



Intel® Omni-Path Fabric Software

Installation Guide

Rev. 5.0

December 2016



You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning Intel products described herein. You agree to grant Intel a non-exclusive, royalty-free license to any patent claim thereafter drafted which includes subject matter disclosed herein.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest Intel product specifications and roadmaps.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Copies of documents which have an order number and are referenced in this document may be obtained by calling 1-800-548-4725 or visit <http://www.intel.com/design/literature.htm>.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at <http://www.intel.com/> or from the OEM or retailer.

No computer system can be absolutely secure.

Intel, the Intel logo, Intel Xeon Phi, and Xeon are trademarks of Intel Corporation in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others.

Copyright © 2016, Intel Corporation. All rights reserved.



Revision History

For the latest documentation, go to: www.intel.com/omnipath/FabricSoftwarePublications

Date	Revision	Description
Dec 2016	5.0	Document has been updated with the following: <ul style="list-style-type: none"> Removed all information pertaining to RHEL 7.1 Updated "Prepare Fabric Manager for Fabric Manager GUI" Added new section "Multi-Rail Usage" to Chapter "Additional Installation and Setup Tasks" Moved Appendix "Multi-Subnet Fabrics" under new Chapter "Additional Installation and Setup Tasks" Updated "HFI UEFI PXE Installation and Configuration" section Removed Appendix "hfdiags User Guide". Moved to <i>Intel® Omni-Path Fabric Host Software User Guide</i>.
Aug 2016	4.0	Document has been updated with the following: <ul style="list-style-type: none"> Added information on UEFI firmware Prerequisites Updated Install TUI menus Updated ./INSTALL command information to include <code>--user-space</code> Updated Fabric Manager GUI Installation information Updated "HFI UEFI PXE Installation and Configuration" section Updated "hfdiags User Guide" section
May 2016	3.0	Document has been updated with the following: <ul style="list-style-type: none"> Updated OS RPMs Installation Prerequisites Added MPI Compiler Prerequisites Added information about minimum disk space requirement for the Fabric Manager GUI installation Replaced "Client/Server Configuration to Boot Over Fabric" section with "HFI UEFI PXE Installation and Configuration"
February 2016	2.0	Document has been updated with the following: <ul style="list-style-type: none"> Updated OS RPMs Installation Prerequisites Added "Client/Server Configuration to Boot Over Fabric" appendix
November 2015	1.0	Initial Release



Contents

Revision History.....	3
Preface.....	8
Intended Audience.....	8
Documentation Set.....	8
Cluster Configurator for Intel® Omni-Path Fabric.....	9
Documentation Conventions.....	9
License Agreements.....	10
Technical Support.....	10
1.0 Overview.....	11
1.1 Installation Packages Overview.....	11
1.1.1 Intel® Omni-Path Fabric Host Software.....	11
1.1.2 Intel® Omni-Path Fabric Suite Software.....	14
1.1.3 Intel® Omni-Path Fabric Suite Fabric Manager GUI.....	15
2.0 Fabric Software Pre-Installation.....	16
2.1 Installation Prerequisites.....	16
3.0 Download and Extract Installation Packages.....	20
3.1 Download the Intel® Omni-Path Software.....	20
3.2 Unpack the Tar File.....	20
4.0 Install the Intel® Omni-Path Software.....	21
4.1 Before You Install.....	21
4.2 Intel® Omni-Path Software Installation Using the Install TUI.....	21
4.3 Intel® Omni-Path Software Installation Using CLI Commands.....	24
4.3.1 Install Command.....	25
4.4 IPoIB IPV4 Configuration Using the Software TUI.....	28
4.5 IPoIB IPV4 Configuration Using CLI Commands.....	29
4.6 Install IPoIB IPV6.....	29
5.0 Configure Intel® Omni-Path Chassis.....	31
5.1 Intel Chassis Configuration Prerequisites.....	31
5.2 Configure Chassis Using Intel® Omni-Path Fabric Suite FastFabric TUI.....	32
5.3 Configure Chassis Using CLI Commands.....	39
5.4 Verify Chassis Configuration Using CLI Commands.....	41
6.0 Configure Firmware on the Externally Managed Intel® Omni-Path Switches.....	44
6.1 Switch Configuration Prerequisites.....	44
6.2 Configure Externally Managed Switches Using FastFabric TUI.....	44
6.3 Configure Externally Managed Switches Using CLI Commands.....	51
7.0 Install Host Software on the Remaining Hosts Using the FastFabric TUI.....	54
8.0 Install Host Software on the Remaining Hosts Using CLI Commands.....	60
9.0 Verify Intel® Omni-Path Fabric Host Software on the Remaining Servers Using the FastFabric TUI.....	63



10.0 Verify Intel® Omni-Path Fabric Host Software on the Remaining Servers Using CLI Commands.....	69
11.0 Install Additional Management Nodes.....	74
12.0 Configure and Initialize Health Check Tools.....	76
13.0 Run High Performance Linpack 2 Benchmark.....	77
14.0 Install Intel® Omni-Path Fabric Suite Fabric Manager GUI.....	80
14.1 Windows* Installation.....	80
14.1.1 System Requirements for a Windows* Environment.....	80
14.1.2 Download the Intel® Omni-Path Fabric Suite Fabric Manager GUI.....	80
14.1.3 Use the Installation Wizard to Install the Fabric Manager GUI.....	81
14.1.4 Changing Database Location.....	84
14.2 Linux* Installation.....	84
14.2.1 System Requirements for a Linux* Environment.....	84
14.2.2 Download the Intel® Omni-Path Fabric Suite Fabric Manager GUI.....	85
14.2.3 Use RPM to Install the Fabric Manager GUI.....	85
14.2.4 Changing Database Location.....	85
14.3 Prepare Fabric Manager for Fabric Manager GUI.....	86
14.4 Start the Fabric Manager GUI Application.....	87
14.4.1 Windows* Procedure.....	87
14.4.2 Linux* Procedure.....	87
14.5 Configure Startup Options.....	87
14.6 Uninstall the Intel® Omni-Path Fabric Suite Fabric Manager GUI.....	88
14.6.1 Windows* Procedure.....	88
14.6.2 Linux* Procedure.....	88
14.7 Clearing the Fabric Manager GUI Cache.....	88
14.7.1 Cleanup Procedure for Windows*.....	88
14.7.2 Cleanup Procedure for Linux*.....	89
15.0 Additional Installation and Setup Tasks.....	90
15.1 Multi-Subnet Fabrics.....	90
15.1.1 Primarily Independent Subnets.....	90
15.1.2 Overlapping Subnets.....	92
15.2 Multi-Rail Usage.....	94
15.2.1 Multi-Rail Overview.....	94
15.2.2 Setting Up Single or Dual Rails for a Single Subnet.....	95
15.2.3 Configuring Dual Rails for Dual Subnets.....	97
16.0 Installation Verification and Additional Settings.....	100
16.1 LED Link and Data Indicators.....	100
16.2 Adapter and Other Settings.....	100
16.3 ARP Neighbor Table Setup for Large Clusters.....	100
16.4 SM Loop Test.....	101
17.0 Upgrade the Management Node.....	102
17.1 Before You Upgrade.....	102
17.2 Intel® Omni-Path Fabric Suite Upgrade.....	102
18.0 Upgrade the Fabric Software.....	105



19.0 Upgrade from IntelOPA-Basic to IntelOPA-IFS.....	108
Appendix A HFI UEFI PXE Installation and Configuration.....	110
A.1 Client Installation.....	110
A.2 Server Configuration.....	111
A.2.1 Intel® Omni-Path Software Installation.....	111
A.2.2 DHCP Server.....	111
A.2.3 Network Interface.....	113
A.2.4 TFTP Server.....	113
A.2.5 PXE Server.....	114
Appendix B UEFI Secure Boot Support.....	115
B.1 UEFI Firmware Support.....	115
B.2 Linux Support.....	115
Appendix C Software Installation Checklist.....	116
C.1 Pre-Installation Checklist.....	116
C.2 Install the Intel® Omni-Path Fabric Software on a Cluster.....	116



Tables

1	Performance Impact.....	66
---	-------------------------	----



Preface

This manual is part of the documentation set for the Intel® Omni-Path Fabric (Intel® OP Fabric), which is an end-to-end solution consisting of Intel® Omni-Path Host Fabric Interfaces (HFIs), Intel® Omni-Path switches, and fabric management and development tools.

The Intel® OP Fabric delivers a platform for the next generation of High-Performance Computing (HPC) systems that is designed to cost-effectively meet the scale, density, and reliability requirements of large-scale HPC clusters.

Both the Intel® OP Fabric and standard InfiniBand* are able to send Internet Protocol (IP) traffic over the fabric, or *IPoFabric*. In this document, however, it is referred to as *IP over IB* or *IPoIB*. From a software point of view, IPoFabric and IPoIB behave the same way and, in fact, use the same `ib_ipoib` driver to send IP traffic over the `ib0` and/or `ib1` ports.

Intended Audience

The intended audience for the Intel® Omni-Path (Intel® OP) document set is network administrators and other qualified personnel.

Documentation Set

The complete end user publications set for the Intel® Omni-Path product includes the following items.

- Hardware Documents:
 - *Intel® Omni-Path Fabric Switches Hardware Installation Guide*
 - *Intel® Omni-Path Fabric Switches GUI User Guide*
 - *Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide*
 - *Intel® Omni-Path Edge Switch Platform Configuration Reference Guide*
 - *Intel® Omni-Path Fabric Managed Switches Release Notes*
 - *Intel® Omni-Path Fabric Externally-Managed Switches Release Notes*
 - *Intel® Omni-Path Host Fabric Interface Installation Guide*
- Software Documents:
 - *Intel® Omni-Path Fabric Software Installation Guide*
 - *Intel® Omni-Path Fabric Suite Fabric Manager User Guide*
 - *Intel® Omni-Path Fabric Suite FastFabric User Guide*
 - *Intel® Omni-Path Fabric Host Software User Guide*
 - *Intel® Omni-Path Fabric Suite Fabric Manager GUI Online Help*
 - *Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide*



- *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*
- *Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide*
- *Intel® Omni-Path Fabric Performance Tuning User Guide*
- *Intel® Omni-Path Host Fabric Interface Platform Configuration Reference Guide*
- *Intel® Omni-Path Fabric Software Release Notes*
- *Intel® Omni-Path Fabric Manager GUI Release Notes*
- *Intel® Omni-Path Storage Router Design Guide*
- *Building Lustre* Servers with Intel® Omni-Path Architecture Application Note*
- *Intel® Omni-Path Fabric Staging Guide*

Documents are available at the following URLs:

- Intel® Omni-Path Switches Installation, User, and Reference Guides
<http://www.intel.com/omnipath/SwitchPublications>
- Intel® Omni-Path Host Fabric Interface Installation, User, and Reference Guides (includes software documents)
<http://www.intel.com/omnipath/FabricSoftwarePublications>
- Drivers and Software (including Release Notes)
<http://www.intel.com/omnipath/Downloads>

Cluster Configurator for Intel® Omni-Path Fabric

The Cluster Configurator for Intel® Omni-Path Fabric is available at: <http://www.intel.com/content/www/us/en/high-performance-computing-fabrics/omni-path-configurator.html>.

This tool generates sample cluster configurations based on key cluster attributes, including a side-by-side comparison of up to four cluster configurations. The tool also generates parts lists and cluster diagrams.

