgrading

If you want to downgrade, please remove both the InfiniPath and OpenFabrics RPMs, then install the older bits. We have found that rpm flags like "`--oldpackage`" will not generate a correct downgrade.

3.21

Installing Newer Drivers from Other Distributions

The driver source now resides in `infinipath-kernel`. This means that newer drivers can be installed as they become available. Those who wish to install newer drivers, for example, from OFED (Open Fabrics Enterprise Distribution), should be able to do so. However, some extra steps need to be taken in order to install properly.

1. Install all InfiniPath RPMs, including `infinipath-kernel`. The RPM `infinipath-kernel` installs into:

    ```
    /lib/modules/$(uname -r)/updates
    ```

    This should not affect any other installed InfiniPath or OpenFabrics drivers.

2. Reload the InfiniPath and OpenFabrics modules to verify that the installation works by using this command (as root):

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

3. Run ipath_checkout or other OpenFabrics test program to verify that the InfiniPath card(s) work properly.

4. Unload the InfiniPath and OpenFabrics modules with the command:

   # **`/etc/init.d/infinipath stop`**

5. Remove the InfiniPath kernel components with the command:

   $ **`rpm -e infinipath-kernel --nodeps`**

   The option `--nodeps` is required because the other InfiniPath RPMs depend on `infinipath-kernel`.

6. Verify that no InfiniPath or OpenFabrics modules are present in the `/lib/modules/$(uname -r)/updates` directory.

7. If not yet installed, install the InfiniPath and OpenFabrics modules from your alternate set of RPMs.

8. Reload all modules by using this command (as root):

   # **`/etc/init.d/infinipath start`**

   An alternate mechanism can be used, if provided as part of your alternate installation.

9. Run an OpenFabrics test program, such as `ibstatus`, to verify that your InfiniPath card(s) work correctly.

This information is repeated in appendix C in the *InfiniPath Interconnect User Guide*.

3.22
Installation Troubleshooting

If you have any difficulties installing the hardware or software, please see appendix C in the *InfiniPath Interconnect User Guide*. This Troubleshooting section contains information on problems related to installation, system administration, and InfiniPath MPI issues.

RPM Descriptions

The following sections show more detailed descriptions of the RPMs for InfiniPath and OpenFabrics.

InfiniPath and OpenFabrics RPMs

For ease of installation, it is recommended that all RPMs are installed on all nodes. However, some RPMs are optional. Since cluster nodes may be used for different functions, it is possible to selectively install. An example would be the `opensm` package for use on the node which 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.

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. We define these nodes in the following way:

- *Front end node*. This node will launch jobs. It is referred to as the *front end node* in this documentation.

- *Compute node*. These are the nodes that are used to run jobs.

- A *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 only one or just a few 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 A-1 for information on installation of software RPMs on specific types of nodes. Note that each set of RPMs uses a build identifier xxx and a distribution identifier yyy. The architecture is designated by `x86_64`, `noarch`, `i386`, or `ppc64`, and is dependent upon the distribution.

The RPMs are now organized in the following way:

```
InfiniPath_license, InfiniPath_release_notes, LEGAL (top level)
Documentation/
InfiniPath/
InfiniPath-Devel/
InfiniPath-MPI/
OpenFabrics/
OpenFabrics-Devel/
OpenSM/
OtherHCAs/
```

The tables below show the sample contents of each of the above directories; the architectures are associated with the `fc3` distribution identifier.

A.3
Documentation and InfiniPath RPMs

Table A-1. Documentation/RPMs

RPM	Front end	Compute	Development
`infinipath-doc-2.0-xxx_yyy.noarch.rpm` InfiniPath man pages and other documents	Optional	Optional	Optional
`mpi-doc-2.0-xxx_yyy.noarch.rpm` Man pages for MPI functions and other MPI documents	Optional	Optional	Optional
`ofed-docs-2.0-xxx.1_1.yyy.x86_64.rpm` OpenFabrics documentation	Optional	Optional	Optional

Table A-2. InfiniPath-Devel/RPMs

RPM	Front end	Compute	Development
`infinipath-devel-2.0-xxx_yyy.x86_64.rpm` Development files for InfiniPath	Optional	Optional	Optional
`mpi-devel-2.0-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 A-3. InfiniPath/RPMs

RPM	Front end	Compute	Development
`infinipath-2.0-`*xxx_yyy*`.x86_64.rpm` Utilities and source code InfiniPath configuration files	Optional	Required	Optional
`infinipath-kernel-2.0-`*xxx_yyy*`.x86_64.rpm` InfiniPath drivers,OpenFabrics kernel modules	Optional	Required	Optional
`infinipath-libs-2.0-`*xxx_yyy*`.i386.rpm` InfiniPath protocol shared libraries for 32-bit and 64-bit systems	Optional	Required	Optional