Documentation Conventions

The following conventions are standard for Intel® Omni-Path documentation:

- **Note:** provides additional information.
- **Caution:** indicates the presence of a hazard that has the potential of causing damage to data or equipment.
- **Warning:** indicates the presence of a hazard that has the potential of causing personal injury.
- Text in **blue** font indicates a hyperlink (jump) to a figure, table, or section in this guide. Links to websites are also shown in blue. For example:
See [License Agreements](#) on page 10 for more information.
For more information, visit www.intel.com.
- Text in **bold** font indicates user interface elements such as menu items, buttons, check boxes, key names, key strokes, or column headings. For example:



Click the **Start** button, point to **Programs**, point to **Accessories**, and then click **Command Prompt**.

Press **CTRL+P** and then press the **UP ARROW** key.

- Text in *Courier* font indicates a file name, directory path, or command line text. For example:

Enter the following command: `sh ./install.bin`

- Text in *italics* indicates terms, emphasis, variables, or document titles. For example:

Refer to *Intel® Omni-Path Fabric Software Installation Guide* for details.

In this document, the term *chassis* refers to a managed switch.

Procedures and information may be marked with one of the following qualifications:

- **(Linux)** – Tasks are only applicable when Linux* is being used.
- **(Host)** – Tasks are only applicable when Intel® Omni-Path Fabric Host Software or Intel® Omni-Path Fabric Suite is being used on the hosts.
- **(Switch)** – Tasks are applicable only when Intel® Omni-Path Switches or Chassis are being used.
- Tasks that are generally applicable to all environments are not marked.

License Agreements

This software is provided under one or more license agreements. Please refer to the license agreement(s) provided with the software for specific detail. Do not install or use the software until you have carefully read and agree to the terms and conditions of the license agreement(s). By loading or using the software, you agree to the terms of the license agreement(s). If you do not wish to so agree, do not install or use the software.

Technical Support

Technical support for Intel® Omni-Path products is available 24 hours a day, 365 days a year. Please contact Intel Customer Support or visit www.intel.com for additional detail.



1.0 Overview

The installation of the Intel® Omni-Path Software is accomplished using a Text User Interface (TUI) that guides you through the installation process. You have the option of using command line interface (CLI) commands to perform the installation or install rpms individually.

1.1 Installation Packages Overview

The following software installation packages are available for an Intel® Omni-Path Fabric:

- Intel® Omni-Path Fabric Host Software – This is the basic installation package that installs the Intel® Omni-Path Fabric Host Software components needed to set up compute, I/O, and Service nodes with drivers, stacks, and basic tools for local configuration and monitoring.
- Intel® Omni-Path Fabric Suite (IFS) Software – This installation package provides special features and includes the Intel® Omni-Path Fabric Host Software package, along with the Intel® Omni-Path Fabric Suite FastFabric Toolset (FastFabric) and the Intel® Omni-Path Fabric Suite Fabric Manager (Fabric Manager).
- Intel® Omni-Path Fabric Suite Fabric Manager GUI (Fabric Manager GUI) – This installation package provides a set of features for viewing and monitoring the fabric or multiple fabrics, and is installed on a computer outside of the fabric.

1.1.1 Intel® Omni-Path Fabric Host Software

The Intel® Omni-Path Fabric Host Software installation package (`IntelOPA-Basic.DISTRO.VERSION.tgz`) includes the following components:

- Intel® Omni-Path Architecture (Intel® OPA) Stack (Deltas relative to the distribution of the OpenFabrics Alliance* (OFA) stack)
- IBACM (Deltas relative to the distribution of the OFA stack)
- Intel HFI Components
- Intel® OPA Tools
- Intel® OPA OFA Development
- IP over IB (Deltas relative to the distribution of the OFA stack)
- MVAPICH2 (hfi, gcc)
- MVAPICH2 (hfi, Intel)
- OpenMPI (hfi, gcc)
- OpenMPI (hfi, Intel)
- GASNet (hfi, gcc)
- OpenSHMEM (hfi, gcc)
- MVAPICH2 (verbs, gcc)



- OpenMPI (verbs, gcc)
- MPI Source
- Pre-Boot components
- OFA Debug Info

Note: There is a separate Intel® Omni-Path Fabric Host Software installation package for each of the supported Linux* distributions. Refer to the release notes of the package version being installed for a list of supported Linux* distributions.

1.1.1.1 OS RPMs Installation Prerequisites

In addition to normal OS installation options, the following OS RPMs must be installed before installing the Intel® Omni-Path software.

Note: Use the version distributed with your OS distribution.

Red Hat Enterprise Linux* (RHEL*)

Install the following packages using yum from the RHEL* distribution.

- libibmad
- libibumad
- libibumad-devel
- libibverbs
- librdmacm
- libibcm
- libpfm.x86 (RHEL* 7.3)
- ibacm
- qperf
- perftest
- rdma
- infinipath-psm
- libhfi1
- expat
- elfutils-libelf-devel
- libstdc++-devel
- gcc-gfortran
- atlas
- tcl
- expect
- tcsh
- sysfsutils
- pciutils
- bc (command line calculator for floating point math)



- opensm-devel
- opensm-libs
- rpm-build
- redhat-rpm-config
- kernel-devel
- papi.x86 (RHEL* 7.3)

For compiling IFS software, the following packages are required.

- libibverbs-devel
- libibmad-devel
- librdmacm-devel
- ibacm-devel
- openssl-devel (1.0.1 or higher)
- libuuid-devel
- expat-devel
- infinipath-psm-devel
- valgrind-devel

SUSE* Linux* Enterprise Server (SLES*)

Install the following packages using zypper from the SLES* distribution.

- libibmad5
- libibumad3
- libibumad-devel
- libibverbs1
- librdmacm1
- libibcm1
- ibacm
- qperf
- perftest
- rdma
- opensm-devel
- opensm-libs3
- libpsm_infinipath1
- libexpat1
- libelf-devel
- gcc-gfortran
- libudev-devel
- texlive-latex
- texlive-babel-english



- texlive-psnfss
- texlive-courier
- bc (command line calculator for floating point math)
- rpm-build
- kernel-devel
- libstdc++-devel

For compiling IFS software, the following packages are required.

- libibverbs-devel
- librdmacm-devel
- ibacm-devel
- libopenssl-devel (1.0.1 or higher)
- libuuid-devel
- libexpat-devel
- infinipath-psm-devel
- rpm-build
- ibacm-devel
- libibumad-devel

Depending on which packages you choose, there may be additional prerequisites. For additional information, refer to the Release Notes for your specific release and installation type.

1.1.1.2 MPI Compiler Prerequisites

Two of the MPIS that are downloaded with the OPA Basic Software package are listed in the Intel OPA Install Menu, MVAPICH2 and OpenMPI were built with Intel compilers. To use these MPIS for compiling and running your applications and benchmarks, you will need to install a version of the Intel Parallel Studio XE 2016 suite that includes the C++, and/or Fortran compilers according to your needs. Intel recommends at a minimum to at least have the Intel Composer Runtime libraries installed to be able to run MPI applications. These runtime libraries may be obtained from: <https://software.intel.com/en-us/articles/redistributables-for-intel-parallel-studio-xe-2016-composer-edition-for-linux>.

1.1.2 Intel® Omni-Path Fabric Suite Software

The Intel® Omni-Path Fabric Suite installation package (`IntelOPA-IFS.DISTRO.VERSION.tgz`) installs the installation package listed in [Intel® Omni-Path Fabric Host Software](#) on page 11, and requires the same prerequisites listed in [OS RPMs Installation Prerequisites](#) on page 12. It also installs:

- FastFabric
- Intel® Omni-Path Fabric Suite Fabric Manager

For details on using the Intel® Omni-Path Fabric Suite FastFabric Toolset, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*. For details on using the Intel® Omni-Path Fabric Suite Fabric Manager, refer to the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide*.



Note: There is a separate IFS installation package for each of the supported Linux* distributions. Refer to the release notes of the version being installed for a list of supported Linux* distributions.

1.1.3 Intel® Omni-Path Fabric Suite Fabric Manager GUI

The Intel® Omni-Path Fabric Suite Fabric Manager GUI (Fabric Manager GUI) installation package installs the application for monitoring the fabric from a computer external to the fabric.



2.0 Fabric Software Pre-Installation

This section provides the information and procedures needed prior to installing, configuring, and verifying the fabric software. Typically, the Site Implementation Engineer performs the tasks described in this section to ensure that the fabric is ready for the software installation. To keep track of steps performed for the installation, a checklist is provided in [Software Installation Checklist](#) on page 116.

2.1 Installation Prerequisites

Design the Fabric

It is important that the design and installation of the hardware be planned carefully prior to the installation and setup of the fabric. The design plan must include the following information:

- Identification of servers that will function as the administration or Management Nodes, where the Intel® Omni-Path Fabric Suite (IFS) will be installed.
 - Server memory requirements based on the software being used:
 - IFS, including the Intel® Omni-Path Fabric Suite Fabric Manager, requires at least 1GB of physical memory for each Fabric Manager instance. When managing a cluster of 10000 nodes or more, 5GB of memory per Fabric Manager instance is required. For very large fabrics, with more than 16000 nodes, 15GB of memory per Fabric Manager instance is required.
 - When running multiple Fabric Manager instances on a single management node, the physical memory requirements should be multiplied by the number of Fabric Manager instances.
 - Swap disk space allowance should follow recommendations for the given version of Linux*. Swap space should be twice the size of the physical memory on the server running the Fabric Manager.
 - All servers with HFIs installed should have Unified Extensible Firmware Interface (UEFI) firmware installed. Refer to the [BIOS table](#), [Appendix B "HFI UEFI PXE Installation and Configuration"](#).

Note: Intel® Omni-Path Architecture Option ROM support requires UEFI firmware and the HFI drivers are based on UEFI instead of Legacy-BIOS.
- Plan the cabling of the fabric and create a cable planning spreadsheet using the sample .cvs files installed into `/usr/lib/opa/samples` on the management node. After a plan is established, `opaxlattopology` or `opaxlittopology_cust` may be used to convert the spreadsheet into a topology XML file that can be used by FastFabric during fabric verification, as well as being potentially used by the Fabric Manager for topology verification at runtime.



Note: The use of cable planning spreadsheets and the resulting topology XML files is highly recommended by Intel as an effective and efficient way to ensure the cluster is assembled and installed as intended. For more information, refer to the **opaxlattopology** or **opaxltopology_cust** sections of the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

- Plan the naming conventions for hosts and switches in the fabric. Intel recommends all switches and hosts be given unique names. Having unique names simplifies operations that are performed using host and switch names. Refer to Appendix C "Node Naming Recommendations" of the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for more information on naming the switches and hosts.

Set Up the Fabric

The following steps provide the information to set up the fabric. For information about the configuration files used by FastFabric, refer to *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

1. **(All)** Ensure all hardware is installed:

- Servers
- Core and edge switches
- Fabric cables

Note: When installing externally managed switches such as the Intel® Omni-Path Edge Switch 100 Series without a management module, the Node GUID could be required. The Node GUID is found on a label on the case of the switch and may be needed to configure and manage the switches with the IFS. You can also use the `opagenswitches` command to request the node GUIDS if there is a topology input file supplied. Intel recommends that you create a topology input file using `opaxlattopology`, `opaxlattopology_cust`, `opagentopology`, or other tools. In this case, the switch names and locations are figured out from the fabric and the GUIDs for each switch name are then determined automatically by tools such as `opagenswitches`.

2. **(All)** Ensure an Intel® Omni-Path Host Fabric Interface (HFI) is installed in each server. Refer to the *Intel® Omni-Path Host Fabric Interface Installation Guide* for instructions.
3. **(All)** The hardware configuration should be reviewed to ensure everything has been installed according to the plan. Following the software installation, Intel® Omni-Path Fabric Suite FastFabric tools may be used to help verify the installation.
4. **(Linux)** Ensure the required Operating System (OS) version (with the same kernel version) is installed on all hosts with the following options:
 - Root user command prompt ends in `"#"` or `"$"`.
Note: There must be a space after `"#"` or `"$"`.
 - Fancy and colored prompts must be disabled
 - TCL and Expect packages installed on all Fabric Management Nodes.

The Management Node(s) that run FastFabric should have a full installation and must include the Tcl and Expect OS RPMs.



For MPI clusters, install the C and Fortran compilers, along with their associated tools, on the Management Node(s).

Note: Refer to the *Intel® Omni-Path Fabric Software Release Notes* for a list of supported OS versions.

5. **(Linux)** Enable remote login as root on each host. In order for FastFabric to manage the hosts, the Management Node must be able to securely log in as root to each host. This can be accomplished using SSH.

Note: FastFabric includes the `opasetupssh` tool, which can help perform the key exchange to establish secure password-less login from the FastFabric node to all other nodes in the fabric. To simplify the use of this tool, Intel recommends initially configuring the same root password on all hosts. After root access through SSH has been set up using FastFabric, the root passwords can be changed.

6. **(All)** Resolve the TCP/IP Host Names.

FastFabric and TCP/IP must resolve host names to the management network and IPoIB IP addresses. If the management network is not IPoIB, each host must have both a management network name and an IPoIB network name. To do this, use the actual host name as the management network name and `HOSTNAME-opa` as the IPoIB network name, where `HOSTNAME` is the management network name of the given host.

Name resolution is accomplished by configuring a DNS server on the management network, with both management network and IPoIB addresses for each host and each Intel internally-managed chassis. Alternatively, an `/etc/hosts` file needs to be created on the Management Node; FastFabric can then propagate this `/etc/hosts` file to all the other hosts.

If using the `/etc/hosts` file approach:

- On the master node, add all the Ethernet* and IPoIB addresses into the `/etc/hosts` file.
- For the IPoIB convention, use `HOSTNAME-opa`.
- The `localhost` line should not be edited.
- The `/etc/hosts` file should not have any node-specific data.
- Copy the file to every node, as follows:

```
opascpall -p -f hostfile /etc/hosts /etc/hosts
```

Note: If DNS is being used, do not perform this task.

If using Domain Name System (DNS):

- Refer to the documentation for the domain name system (DNS) server being used. Make sure to edit the `/etc/resolv.conf` configuration file on the Management Node to use the proper DNS server.
- Refer to the Linux* OS documentation for more information about configuring the `/etc/resolv.conf` file. This file is typically configured during OS installation.
- If `/etc/resolv.conf` must be manually configured for each host, FastFabric can aid in copying the file to all the hosts.



- The `/etc/resolv.conf` file created on the Management Node must not have any node-specific data and must be appropriate for use on all hosts.
- Copying the `/etc/resolv.conf` file to all the nodes is accomplished during the OS installation.
- If the `/etc/resolv.conf` file was not set up on all the hosts during the OS installation, the **FastFabric Copy a file to all hosts** operation can be used during the [Install Host Software on the Remaining Hosts Using the FastFabric TUI](#) procedures to copy the `/etc/resolv.conf` file from the Management Node to all the other nodes.

7. **(All)** Set up a Network Time Protocol (NTP) server.

Configure an NTP server for the cluster, and set all Linux* hosts and internally managed chassis to sync to the NTP server. The setup of the internally managed chassis is described in [Configure Intel® Omni-Path Chassis](#) on page 31.

8. **(All)** Assign HFI Node Description Names.

Node Description names can be configured in many ways. For Intel® Omni-Path, Intel recommends the use of the `rdma-ndd` daemon (part of the `infiniband-diags` package) to keep the Node Description up to date with the `hostname` of the node. Once set up to assign node descriptions, `rdma-ndd` automatically assigns node descriptions to RDMA devices whenever a node is restarted, a RDMA device comes online, or the hostname changes. For details on `rdma-ndd`, see the man page.

Some distributions use `rdma-ndd` by default. For Intel® Omni-Path Fabric Suite-supported distributions that lack the `rdma-ndd` daemon, the latest version of `infiniband-diags` is supplied with IFS.



3.0 Download and Extract Installation Packages

This section discusses downloading the software package from an Intel web page or other Intel secured location, and unpacking the included tar file.

3.1 Download the Intel® Omni-Path Software

1. Using a web browser, type `downloadcenter.intel.com` in the address field and press Enter.
2. In the "Search downloads" field type `Omni-Path`.
3. From the search drop-down select the Intel® Omni-Path Fabric Software
4. In the Available Downloads list select the file(s) you need for the OS you have installed on your fabric.
5. Agree to the "Intel Software License Agreement".
6. Save the download to your hard drive.

3.2 Unpack the Tar File

Use the following procedure to unpack the tar file.

Basic

```
IntelOPA-Basic.DISTRO.VERSION.tgz
```

IFS

```
IntelOPA-IFS.DISTRO.VERSION.tgz
```

1. Open an SSH client session and log into the host where the package is being installed.
2. Copy the tar file to the `/root` directory.
3. Change directory to `/root`.

```
cd /root
```

4. Unpack the tar file.

Basic

```
tar xvfz IntelOPA-Basic.DISTRO.VERSION.tgz
```

IFS

```
tar xvfz IntelOPA-IFS.DISTRO.VERSION.tgz
```



4.0 Install the Intel® Omni-Path Software

This section provides information and procedures to install the Intel® Omni-Path Software on the Management Node or a host node in the fabric. There are two ways of installing the software:

- Install using the Install TUI (recommended). Refer to [Intel® Omni-Path Software Installation Using the Install TUI](#).
- Install using CLI commands. Refer to [Intel® Omni-Path Software Installation Using CLI Commands](#) on page 24.

All pre-installation requirements must be met prior to installing the software. When installing a fabric/cluster, Intel recommends installing the software on the Management Node using the Install TUI, and then using FastFabric to configure the Management Node. Once the Management Node has been configured, the Basic software can be installed on all the remaining hosts using the FastFabric TUI, or using a provisioning or diskless boot mechanism. If using a provisioning system, consult the documentation that comes with the provisioning system.

The [Software Installation Checklist](#) on page 116 provides a checklist for tracking the installation, configuration, and verification processes.

The installation, configuration, and verification procedures provide step-by-step instructions for a typical, single subnet fabric. For information about the installation and verification of multiple subnet fabrics, see [Multi-Subnet Fabrics](#) on page 90.

4.1 Before You Install

Refer to the Release Notes for a list of compatible operating systems.

4.2 Intel® Omni-Path Software Installation Using the Install TUI

The following procedure goes step-by-step through the installation and defines the different items for the basic and IFS installation. When needed, the procedure specifies **Basic** or **IFS** accordingly. This procedure presumes that you are logged in and have completed all pre-installation requirements.

Perform the following procedure to install the Intel® OP Software:

1. Change directory to `IntelOPA-[Basic|IFS].DISTRO.VERSION` directory.

Basic

```
cd IntelOPA-Basic.DISTRO.VERSION
```



IFS

```
cd IntelOPA-IFS.DISTRO.VERSION
```

2. Start the Install TUI:

```
./INSTALL
```

The **Intel OPA VERSION Software** main menu displays.

```
Intel OPA VERSION Software

1) Install/Uninstall Software
2) Reconfigure OFA IP over IB
3) Reconfigure Driver Autostart
4) Generate Supporting Information for Problem Report
6) FastFabric (Host/Chassis/Switch Setup/Admin)

X) Exit
```

3. Select 1) Install/Uninstall Software.

Screen 1 of 2 of the **Intel OPA Install Menu** is display.

```
Intel OPA Install (VERSION release) Menu

Please Select Install Action (screen 1 of 2):
0) OFA OPA Stack      [  Install  ][Available] VERSION
1) OFA IBACM          [  Install  ][Available] VERSION
2) Intel HFI Components[  Install  ][Available] VERSION
3) OPA Tools          [  Install  ][Available] VERSION
4) OFA OPA Development[  Install  ][Available] VERSION
5) FastFabric         [  Install  ][Available] VERSION
6) OFA IP over IB     [  Install  ][Available] VERSION
7) OPA FM             [  Install  ][Available] VERSION
8) MVAPICH2 (hfi,gcc) [  Install  ][Available] VERSION
9) MVAPICH2 (hfi,Intel)[  Install  ][Available] VERSION
a) OpenMPI (hfi,gcc)  [  Install  ][Available] VERSION
b) OpenMPI (hfi,Intel)[  Install  ][Available] VERSION
c) GASNet (hfi,gcc)   [  Install  ][Available] VERSION
d) OpenSHMEM (hfi,gcc)[  Install  ][Available] VERSION

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

Note: The previous menu is an example of the default IntelOPA-IFS Install TUI. The Basic Install TUI would show 6) FastFabric and 8) OPA FM as [Not Avail].

Note: Intel HFI Components selection contains the enhanced Intel® Omni-Path Host Fabric Interface (HFI) driver optimized stack for MPI (PSM2) on HFIs and OpenMPI, as well as user tools.

4. Review the items to be installed; the default value is in brackets (Install or Don't Install). To change a value, enter the alphanumeric character associated with the item.
5. Press N to go to the next screen.



Screen 2 of 2 of the **Intel OPA Install Menu** is displayed.

```
Intel OPA Install (VERSION release) Menu

Please Select Install Action (screen 2 of 2):
0) MVAPICH2 (verbs,gcc) [Don't Install][Available] VERSION
1) OpenMPI (verbs,gcc) [Don't Install][Available] VERSION
2) MPI Source [ Install ][Available] VERSION
3) Pre-Boot Components [Don't Install][Available] VERSION
4) OFA Debug Info [Don't Install][Available] VERSION

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

6. Review the items to be installed; the default value is in brackets (Install or Don't Install). To change a value, enter the alphanumeric character associated with the item.

7. Press **P** to perform the selected actions from the two screens.

The system prompts:

```
Preparing OFA VERSION release for Install...
Rebuild OFA SRPMs (a=all, p=prompt per SRPM, n=only as needed?) [n]:
```

8. Select the default by pressing **Enter**.

The system displays prompts that require your response as it is going through the installation.

9. For each system prompt, select the default by pressing **Enter**.

The **Intel OPA Autostart Menu** displays. The following is an example of the Intel OPA Autostart Menu for an IFS installation.

```
Intel OPA Autostart (VERSION release) Menu

Please Select Autostart Option:
0) OFA OPA Stack () [Enable ]
1) OFA IBACM (ibacm) [Disable]
2) Intel HFI Components () [Enable ]
3) OFA IP over IB () [Enable ]
4) OFA FM (opafm) [Disable]

P) Perform the autostart changes
S) Autostart All R) Autostart None
X) Return to Previous Menu (or ESC)
```

Note: The previous menu is an example of the IntelOPA-IFS Autostart Menu. The IntelOPA-Basic Autostart Menu does not show 4) OFA FM (opafm).

10. Review the items for Autostart; the default value is in brackets (Enable or Disable). To change a value, enter the alphanumeric character associated with the item.

Intel recommends leaving all of the Autostart selections set to the default values. However, on the Fabric Management Nodes where you plan to run the Intel® Omni-Path Fabric Suite Fabric Manager, make sure to set the OPA FM to enable autostart.