Table A-4. InfiniPath/RPMs

RPM	Front end	Compute	Development
`infinipath-2.0-`*xxx_yyy*`.x86_64.rpm` Utilities and source code InfiniPath configuration files	Optional	Required	Optional
`infinipath-kernel-2.0-`*xxx_yyy*`.x86_64.rpm` InfiniPath drivers,OpenFabrics kernel modules	Optional	Required	Optional
`infinipath-libs-2.0-`*xxx_yyy*`.i386.rpm` InfiniPath protocol shared libraries for 32-bit and 64-bit systems	Optional	Required	Optional

Table A-5. InfiniPath-MPI/RPMs

RPM	Front end	Compute	Development
`mpi-benchmark-2.0-`*xxx_yyy*`.x86_64.rpm` MPI benchmark binaries	Optional	Required[a]	Optional
`mpi-frontend-2.0-`*xxx_yyy*`.i386.rpm` MPI job launch scripts and binaries, including mpirun and MPD	Required	Required[b]	Optional
`mpi-libs-2.0-`*xxx_yyy*`.i386.rpm` Shared libraries for MPI	Optional	Required	Required

Table Notes

[a]Required if using `ipath_checkout`. Otherwise, installation is optional.

[b]Required on nodes that run `mpirun` or MPD. Otherwise installation is optional.

A.4
OpenFabrics RPMs

OpenFabrics and OpenSM are optional components. For ease of installation, it is recommended that the OpenFabrics RPMs listed below are installed on all nodes. The exception is the `opensm` package, which should only be installed on the node which will be used as a subnet manager.

Table A-6. OpenFabrics/RPMs

RPM name	Comments
`dapl-2.0-xxx.1_2_0.yyy.x86_64.rpm` uDAPL support	Optional for OpenFabrics[b]
`ibutils-2.0-xxx.1_0.yyy.x86_64.rpm` ibutils provides IB network and path diagnostics	Optional for OpenFabrics[b]
`ipoibtools-2.0-xxx.1_0.yyy.x86_64.rpm` IPoIB High Availability daemon	Optional for OpenFabrics
`libibcm-2.0-xxx.0_9_0.yyy.x86_64.rpm` Along with the OpenFabrics kernel drivers, libibcm provides a userspace InfiniBand Connection Managment API	Optional for OpenFabrics[a]
`libibcommon-2.0-xxx.1_0.yyy.x86_64.rpm` Common utility functions for the IB diagnostic and management tools	Required for OpenSM
`libibmad-2.0-xxx.1_0.yyy.x86_64.rpm` Low layer IB functions for use by the IB diagnostic and management programs. These include MAD, SA, SMP, and other basic IB functions.	Required for OpenSM
`libibumad-2.0-xxx.1_0.yyy.x86_64.rpm` Provides the user MAD library functions which sit on top of the user MAD modules in the kernel. These are used by the IB diagnostic and management tools, including OpenSM	Required for OpenSM
`libibverbs-2.0-xxx.1_0_4.yyy.x86_64.rpm` Library that allows userspace processes to use InfiniBand "verbs" as described in the InfiniBand Architecture Specification. This includes dierct hardware access for fast path operations. For this library to be usefule, a device-specfic plug-in module should also be installed.	Required for OpenFabrics
`libibverbs-utils-2.0-xxx.1_0_4.yyy.x86_64.rpm` Useful libibverbs example programs such as ibv_devinfo, which displays information about InfiniBand devices	Required for OpenFabrics
`libipathverbs-2.0-xxx.1_0.yyy.x86_64.rpm` Provides device-specific userspace driver for QLogic HCAs	Required for OpenFabrics

Table A-6. OpenFabrics/RPMs (Continued)