11. Press **P** to perform the selected actions from the screen.



The system prompts:

```
Hit any key to continue...
```

12. Press any key to proceed with the installation.

The system prompts:

```
Firmware Update unnecessary  
Hit any key to continue...
```

13. Press any key to proceed with the installation.

The system prompts:

```
A System Reboot is recommended to activate the software changes  
Hit any key to continue...
```

14. Press any key to proceed with the installation.

The installation completes and returns to the main menu.

15. Press **X** to exit.

16. Reboot the server.

If you have finished the IntelOPA-IFS installation of the Management Node, continue setting up your Management Node by proceeding to [Configure Intel® Omni-Path Chassis](#).

4.3 Intel® Omni-Path Software Installation Using CLI Commands

The `./INSTALL` command has many options including installing single components, and enabling and disabling autostart of components. For more information on the `./INSTALL` syntax and options, refer to the [Install Command](#) section.

The following procedure steps through the installation and defines the different items for the Basic and IFS installations. When needed, the procedure specifies **Basic** or **IFS** accordingly. This procedure presumes that you are logged in and have completed all pre-installation requirements.

Perform the following procedure to install the Intel® OP Software in a default configuration:

1. Change directory to `IntelOPA-[Basic|IFS].DISTRO.VERSION` directory.

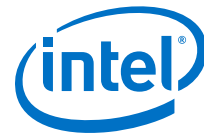
Basic

```
cd IntelOPA-Basic.DISTRO.VERSION
```

IFS

```
cd IntelOPA-IFS.DISTRO.VERSION
```

2. Start the default installation by typing `./INSTALL -n`.



The installation completes with the following system prompt:

```
A System Reboot is recommended to activate the software changes
Done Installing OPA Software.
```

3. Reboot the server.

```
reboot
```

If you have finished the IntelOPA-IFS installation of the Management Node, continue setting up your Management Node by proceeding to [Configure Intel® Omni-Path Chassis](#) on page 31.

4.3.1 Install Command

The `./INSTALL` command for the Basic and IFS installations are issued from the following directories:

- Intel Basic directory:

```
IntelOPA-Basic.DISTRO.VERSION
```

- Intel IFS directory:

```
IntelOPA-IFS.DISTRO.VERSION
```

Basic

```
cd IntelOPA-Basic.DISTRO.VERSION
```

IFS

```
cd IntelOPA-IFS.DISTRO.VERSION
```

Syntax

```
./INSTALL [-r root] [-v|-vv] [-a|-n|-U|-u|-s|-O|-N|-i comp|-e comp]
[-E comp] [-D comp] [--user-space] [--user_configure_options 'options']
[--kernel_configure_options 'options'] [--prefix dir]
[--without-depcheck] [--rebuild] [--force] [--answer keyword=value]
```

or

```
./INSTALL -C
```

or

```
./INSTALL -V
```

Options

No option selected

Displays the Intel OPA Software TUI.



<code>-a</code>	Installs all ULPs and drivers with the default options.
<code>-n</code>	Installs all ULPs and drivers with the default options, but does not change the autostart options.
<code>-U</code>	Upgrades/re-installs all presently installed ULPs and drivers with the default options, and does not change the autostart options.
<code>-i comp</code>	Installs the given component with the default options. This option can appear multiple times on a command line.
<code>-u</code>	Uninstalls all ULPs and drivers with the default options.
<code>-s</code>	Enables autostart for all installed software.
<code>-e comp</code>	Uninstalls the given component with the default options. This option can appear multiple times on a command line.
<code>-E comp</code>	Enables autostart of given component. This option can appear with <code>-D</code> or multiple times on a command line. <i>Note:</i> To control which installed software is configured for autostart, combine this option with <code>-a</code> , <code>-n</code> , <code>-i</code> , <code>-e</code> , and <code>-U</code> options.
<code>-D comp</code>	Disables autostart of the given component. This option can appear with <code>-E</code> or multiple times on a command line. <i>Note:</i> To control which installed software is configured for autostart, combine this option with <code>-a</code> , <code>-n</code> , <code>-i</code> , <code>-e</code> , and <code>-U</code> options.
<code>-v</code>	Verbose logging. Logs to the <code>/var/log/opa.log</code> file.
<code>-vv</code>	Very verbose debug logging. Logs to the <code>/var/log/opa.log</code> file.
<code>-C</code>	Shows the list of supported component names.
<code>-V</code>	Outputs the version number of the software.



<code>-r dir</code>	Specifies an alternate root directory. The default is <code>/</code> .
	<i>Note:</i> This option permits boot images to be constructed that include Intel® Omni-Path Software so that the boot images can later be used for network boot of Intel® Omni-Path Fabric enabled nodes.
	<i>Note:</i> Intel® Omni-Path Fabric Suite FastFabric use is not permitted in this environment.
<code>-O</code>	Keeps the current modified rpm configuration file.
<code>-N</code>	Uses a new default rpm configuration file.
<code>--rebuild</code>	Forces a rebuild of OFED_DELTA srpms.
<code>--user_queries</code>	Permits non-root users to query the fabric. This is the default.
<code>--no_user_queries</code>	Non-root users cannot query the fabric.
<code>--user-space</code>	Skip kernel space components during installation
<code>--user_configure_options options</code>	Specifies additional OFED build options for user space srpms. This forces a rebuild of all user srpms.
<code>--kernel_configure_options options</code>	Specifies additional OFED build options for driver srpms. This forces a rebuild of all driver srpms.
<code>--prefix dir</code>	Specifies an alternate directory prefix for the OFED_Delta installation. Default is <code>/usr</code> . This causes a rebuild of needed srpms.
<code>--without-depcheck</code>	Disables the check of OS dependencies.
<code>--force</code>	Forces the installation, even if the distributions do not match. Use of this option can result in undefined behaviors.
<code>--answer keyword=value</code>	Provides an answer to a question which might occur during the operation. Answers to questions that are not asked are ignored. Invalid answers result in prompting for interactive installations or use of the default for non-interactive installations.

Possible Questions:

`UserQueries` Allow non-root users to access the UMAD interface.



Note: Allowing access to `umadX` device files may present a security risk. However, this allows tools such as `opasaquery` and `opaportinfo` to be used by non-root users.

Default options retain existing configuration files.

Other Information

Supported component (`comp`) names:

```
opa_stack, ibacm, mpi_selector, intel_hfi, oftools, opa_stack_dev, fastfabric,
delta_ipoib, opafm, mvapich2_gcc_hfi, mvapich2_intel_hfi, openmpi_gcc_hfi,
openmpi_intel_hfi, gasnet, openshmem, mvapich2, openmpi, delta_mpirsrc,
hfi1_uefi, delta_debug
```

Supported component (`comp`) name aliases:

```
opa, ipoib, mpi, psm_mpi, verbs_mpi, pgas, mpisrc, opadev
```

4.4 IPoIB IPV4 Configuration Using the Software TUI

Prerequisites: Have a list of the IP addresses and netmasks for each interface you are going to set up.

The following procedure can be done during the Software TUI installation or after the installation is complete. If you are configuring the IPoIB IPV4 during the Software TUI installation, start with Step 2. When complete, go back to the Software TUI installation and continue where you left off.

The following procedure presumes you are logged on and at the **Intel OPA Software** main menu.

1. Select 2) Reconfigure OFA IP over IB from the **Intel OPA Software** main menu.

The system prompts:

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

2. Press **y** to configure OFA IP over IB IPV4.

The system prompts:

```
You may configure an OFA IP over IB interface for each HFI port
Or you may select to have OFA IP over IB only run on some HFI ports
Or you may select to configure redundant HFI ports with a
pair of HFI ports running a single OFA IP over IB interface
How many OFA IP over IB interfaces would you like to configure? [1]:
```



3. Press **Enter** for the default of 1 or type in the number of interfaces you are going to configure.

If you selected the default, the system prompts:

```
About to create OFA IP over IB ifcfg files in
/etc/sysconfig/network-scripts
Use interface name ib0? [y]:
```

If you selected something other than the default, the system prompts:

```
About to create OFA IP over IB ifcfg files in
/etc/sysconfig/network-scripts
Configure interface names sequentially starting with ib0? [y]:
```

4. Continue the setup by answering and following the prompts to set up the OFA IP over IB interfaces.
Setup completes and returns to the main menu.

4.5 IPoIB IPV4 Configuration Using CLI Commands

Prerequisites: Have a list of the IP addresses and netmasks for each interface you are going to set up.

The following procedure sets up the IPoIB IPV4 network interface for the local host and for the remote host.

1. Manually edit/create the `/etc/sysconfig/network-scripts/ifcfg-ibX` file using the OS distribution supplied instructions for setting up network interfaces.

The following is a sample of the `ifcfg-ibX`

```
DEVICE=ib0
TYPE=Infiniband
BOOTPROTO=dhcp
ONBOOT=yes
NM_CONTROLLED=no
CONNECTED_MODE=yes
MTU=65520
```

2. Type `opahostadmin -f /etc/sysconfig/opa/hosts configipoib`.
This creates the `ifcfg-ibX` configuration file for each host, using the host IP addresses from the `/etc/hosts` file.

4.6 Install IPoIB IPV6

To install IPoIB for IPV6 on the management node, use the following procedures for the particular OS installed on the node.

Red Hat* Enterprise Linux:

1. Edit file `/etc/sysconfig/network` to add the following line:

```
NETWORKING_IPV6=yes
```



2. Edit file `ifcfg-ifname` to add the following lines:

```
IPV6INIT=yes  
IPV6ADDR="ipv6addr/prefix-length"
```

IPv6 address should look like the following example:

```
3ffe::6/64
```

3. Restart the network.

SUSE* Linux Enterprise Server:

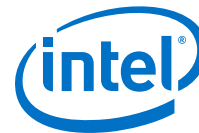
1. Edit `ifcfg-ifname` to add the following line:

```
IPADDR="ipv6addr/prefix-length"
```

IPv6 address should look like the following example:

```
3ffe::6/64
```

2. Restart the IPoIB network interfaces.



5.0 Configure Intel® Omni-Path Chassis

This section describes using the Intel® Omni-Path Fabric Suite FastFabric to install and configure internally managed switches, such as the Intel® Omni-Path Director Class Switch 100 Series. See the switch documentation for information about installing and configuring switches made by other manufacturers.

5.1 Intel Chassis Configuration Prerequisites

The *Intel® Omni-Path Fabric Switches Hardware Installation Guide* contains additional details about steps that must be performed before completing this procedure.

Configure the internally managed switches to use the Intel® Omni-Path Fabric Suite FastFabric tool set by performing the following steps:

1. **(Switch)** Connect each chassis to the management network through its Ethernet management port. For chassis with redundant management, connect both Ethernet management ports.
2. **(Switch)** Set up the netmask and gateway addresses on each Intel chassis, following the procedures in the *Intel® Omni-Path Fabric Switches GUI User Guide*.
3. **(Switch)** Assign each Intel chassis a unique IP address, and appropriately configure the Ethernet management port network settings.
4. **(Switch)** For a chassis with redundant management, assign a unique IP address for each Intel Management Module or Intel Management Spine, and configure their Ethernet management port network settings.
5. **(Switch)** Select a unique name for each Intel chassis, Management Module, and Spine. This name should be configured in DNS or `/etc/hosts` as the TCP/IP name for the Ethernet management port.

Note: The chassis node description is set later in this process.

6. **(Switch)** Configure the administrator password on each Intel chassis.