RPM name	Comments
`libopensm-2.0-xxx.2_0_0.yyy.x86_64.rpm` libopensm provides the library for OpenSM	Optional for OpenFabrics Used by OpenSM
`libosmcomp-2.0-xxx.2_0_0.yyy.x86_64.rpm` libosmcomp provides the OS component library for OpenSM	Optional for OpenFabrics Used by OpenSM
`libosmvendor-2.0-xxx.2_0_0.yyy.x86_64.rpm` libosmvendor provides the vendor library for OpenSM	Optional for OpenFabrics Used by OpenSM
`librdmacm-2.0-xxx.0_9_0.yyy.x86_64.rpm` Support for the new connection manager	Optional for OpenFabrics
`librdmacm-utils-2.0-xxx.0_9_0.yyy.x86_64.rpm` Utilities for the new connection manager	Optional for OpenFabrics
`libsdp-2.0-xxx.1_1_0.yyy.x86_64.rpm` Can be LD_PRELOAD-ed to have a sockets application use InfiniBand Sockets Direct Protocol (SDP) instead of TCP, transparently and without recompiling the application	Required for OpenFabrics
`ofed-scripts-2.0-xxx.1_1.yyy.x86_64.rpm` OpenFabrics scripts	Optional for OpenFabrics
`openib-diags-2.0-xxx.1_1_0.yyy.x86_64.rpm` Useful programs for troubleshooting and checking state.	Optional for OpenFabrics
`perftest-2.0-xxx.1_0.yyy.x86_64.rpm` IB Performance tests	Optional for OpenFabrics
`srptools-2.0-xxx.0_0_4.yyy.x86_64.rpm` Support for SRP	Optional for OpenFabrics

<u>Table Notes</u>

[a] `libibcm-utils-*`, the utilities for the `libibcm` library, are no longer installed. Checking for dependencies on this RPM is advised.

[b] Not available on ppc64 (Power)

Table A-7. OpenFabrics-Devel/RPMs

RPM name	Comments
`dapl-devel-2.0-xxx.1_2_0.yyy.x86_64.rpm` Development files for uDAPL support	Optional for OpenFabrics[b]
`libibcm-devel-2.0-xxx.0_9_0.yyy.x86_64.rpm` Development files for the libibcm library	Optional for OpenFabrics
`libibcommon-devel-2.0-xxx.1_0.yyy.x86_64.rpm` Development files for the libibcommon library	Optional for OpenFabrics[a]

Table A-7. OpenFabrics-Devel/RPMs

RPM name	Comments
`libibmad-devel-2.0-xxx.1_0.`*`yyy`*`.x86_64.rpm` Development files for the libibmad library	Optional for OpenFabrics
`libibumad-devel-2.0-xxx.1_0.`*`yyy`*`.x86_64.rpm` Development files for the libibumad library	Optional for OpenFabrics[a]
`libibverbs-devel-2.0-xxx.1_0_4.`*`yyy`*`.x86_64.rpm` Static libraries and header files for the libibverbs verbs library	Optional for OpenFabrics[b]
`libipathverbs-devel-2.0-xxx.1_0.`*`yyy`*`.x86_64.rpm` Static libraries and header files for the libibverbs verbs library	Optional for OpenFabrics
`libmthca-devel-2.0-xxx.1_0_3.`*`yyy`*`.x86_64.rpm` Static version of libmthca that may be linked directly to an application, which may be useful for debugging	Optional for OpenFabrics
`libopensm-devel-2.0-xxx.2_0_0.`*`yyy`*`.x86_64.rpm` Static component libraries for OpenSM	Optional for OpenFabrics
`libosmcomp-devel-2.0-xxx.2_0_0.`*`yyy`*`.x86_64.rpm` Static component libraries for OpenSM	Optional for OpenFabrics
`libosmvendor-devel-2.0-xxx.2_0_0.`*`yyy`*`.x86_64.rpm` Static vendor libraries for OpenSM	Optional for OpenFabrics
`librdmacm-devel-2.0-xxx.0_9_0.`*`yyy`*`.x86_64.rpm` Development files for the new connection manager.	Optional for OpenFabrics

Table Notes

[a] `libibcm-utils-*`, the utilities for the `libibcm` library, are no longer installed. Checking for dependencies on this RPM is advised.

[b] Only needed if compiling software against the packages.

Table A-8. OpenSM/RPMs

RPM name	Comments
`opensm-2.0-xxx.2_0_0.`*`yyy`*`.x86_64.rpm` OpenSM provides an implementation of an InfiniBand Subnet Manager and Administrator. At least one per each InfiniBand subnet is required in order to initialize the InfiniBand hardware.	Required for OpenSM

QLOGIC

Table A-9. OtherHCAs/RPMs

RPM name	Comments
`libmthca-2.0-xxx.1_0_3.`*yyy*`.x86_64.rpm` Provides a device-specific userspace driver for Mellanox HCAs (MT23108 InfiniHost and MT25208 InfiniHost III Ex) for use with the libibverbs library	Optional
`mstflint-2.0-xxx.1_0.`*yyy*`.x86_64.rpm` This package contains a burning tool for Mellanox manufactured HCA cards. It also provides access to the relevant source code.	Optional
`tvflash-2.0-xxx.0_9_0.`*yyy*`.x86_64.rpm` tvflash is used to query and update the firmware flash memory attached to Mellanox InfiniBand HCAs	Optional

Index