Note: All versions of Intel® Omni-Path Chassis 100 Series firmware permit SSH keys to be configured within the chassis for secure password-less login. Intel® Omni-Path Fabric Suite FastFabric includes a tool (`opasetupssh`) that can help perform the key exchange to establish secure password-less login from the Management Node to all the internally managed switches in the fabric. For simplicity, Intel recommends you initially configure the same administrator password on all switches. (Another method is to leave the default “adminpass” until FastFabric has set up SSH keys.) After SSH has been set up using FastFabric, Intel recommends you change the admin password.

7. **(Switch)** Copy the relevant chassis firmware files onto the Management Node. When performing the Chassis Configuration procedures, the `*.spkg` files are used to upgrade the firmware on each chassis.



Note: Place all files at a given firmware level into a single directory. Intel recommends that you include the firmware revision number in the directory name.

5.2 Configure Chassis Using Intel® Omni-Path Fabric Suite FastFabric TUI

Refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide* for more information on how to use the FastFabric TUI.

1. **(Switch)** Type `opafastfabric` and press **Enter**.

The **Intel FastFabric OPA Tools** menu is displayed.

```
Intel FastFabric OPA Tools
Version: X.X.X.X.X

  1) Chassis Setup/Admin
  2) Externally Managed Switch Setup/Admin
  3) Host Setup
  4) Host Verification/Admin
  5) Fabric Monitoring

X) Exit
```

2. **(Switch)** Press **1**.

The **FastFabric OPA Chassis Setup/Admin Menu** is displayed.

```
FastFabric OPA Chassis Setup/Admin Menu
Chassis File: /etc/sysconfig/opa/chassis

Setup:
0) Edit Config and Select/Edit Chassis File [ Skip ]
1) Verify Chassis via Ethernet Ping [ Skip ]
2) Update Chassis Firmware [ Skip ]
3) Setup Chassis Basic Configuration [ Skip ]
4) Setup Password-Less SSH/SCP [ Skip ]
5) Reboot Chassis [ Skip ]
6) Get Basic Chassis Configuration [ Skip ]
7) Configure Chassis Fabric Manager (FM) [ Skip ]
8) Update Chassis FM Security Files [ Skip ]
9) Get Chassis FM Security Files [ Skip ]
Admin:
a) Check OPA Fabric Status [ Skip ]
b) Control Chassis Fabric Manager (FM) [ Skip ]
c) Generate All Chassis Problem Report Info [ Skip ]
d) Run a Command on All Chassis [ Skip ]
Review:
e) View opachassisadmin Result Files [ Skip ]

P) Perform the Selected Actions N) Select None
X) Return to Previous Menu (or ESC)
```

3. **(Switch)** Select the items in the **Setup** section that are required or needed for your fabric.

Type the alphanumeric character associated with the item to toggle the selection from **Skip** to **Perform**.

4. Press **P**.

Perform the items that were selected in the sub-sections as follows.



Edit the Configuration and Select/Edit Chassis File

(Switch) The **Edit Config and Select/Edit Chassis File** selection allows the chassis, ports, and FastFabric configuration files to be edited.

- When placed in the editor for `opafastfabric.conf`, review the settings.
 - Especially review the `FF_CHASSIS_LOGIN_METHOD` and `FF_CHASSIS_ADMIN_PASSWORD`. Refer to "Configuration Files for FastFabric" section of the *Intel® Omni-Path Fabric Suite FastFabric User Guide* for more information about the `opafastfabric.conf` file.

Note: FastFabric provides the opportunity to enter the chassis password interactively when needed. It is not necessary to place it within `opafastfabric.conf`. If the Intel chassis admin password is placed in `opafastfabric.conf`, change the `opafastfabric.conf` permissions to be 0x600 (root-only access).

Note: All versions of Intel® Omni-Path Chassis 100 Series firmware permit SSH keys to be configured within the chassis for secure password-less login. There is no need to configure a `FF_CHASSIS_ADMIN_PASSWORD`, and `FF_CHASSIS_LOGIN_METHOD` can be set to SSH (the default).

- Select the location for the result files from FastFabric with the `FF_RESULT_DIR` parameter. The default is the directory from which a given session of FastFabric is invoked. Alternatively, it can be set to a directory relative to your home directory. For example:

```
export FF_RESULT_DIR=${FF_RESULT_DIR:-$HOME/
fastfabric_results}
```

Note: Refer to "Configuration Files for FastFabric" section of the *Intel® Omni-Path Fabric Suite FastFabric User Guide* for more information about the `opafastfabric.conf` file.

- When placed in the editor for `ports`, review the file. For typical single-subnet clusters, the default of "0:0" may be used. This uses the first active port on the Management Node to access the fabric. For more information on configuring a cluster with multiple subnets, see [Multi-Subnet Fabrics](#) on page 90. For further details about the `ports` file format, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.
- When placed in the editor for `chassis`, create the file with a list of the chassis names (the TCP/IP Ethernet management port names assigned) or IP addresses. Intel recommends you use chassis names. Enter one chassis name or IP address per line. For example:

```
Chassis1
Chassis2
```

Note: Do not list externally managed switches in this file. Those are covered in section [Configure Firmware on the Externally Managed Intel® Omni-Path Switches](#) on page 44.

For further details about the `chassis` file format, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.



Verify Chassis via Ethernet Ping

(Switch) The **Verify Chassis via Ethernet Ping** selection pings each selected chassis over the management network. If all chassis were found, continue to the next step. If some chassis were not found, exit the menu and review the following list for those chassis which were not found:

- Is chassis powered on and booted?
- Is chassis connected to management network?
- Are chassis IP address and network settings consistent with DNS or `/etc/hosts`?
- Is Management node connected to the management network?
- Are Management node IP address and network settings correct?
- Is management network itself up (including switches, routers, and others)?
- Is correct set of chassis listed in the chassis file? You may need to repeat the previous step to review and edit the file.

Update Chassis Firmware

(Switch) The **Update Chassis Firmware** selection permits the chassis firmware version to be verified and updated as needed.

Notes:

Before continuing, refer to the relevant switch release notes for any prerequisites:

- *Intel® Omni-Path Fabric Managed Switches Release Notes*
- *Intel® Omni-Path Fabric Externally-Managed Switches Release Notes*

1. When this procedure is started, the following system prompt is displayed:

```
Multiple Firmware files and/or Directories may be space separated
Shell wildcards may be used
For Directories all .spkg files in the directory tree will be used
Enter Files/Directories to use (or none):
```

2. Specify the directory where the relevant firmware files have been stored and press **Enter**.

This can be the mount point of the CD or the directory to which the files were copied in a previous step.

System prompts:

```
Would you like to run the firmware now? [n]:
```

3. Type **y** and press **Enter**.

FastFabric ensures that all chassis are running the firmware level provided, and installs and/or reboots each chassis as needed.

If any chassis fails to be updated, use the **View opachassisadmin Result Files** option to review the result files from the update. Refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for more details.

Set Up Chassis Basic Configuration

(Switch) The **Setup Chassis Basic Configuration** permits the typical chassis setup operations to be performed for all chassis.



1. Follow the system prompts and accept the defaults.
2. Continue with [Set up Password-Less SSH/SCP](#).

Set up Password-Less SSH/SCP

(Switch) The **Set up Password-Less SSH/SCP** selection sets up secure password-less SSH, such that the Management Node can securely log into all the chassis as `admin` through the management network, without requiring a password.

Reboot Chassis

(Switch) The **Reboot Chassis** selection reboots all the selected chassis and ensures they reboot fully (as verified through ping over the management network). When the chassis come back up following the reboot, they are running with all the new configuration settings.

Get Basic Chassis Configuration

(Switch) The **Get Basic Chassis Configuration** selection retrieves basic information from chassis such as syslog, NTP configuration, time zone, node description, and other information. The following is an example of the output from this selection:

```
Performing Chassis Admin: Get basic Chassis configuration
Executing: /usr/sbin/opachassisadmin -F /etc/sysconfig/opa/chassis getconfig
Executing getconfig Test Suite (getconfig) day mmm dd hh:mm:ss timezone yyyy ...
Executing TEST SUITE getconfig CASE (getconfig.xx.xx.xx.xx.getconfig) get
xx.xx.xx.xx ...
TEST SUITE getconfig CASE (getconfig.xx.xx.xx.xx.getconfig) get xx.xx.xx.xx
xx.xx.xx.xx:
  Firmware Active       : xx.xx.xx.xx
  Firmware Primary      : xx.xx.xx.xx
  Syslog Configuration  : Syslog host set to: 0.0.0.0 port 514 facility 22
  NTP                   : Configured to use the local clock
  Time Zone             : Time zone offset has not been configured
  LinkWidth Support     : 4X
  Node Description      : Node_Name
  Link CRC Mode         : 48b_or_14b_or_16b
PASSED
TEST SUITE getconfig: 1 Cases; 1 PASSED
TEST SUITE getconfig PASSED
Done getconfig Test Suite day mmm dd hh:mm:ss timezone yyyy

Hit any key to continue (or ESC to abort)...
```

Configure Chassis Fabric Manager (FM)

(Switch) The **Configure Chassis Fabric Manager (FM)** selection configures the Fabric Manager for any Intel® Omni-Path Chassis 100 Series.

System prompts:

```
Performing Chassis Admin: Configure Chassis Fabric Manager
Enter FM Config file to use (or none or generate):
```

1. Enter generate.

This performs the `config_generate` operation to guide you through selecting FM configuration options. See the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for more information about `config_generate`.



2. After responding to the prompts for `config_generate`, the following system prompt is displayed:

```
You have selected to use: ./opafm.xml
Syntax Checking ./opafm.xml...
Executing: /opt/hfi/fm_tools/config_check -s -c ./opafm.xml
Valid FM Config file: ./opafm.xml
After push, the FM may be started/restarted
Would you like to restart the FM? [n]:
```

3. Enter `y`.

This causes the FM to be started with the new configuration.

System prompts:

```
Would you like to run the FM on slave MMs? [n]:
```

4. Refer to the following If/Then table:

If	Then
Your fabric has a single chassis running the Fabric Manager. You can run the Fabric Manager on the slave management module (MM). This causes the Fabric Manager to be started in the applicable chassis.	Enter <code>y</code>
Your fabric has multiple chassis running the Fabric Manager. Intel recommends you run Fabric Manager on the master management module. This causes the Fabric Manager to be started only on the master management module in the applicable chassis.	Enter <code>n</code>

System prompts:

```
There will be a disruption as FMs are restarted
Doing the operation in parallel (on multiple chassis) will finish the fastest
Doing it serially may reduce disruption
Would you like to do the operation in parallel? [y]:
```

5. Press **Enter** to select the default option: `y`.

Intel recommends doing the operation in parallel.

System prompts:

```
You have selected to perform the push and FM restart in parallel
Would you like to enable FM start at boot? [n]:
```

6. Enter `y`.

This causes the Fabric Manager to be started on all applicable chassis each time those chassis boot.

System prompts:

```
Would you like to enable FM start on slave MMs at boot? [n]:
```

7. Refer to the following If/Then table:



If	Then
Your fabric has a single chassis running the Fabric Manager. You can run the Fabric Manager on the slave management module. This causes the Fabric Manager to be started in the applicable chassis.	Enter y
Your fabric has multiple chassis running the Fabric Manager. Intel recommends you run Fabric Manager on the master management module. This causes the Fabric Manager to only be started on the master management module in the applicable chassis.	Enter n

System prompts:

```
Would you like to be prompted for chassis' password? [n]:
```

8. Press **Enter** to select the default n option.

System prompts:

```
Are you sure you want to proceed? [n]:
```

9. Enter y.

This updates the Fabric Manager.

System prompts:

```
Hit any key to continue (or ESC to abort)...
```

10. Press any key to complete this procedure.

Update Chassis FM Security Files

(Switch) The **Update Chassis FM Security Files** selection runs the `opachassisadmin fmsecurityfiles` command to verify and update the chassis security files.

Note: The FM security files are the private key, public key, and certificate files required by the FM, in order to support secure socket connection via OpenSSL. Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for instructions on the administration tasks required to support these files.

Prompts guide you through the options:

- `push` - Ensures given security files are pushed to each chassis.
- `restart` - After push, restart FM on master, stop on slave.
- `restartall` - After push, restart FM on all MM.

Additional options prompted for:

- Selection of security files or directory containing pem files
- Parallel vs. serial update
- Chassis password (default is to have password in `fastfabric.conf` or to use password-less SSH)



If any chassis fails to be updated, use the `View opachassisadmin results files` option to review the result files from the update. Refer to [View opachassisadmin Result Files](#) on page 39 for more information.

Refer to *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for more details on the `opachassisadmin update` command.

Get Chassis FM Security Files

(Switch) The **Get Chassis FM Security Files** selection runs the `opachassisadmin fmgetsecurityfiles` command to retrieve the chassis FM security files from the chassis.

Check OPA Fabric Status

(All) The **Check OPA Fabric Status** selection prompts you to:

- Perform a fabric error analysis.
- Clear the error counters after generating a report.
- Perform a fabric link speed error analysis.
- Check for links configured to run slower than supported.
- Check for links connected with mismatched speed potential.
- Enter a filename for the results or save the results to the default location, which is: `/root/ffres/linkanalysis.res`

For more information, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

Control Chassis Fabric Manager (FM)

(Switch) The **Control Chassis Fabric Manager (FM)** selection prompts you to:

- Restart the Fabric Manager.
- Run the Fabric Manager on slave management modules.
- Restart the Fabric Manager on all management modules.
- Perform this operation in parallel.
- Change the Fabric Manager boot state to enable the Fabric Manager to start at boot.
- Enable Fabric Manager to start on slave management modules at boot.
- Be prompted for chassis password.

For more information, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

Generate All Chassis Problem Report Information

(Switch) The **Generate All Chassis Problem Report Info** selection generates the chassis problem report. For more information, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.



Run a Command on All Chassis

(Switch) If there are any other operations that need to be performed on all chassis, they may be performed using the **Run a Command on All Chassis** option. Each time this is executed, a single chassis CLI command may be specified to be executed against all selected chassis. When using these commands, additional setup or verification of the chassis may be performed.

View opachassisadmin Result Files

(Switch) The **View opachassisadmin Result Files** selection opens the `punchlist.csv`, `test.res`, and `test.log` files to be viewed. For more information, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

5.3 Configure Chassis Using CLI Commands

Prerequisites: Refer to [Intel Chassis Configuration Prerequisites](#) on page 31 for details.

This procedure provides step-by-step information for configuring the chassis using CLI commands.

1. (Optional) Edit or review the `/etc/sysconfig/opa/opafastfabric.conf` file.

Review the complete file. Especially review the `FF_CHASSIS_LOGIN_METHOD` and `FF_CHASSIS_ADMIN_PASSWORD` sections. Refer to "Configuration Files for FastFabric" section of the *Intel® Omni-Path Fabric Suite FastFabric User Guide* for more information about the `opafastfabric.conf` file.

2. (Optional) Edit or review the `/etc/sysconfig/opa/ports` file.

For typical single-subnet clusters, the default of "0:0" may be used. This uses the first active port on the Management Node to access the fabric. For more information on configuring a cluster with multiple subnets, see [Multi-Subnet Fabrics](#) on page 90. For further details about the file format, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

3. (Optional) Create, edit, or review the `/etc/sysconfig/opa/chassis` file.

Create/review the file with a list of the chassis names (the TCP/IP Ethernet management port names assigned) or IP addresses. Intel recommends you use names. Enter one chassis name or IP address per line. For example:

```
Chassis1
Chassis2
```

Note: Do not list externally managed switches in this file. Those are covered in [Configure Firmware on the Externally Managed Intel® Omni-Path Switches](#) on page 44.

For further details about the `chassis` file format, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

4. Verify the chassis is accessible over the management network using the command:

```
opapingall -C -p -f /etc/sysconfig/opa/chassis
```

If some chassis were not found, review the following list for those chassis:



- Is chassis powered on and booted?
- Is chassis connected to management network?
- Are chassis IP address and network settings consistent with DNS or `/etc/hosts`?
- Is Management node connected to the management network?
- Are Management node IP address and network settings correct?
- Is the management network itself up (including switches, routers, and others)?
- Is correct set of chassis listed in the chassis file? You may need to repeat the previous step to review and edit the file.

5. Update the chassis firmware.

Parallel method:

```
opachassisadmin -S -F chassisfile -P package -a run upgrade
```

Serial method:

```
FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -P package  
-a run upgrade
```

During the firmware update, the following files are produced:

- `test.res` - appended with summary results of run
- `test.log` - appended with detailed results of run
- `save_tmp/` - contains a directory per failed operation with detailed logs
- `test_tmp*/` - intermediate result files while operation is running

If any chassis fails to be updated, refer to the **View opachassisadmin result files** section to review the result files from the update.

Refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for more details.

6. Set up the chassis with the basic configuration settings using the command:

```
opachassisadmin -S -F chassisfile configure
```

This wizard gives you the option of setting up the following items on the chassis:

- Syslog server
- NTP server
- Timezone and DST information
- Chassis link width
- Configure Node Desc to match Ethernet* chassis name
- Configure the Link CRC Mode

Follow the system prompts to select and configure the items required.

7. Set up secure password-less SSH, such that the Management Node can securely log into all the chassis as `admin` through the management network, without requiring a password. Use the command:

```
opasetupssh -p -S -C -F chassisfile
```




8. Reboot all the selected chassis and ensure they reboot fully, as verified through ping over the management network. Use the command:


```
opachassisadmin -S -F chassisfile reboot
```

When the chassis come back up following the reboot, they are running with all the new configuration settings.
9. Retrieve basic information from chassis such as syslog, NTP configuration, time zone, node description, and other information. Use the command:


```
opachassisadmin -F chassisfile getconfig
```
10. Configure the Chassis Fabric Manager for any Intel® Omni-Path Chassis 100 Series chassis.

Parallel method:

```
opachassisadmin -S -F chassisfile -P package -a run fmconfig
```

Serial method:

```
FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -P package -a run fmconfig
```
11. Update Chassis Fabric Manager Security Files to permit the chassis security files to be verified and updated as needed.

Note: The FM security files are the private key, public key, and certificate files required by the FM, to support secure socket connection via OpenSSL. Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for instructions on the administration tasks required to support these files.

Parallel method:

```
opachassisadmin -S -F chassisfile -s securityFiles -a push fmsecurityfiles
```

Serial method:

```
FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -s securityFiles -a push fmsecurityfiles
```
12. Retrieve the Chassis Fabric Manager Security Files from the chassis using the command:


```
opachassisadmin -F chassisfile fmgetsecurityfiles
```

5.4 Verify Chassis Configuration Using CLI Commands

Prerequisites: Chassis has been configured.

This procedure provides step-by-step information to verify the chassis configuration and view the result files.

1. Check the OPA fabric status using the command:


```
opalinkanalysis -U -x snapshot_suffix reports verifyall > results_file 2>&1
```

This step performs the following operations:

- Perform a fabric error analysis.
- Clear the error counters after generating a report.



- Perform a fabric link speed error analysis.
 - Check for links configured to run slower than supported.
 - Check for links connected with mismatched speed potential.
 - Enter a filename for the results or save the results to the default location which is: /root/ffres/linkanalysis.res
2. To control the Chassis Fabric Manager, you can perform any or all of the following steps:

- a. (Optional) Restart all of the Chassis Fabric Managers.

Note: Using the serial method may reduce disruption.

Restart all the Chassis Fabric Managers and run the Fabric Manager on the slave Management Module using:

Parallel method:

```
opachassisadmin -S -F chassisfile -a restartall fmcontrol
```

Serial method:

```
FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -a  
restartall fmcontrol
```

There is a disruption as FMs are restarted.

- b. (Optional) Restart the master Chassis Fabric Manager and stop the slaves.

Note: Using the serial method may reduce disruption.

Restart the Master Chassis Fabric Managers and stop the Fabric Manager on the slave Management Modules using:

Parallel method:

```
opachassisadmin -S -F chassisfile -a restart fmcontrol
```

Serial method:

```
FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -a  
restart fmcontrol
```

- c. (Optional) Stop the Chassis Fabric Managers from running.

Stop the Master Chassis Fabric Managers from running using:

Parallel method:

```
opachassisadmin -S -F chassisfile -a stop fmcontrol
```

Serial method:

```
FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -a stop  
fmcontrol
```

- d. (Optional) Make sure all of the Chassis Fabric Managers are running.

Ensure all of the Chassis Fabric Managers are running, including the slaves, using:

Parallel method:

```
opachassisadmin -S -F chassisfile -a runall fmcontrol
```

Serial method:



```
FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -a
runall fmcontrol
```

- e. (Optional) Make sure the Master Chassis Fabric Manager is running.
Ensure the Master Chassis Fabric Manager is running and the slaves are stopped, using:

Parallel method:

```
opachassisadmin -S -F chassisfile -a run fmcontrol
```

Serial method:

```
FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -a run
fmcontrol
```

- f. (Optional) Change the Fabric Manager boot state to enable the Master Chassis Fabric Manager to start at boot using the command:

```
opachassisadmin -S -F chassisfile -I enable
```

- g. (Optional) Change the Fabric Manager boot state to enable the all of the Chassis Fabric Managers to start at boot using the command:

```
opachassisadmin -S -F chassisfile -I enableall
```

3. Generate all chassis problem report information using the command:

```
opacaptureall -p -D 4 -f hostfile
```

4. Run a command on all chassis using:

```
opacmdall -C -S -p -T timelimit -F chassisfile STDIN
```

5. View opachassisadmin result files using:

```
editor result_dir/result_file
```

In the example, *editor* indicates the command line editor, for example, vi.

The following default files are created:

- punchlist.csv
- test.res
- test.log



6.0 Configure Firmware on the Externally Managed Intel® Omni-Path Switches

This section describes using Intel® Omni-Path Fabric Suite FastFabric to install and configure Intel® Omni-Path Switch 100 Series externally managed switches.

6.1 Switch Configuration Prerequisites

Prior to using Intel® Omni-Path Fabric Suite FastFabric, the following minimal steps need to be performed:

1. **(Switch)** Select a unique name to be used for each switch. This name is configured as the Switch Description for the switch in the following steps.
Note: Externally managed switches do not have an Ethernet port and therefore do not have a TCP/IP name.
2. **(Switch)** Copy the relevant switch firmware files onto the Intel® Omni-Path Fabric Suite FastFabric management node. For the following steps, the *.emfw files are used to upgrade the firmware on each switch.

Note: Place all files at a given firmware level into a single directory. Intel recommends that you include the firmware revision number in the directory name.

6.2 Configure Externally Managed Switches Using FastFabric TUI

Refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide* for more information on how to use the FastFabric TUI.

1. **(Switch)** If the **Intel FastFabric OPA Tools** menu is not displayed, type `opafastfabric` at a command prompt and press **Enter**.



2. **(Switch)** Press **2** to display the **FastFabric OPA Switch Setup/Admin Menu**.

```
FastFabric OPA Switch Setup/Admin Menu
Externally Managed Switch File: /etc/sysconfig/opa/switches

Setup:
0) Edit Config and Select/Edit Switch File      [ Skip ]
1) Generate or Update Switch File               [ Skip ]
2) Test for Switch Presence                     [ Skip ]
3) Verify Switch Firmware                      [ Skip ]
4) Update Switch Firmware                      [ Skip ]
5) Setup Switch Basic Configuration             [ Skip ]
6) Reboot Switch                              [ Skip ]
7) Report Switch Firmware & Hardware Info      [ Skip ]
8) Get Basic Switch configuration               [ Skip ]
Admin:
9) Report Switch VPD Information                [ Skip ]
Review:
a) View opaswitchadmin result files            [ Skip ]

P) Perform the selected actions                 N) Select None
X) Return to Previous Menu (or ESC)
```

3. **(Switch)** Select the items in the **Setup** section that are required or needed for your fabric.
4. Press **P**.
5. Perform the items selected using the following sections.

Edit Configuration and Select/Edit Switch File

(Switch) The **Edit Config and Select/Edit Switch File** selection starts the edit mode, where you can edit the switches, ports, and FastFabric configuration files. While in the editor, review all the settings in `opafastfabric.conf`. Refer to *Intel® Omni-Path Fabric Suite FastFabric User Guide* Configuration Files section for more information about `opafastfabric.conf`.

When in the editor for ports, review the file. For typical single-subnet clusters, the default of "0:0" may be used. This uses the first active port on the Management Node to access all externally managed switches. For more information on configuring a cluster with multiple subnets, see [Multi-Subnet Fabrics](#) on page 90. For further details about the file format, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

When in the editor for switches, create the file with a list of the switch node GUID and required switch names. Enter one switch node GUID and required switch name per line. Do not use any spaces before or after the comma separating the switch node GUID and the name, as shown in this example:

```
0x00117500d9000138,edge1
0x00117500d9000139,edge2
```

Note: Do not list internally managed chassis in this file. Those are covered in [Configure Intel® Omni-Path Chassis](#) on page 31.



Note: The **Generate or Update Switch File** menu item or `opagenswitches` may be used to generate a list of the externally managed switches presently in the fabric. For example, when using the vi editor, the command `:r ! opagenswitches` may be used to add the output from this command to the file.

Note: Intel® Omni-Path Fabric Suite FastFabric is topology-aware when updating externally managed switch firmware or resetting the switches. The update or restart starts at the switches farthest from the FastFabric node and then works toward the FastFabric node. This way, switches that are rebooted are not in the path between the FastFabric node and others that are being updated or reset.

For further details regarding the file format, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

If needed, an SA query can be used to get a list of all switches. This includes both internally and externally managed switches. Consequently, the output must be edited to leave only the Intel externally managed switches. An example SA query is:

```
opasaquery -t sw -o nodeguid
```

Generate or Update Switch File

(Switch) The **Generate or Update Switch File** selection generates or updates the switches file. It can also update switch names in the switches file by comparing the actual fabric to topology xml data.

Test for Switch Presence

(Switch) The **Test for Switch Presence** selection verifies that each Externally Managed Switch specified in the switches file can be accessed by the Management Node through the Fabric Network. If all switches were found, continue to the next step. If some switches were not found, exit the menu and review the following for switches that were not found:

- Is switch powered on and booted?
- Is switch connected to Intel® Omni-Path Fabric?
- Is Subnet Manager running?
- Is Management Node's Port active?
- Is Management Node connected to the correct Intel® Omni-Path Fabric?
- Is FM Switch LED activated on the switch port to which the Fabric Management node is connected? For more information, refer to the "FM Switch" section in the *Intel® Omni-Path Fabric Switches Hardware Installation Guide*.
- Is the correct set of switches listed in the switches file? You may need to repeat the previous step to review and edit the file.

Verify Switch Firmware

(Switch) The **Verify Switch Firmware** selection checks that each externally managed switch is operational and that its firmware is valid and accessible.



Update Switch Firmware

(Switch) The **Update Switch Firmware** selection updates the switch firmware version and sets the switch node name. When this selection is performed, the following message is displayed:

```
Multiple Firmware files and/or Directories may be space separated
Shell wildcards may be used
For Directories all .emfw files in the directory tree will be used
Enter Files/Directories to use (or none):
```

Note: Refer to the *Intel® Omni-Path Fabric Externally-Managed Switches Release Notes* to ensure that any prerequisites for the upgrade to the new firmware level have been met prior to performing the upgrade through FastFabric.

Perform the following procedure:

1. Specify the directory where the relevant firmware files are located (copied in a previous step).

The following message displays:

```
After upgrade, the switch may be optionally rebooted.
Would you like to reboot the switch after the update? [n]:
```

2. Type **y**.

The following message displays:

```
The firmware on the switch will be checked, and if the running version is the
same
as the version being used for the update, the update operation will be
skipped.
Would you like to override this check, and force the update to occur? [n]:
```

3. Press **Enter** to select default (n).

The fabric is not yet operational.

The following message displays:

```
You have selected to update the switch firmware and reboot.
There will be a disruption as switch or switches are rebooted.
Doing the operation in parallel (on multiple switches) will finish the
fastest.
Doing it serially may reduce disruption.
Would you like to do the operation in parallel? [y]:
```

Note: Because the Intel® Omni-Path Fabric itself is used to update externally managed switches, updating multiple switches with the reboot option may disrupt parallel update operations. If there are not any selected externally managed switches in the path from the Management Node to any other externally managed switch, parallel operations can be established. For example, if the Management Node is connected directly to a core switch and externally managed switches are only at the edges. To control the order of the rebooting of externally managed switches by FastFabric, see the discussion of the distance option for the switches file in the *Intel® Omni-Path Fabric Suite FastFabric User Guide*, “Externally Managed Switch List File” sub-section.



4. Press **Enter**.

If you do not want to perform the update in parallel, type **n** and press **Enter**.

Note: Be aware that non-parallel operation for a fabric with many externally managed switches can take a significant amount of time.

FastFabric updates the firmware on all switches and sets the node names, as per the switches file created in a previous step. Each switch is then rebooted.

If any switch fails to be updated, use the **View opaswitchadmin result files** option to review the result files from the update. Refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for more details.

Set Up Switch Basic Configuration

(Switch) The **Setup Switch Basic Configuration** selection performs typical switch setup operations using a wizard to configure all switches.

The following steps walk you through the procedure:

1. When this procedure is started, the following message displays:

```
Do you wish to configure the switch Link Width Options? [n]:
```

2. Press **Enter** to select default (n).

This causes the default switch Link Width Options to be used for all switches. If switches have previously been manually configured for different switch Link Width Options, this option keeps the previously configured switch Link Width Options. See the *Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide* for more information.

Selecting (y) prompts for setting the switch link width supported setting for all ports on all switches.

Note: This operation is only applicable to Intel® Omni-Path Edge Switch 100 Series switches.

The following message displays:

```
Do you wish to configure the switch Node Description as it is set in the switches file? [n]:
```

3. Press **Enter** to select default (n).

This causes the default switch Node Description on each switch to be used. If the switches have previously been manually configured for a customized switch Node Description, this option keeps the previously configured switch Node Descriptions. See the *Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide* for more information.

Selecting (y) causes the Node Description on each switch to be updated as specified by the switches file.

Note: Only node descriptions on Intel® Omni-Path Edge Switch 100 Series switches can be changed in this step.

The following message displays:

```
Do you wish to configure the switch FM Enabled option? [n]:
```




4. Press **Enter** to select default (n).

This causes all of the externally managed switch ports to stay FM disabled. Selecting (y) prompts for setting the switch FM enabled capability for all ports on all switches. Setting it to enabled allows the FM to be connected to any port on any externally managed switch. If this is not desired, then select the default for the answer (disabled) and set the desired ports on the externally managed switch to be FM enabled using the FM switch. Refer to the "FM Switch" section in the *Intel® Omni-Path Fabric Switches Hardware Installation Guide* to set the port to FM enabled.

Note: This operation is only applicable to Intel® Omni-Path Edge Switch 100 Series switches.

The following message displays:

```
Do you wish to configure the switch Link CRC Mode? [n]:
```

5. Press **Enter** to select default (n).

This causes all of the externally managed switch ports Link CRC Mode to stay disabled. Selecting (y) prompts for setting the link CRC Mode for all ports on all switches. Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide* for more information.

Reboot Switch

(Switch) The **Reboot Switch** selection reboots all switches, ensuring that all the configuration changes become effective and are discovered by the Intel® Omni-Path Fabric Suite Fabric Manager.

Report Switch Firmware and Hardware Information

(Switch) The **Report Switch Firmware & Hardware Info** selection reports the firmware and hardware versions for each switch, along with other information for all of the externally managed switches. Review the results against the expected models and firmware versions.

(Switch) If any Intel® Omni-Path Switch 100 Series switches were purposely skipped, these sections should be repeated for those switches. In this case, Intel recommends that you create a separate file with a name other than switches.

The following is an example of the results:

```
Performing Switch Admin: Report Switch firmware & hardware info
Executing: /usr/sbin/opaswitchadmin -L /etc/sysconfig/opa/switches info
Executing report switch info Test Suite (switchinfo) day mmm dd hh:mm:ss timezone
YYYY ...
Executing TEST SUITE report switch info CASE (switchinfo.
0x00117500ff513121,Node_Name.i2c.extmtd.switchinfo)
  retrieve switch 0x00117500ff513121,Node_Name ...
TEST SUITE report switch info CASE (switchinfo.
0x00117500ff513121,Node_Name.i2c.extmtd.switchinfo)
  retrieve switch 0x00117500ff513121,Node_Name
0x00117500ff513121,hds1swb8171:
  F/W ver:xx.xx.xx.xx H/W ver:XXX H/W pt num:NNNNNN-NNN
  Fan status:Normal/Normal/Normal/Normal/Normal/Normal PS1 Status:N/A PS2
Status:ONLINE
PASSED
```



```
TEST SUITE report switch info: 1 Cases; 1 PASSED
TEST SUITE report switch info PASSED
Done report switch info Test Suite day mmm dd hh:mm:ss timezone yyyy
```

Get Basic Switch Configuration

(Switch) The **Get Basic Switch Configuration** selection returns the switch configuration report for all of the ports. The results show the number of cases, how many of the cases passed, and how many of the cases failed. It also gives an overall summary of configuration and passed or failed, as shown in the following example:

```
Performing Switch Admin: Get basic Switch configuration
Executing: /usr/sbin/opaswitchadmin -L /etc/sysconfig/opa/switches getconfig
Executing report switch getconfig Test Suite (switchgetportconfig) day mmm dd
hh:mm:ss timezone yyyy ...
Executing TEST SUITE report switch getconfig CASE (switchgetportconfig.
0x00117500ff513121,hds1swb8171.i2c
.extmgd.switchgetportconfig) retrieve switch 0x00117500ff513121,Node_Name ...
TEST SUITE report switch getconfig CASE (switchgetportconfig.
0x00117500ff513121,Node_Name.i2c
.extmgd.switchgetportconfig) retrieve switch 0x00117500ff513121,Node_Name
Link Width : 1,2,3,4
Link Speed : 25Gb
FM Enabled : Yes
Link CRC Mode : 14-bit,16-bit,48-bit
vCU : 0
External Loopback Allowed : Yes
Node Description : Node_Name

PASSED
TEST SUITE report switch getconfig: 1 Cases; 1 PASSED
TEST SUITE report switch getconfig PASSED
Done report switch getconfig Test Suite day mmm dd hh:mm:ss timezone yyyy
```

Report Switch VPD Information

(Switch) The **Report Switch VPD Information** selection returns the vital product data (VPD) for all of the nodes listed in /etc/sysconfig/opa/switches. The results display the VPD hardware information as shown in the following example:

```
Performing Switch Admin: Report Switch VPD information
Executing: /usr/sbin/opaswitchadmin -L /etc/sysconfig/opa/switches hwvpd
Executing report switch hwvpd Test Suite (switchhwvpd) day mmm dd hh:mm:ss
timezone yyyy ...
Executing TEST SUITE report switch hwvpd CASE (switchhwvpd.
0x00117500ff513121,Node_Name.i2c
.extmgd.switchhwvpd) retrieve switch 0x00117500ff513121,Node_Name ...
TEST SUITE report switch hwvpd CASE (switchhwvpd.0x00117500ff513121,Node_Name.i2c
.extmgd.switchhwvpd) retrieve switch 0x00117500ff513121,Node_Name

0x00117500ff513121,hds1swb8171: H/W VPD serial number: USFU13150000D
0x00117500ff513121,hds1swb8171: H/W VPD part number : NNNNNN-NNN
0x00117500ff513121,hds1swb8171: H/W VPD model : 100SWE48QF2
0x00117500ff513121,hds1swb8171: H/W VPD h/w version : 004
0x00117500ff513121,hds1swb8171: H/W VPD manufacturer : Intel Corporation
0x00117500ff513121,hds1swb8171: H/W VPD prod desc : 100 OP Edge 48p Q7
forward 2PSU
0x00117500ff513121,hds1swb8171: H/W VPD mfg id : 001175
0x00117500ff513121,hds1swb8171: H/W VPD mfg date : m-dd-yyyy
0x00117500ff513121,hds1swb8171: H/W VPD mfg time : hh:mm

PASSED
TEST SUITE report switch hwvpd: 1 Cases; 1 PASSED
TEST SUITE report switch hwvpd PASSED
Done report switch hwvpd Test Suite day mmm dd hh:mm:ss timezone yyyy
```



View opaswitchadmin Result Files

(Switch) The **View opaswitchadmin Result Files** selection starts the editor to view opaswitchadmin result files. The files that are opened in the editor are:

- /root/punchlist.csv
- /root/test.res
- /root/test.log

When you exit the vi editor, the following question is displayed:

```
Would you like to remove test.res test.log test_tmp* and save_tmp
in /root ? [n]:
```

After you answer the question, it takes you back to the TUI.

6.3 Configure Externally Managed Switches Using CLI Commands

Prerequisites: Refer to [Switch Configuration Prerequisites](#) for details.

This procedure provides step-by-step information to configure the chassis.

1. (Optional) Edit or review the `/etc/sysconfig/opa/opafastfabric.conf` file.

Review the complete file. Refer to "Configuration Files for FastFabric" section of the *Intel® Omni-Path Fabric Suite FastFabric User Guide* for more information.

2. (Optional) Edit or review the `/etc/sysconfig/opa/ports` file.

For typical single-subnet clusters, the default of "0:0" may be used. This uses the first active port on the Fabric Management Node to access all externally managed switches. For more information on configuring a cluster with multiple subnets, see [Multi-Subnet Fabrics](#) on page 90. For further details about the file format, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

3. (Optional) Create, edit, or review the `/etc/sysconfig/opa/switches` file.

When in the editor for `switches`, create the file with a list of the switch node GUIDs with the required switch names. Enter one switch node GUID and required switch name per line. Optionally you can enter the relative distance of each switch from the Management Node at the end of the line. Do not use any spaces before or after the comma separating the switch node GUID, the name, and the relative distance (optional), as shown in this example:

```
0x00117500d9000138,edge1,1
0x00117500d9000139,edge2,2
```

Note: Do not list internally managed chassis in this file. Those are covered in [Configure Intel® Omni-Path Chassis](#) on page 31.

Note: Refer to the following step to generate a list of the externally managed switches presently in the fabric.



Note: Intel® Omni-Path Fabric Suite FastFabric is topology-aware when updating externally managed switch firmware or resetting the switches. The update or restart starts at the switches farthest from the FastFabric node and then works toward the FastFabric node. This way, switches that are rebooted are not in the path between the FastFabric node and others that are being updated or reset.

For further details regarding the `switches` file format, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*, "Configuration Files for FastFabric" section.

If needed, an SA query can be used to get a list of all switches, including both internally and externally managed switches. Consequently, the output must be edited to leave only the Intel externally managed switches. An example SA query is:

```
opasaquery -t sw -o nodeguid
```

4. (Optional) Generate or update a list of the externally managed switches presently in the fabric and save the list to the `switches` file. You can also update switch names in the `switches` file by comparing the actual fabric to topology xml data.

```
opagenswitches -s -o switches
```

5. Test to see if each externally managed switch is present.

```
opaswitchadmin -L switchfile ping
```

If some switches were not found, review the following list for the switches that were not found:

- Is switch powered on and booted?
- Is switch connected to Intel® Omni-Path Fabric?
- Is Subnet Manager running?
- Is Fabric Management node's Port active?
- Is Fabric Management node connected to the correct Intel® Omni-Path Fabric?
- Is FM Switch LED activated on the switch port to which the Fabric management node is connected? Refer to the "FM Switch" section in the *Intel® Omni-Path Fabric Switches Hardware Installation Guide*.
- Is correct set of switches listed in the switches file? You may need to repeat the previous step to review and edit the file.

6. Verify the switch firmware to check that each externally managed switch is operational and that its firmware is valid and accessible.

```
opaswitchadmin -L switchfile fwverify
```

7. Update the switch firmware version and the switch node name set.

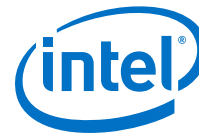
Run the firmware update in parallel using:

```
opaswitchadmin -O -L switchfile -P packages -a run upgrade
```

Run the firmware update in serial using:

```
FF_MAX_PARALLEL=0 opaswitchadmin -O -L switchfile -P packages  
-a run upgrade
```

Note: Be aware that non-parallel operation for a fabric with many externally managed switches can take a significant amount of time.



8. Reboot the externally managed switches using the command:

```
opaswitchadmin -L switchfile reboot
```
9. Set up the switch basic configuration for all externally managed switches using the command:

```
opaswitchadmin -L switchfile configure
```

This command runs a wizard to set up the externally managed switch configuration.
10. Run a report for the firmware and hardware versions for each switch, along with the part number and data rate capability of all of the externally managed switches.

```
opaswitchadmin -L switchfile info
```
11. Run the command to get the externally managed switch configuration report for all of the ports.

```
opaswitchadmin -L switchfile getconfig
```
12. Run the report that returns externally managed switch hardware vital product data (VPD) for all of the nodes listed in the `/etc/sysconfig/opa/switches` file.

```
opaswitchadmin -L switchfile hwvpd
```



7.0 Install Host Software on the Remaining Hosts Using the FastFabric TUI

This section describes using the Intel® Omni-Path Fabric Suite FastFabric TUI to install and configure the remaining hosts with the Host Software.

Note: Intel® Omni-Path Fabric Suite FastFabric is also used to install the Intel® Omni-Path Fabric Stack Tools on the remaining hosts when using other variations of OFA. In this case, OFA software must be installed on each host prior to installing the OPA software.

1. **(All)** If the **Intel FastFabric Tools** menu is not displayed, type `opafastfabric` at a command prompt and press **Enter**.
2. **(Linux)** Press **3** to select Host Setup.

The **FastFabric OPA Host Setup Menu** displays:

```
FastFabric OPA Host Setup Menu
Host File: /etc/sysconfig/opa/hosts

Setup:
0) Edit Config and Select/Edit Host File      [ Skip ]
1) Verify Hosts Pingable                     [ Skip ]
2) Set Up Password-Less SSH/SCP               [ Skip ]
3) Copy /etc/hosts to All Hosts               [ Skip ]
4) Show uname -a for All Hosts                [ Skip ]
5) Install/Upgrade OPA Software               [ Skip ]
6) Configure IPoIB IP Address                 [ Skip ]
7) Build Test Apps and Copy to Hosts          [ Skip ]
8) Reboot Hosts                              [ Skip ]
Admin:
9) Refresh SSH Known Hosts                   [ Skip ]
a) Rebuild MPI Library and Tools              [ Skip ]
b) Run a Command on All Hosts                 [ Skip ]
c) Copy a File to All Hosts                   [ Skip ]
Review:
d) View opahostadmin Result Files             [ Skip ]

P) Perform the Selected Actions               N) Select None
X) Return to Previous Menu (or ESC)
```

3. Select items **0** through **2** and **4** through **8**.

Note: If `/etc/hosts` is used for name resolution (as opposed to using DNS), also select 3) Copy `/etc/hosts` to all hosts

4. Press **P**.
5. Perform the items selected using the descriptions in the following sections.

Edit Config and Select/Edit Host File

(All) The **Edit Config and Select/Edit Host File** selection permits the hosts and FastFabric configuration files to be edited. When placed in the editor for `opafastfabric.conf`, review all the settings. In particular:

- `FF_IPOIB_SUFFIX`



- FF_IPOIB_NETMASK
- FF_IPOIB_CONFIG
- FF_PRODUCT
- FF_PACKAGES
- FF_INSTALL_OPTIONS
- FF_UPGRADE_OPTIONS

Refer to "Configuration Files for FastFabric" section of the *Intel® Omni-Path Fabric Suite FastFabric User Guide* for more information.

Note: During setup of password-less SSH, FastFabric provides the opportunity to enter the host root password interactively when needed. Therefore, Intel recommends that you do not place it within the `opafastfabric.conf` file. If you are required to keep the root password for the hosts in the `opafastfabric.conf` file, Intel recommends that you change the `opafastfabric.conf` permissions to be 0x600 (root-only access).

When placed in the editor for hosts, create the file with a list of the hosts names (the TCP/IP management network names), except the Management Node from which FastFabric is presently being run. Enter one host's name per line. For example:

```
host1
host2
```

Note: Do not list the Management Node itself (the node where FastFabric is currently running).

If additional Management Nodes are to be used, they may be listed at this time, and FastFabric can aid in their initial installation and verification.

For further details about the file format, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

Verify Hosts Pingable

(All) The **Verify Hosts Pingable** selection pings each selected host over the management network.

The system prompts:

```
Performing Host Setup: Verify Hosts Pingable
Would you like to verify hosts are ssh-able? [n]:
```

Press **ENTER** to select the default (n).

If all hosts were found, continue to the next step. If some hosts were not found, exit the menu and review the following for those hosts that were not found:

- Host powered on and booted?
- Host connected to management network?
- Host management network IP address and network settings consistent with DNS or `/etc/hosts`?
- Management node connected to the management network?



- Management node IP address and network settings correct?
- Management network itself up (including switches, routers, and others)?
- Correct set of hosts listed in the hosts file? You may need to repeat the previous step to review and edit the file.

Set Up Password-Less SSH/SCP

(Linux) The **Setup Password-less ssh/scp** selection sets up secure password-less SSH (root password) such that the Management Node can securely log in to all the other hosts as root through the management network without requiring a password.

The system prompts:

```
Performing Host Setup: Setup Password-less ssh/scp
Executing: /usr/sbin/opasetupssh -S -p -i '' -f /etc/sysconfig/opa/hosts
Password for root on all hosts:
```

Type the password for root on all hosts and press **ENTER**.

Password-less SSH is required by Intel® Omni-Path Fabric Suite FastFabric, MPI test applications, and most versions of MPI (including OpenMPI and MVAPICH2).

Copy /etc/hosts to All Hosts

(Linux) The **Copy /etc/hosts to all hosts** selection copies the `/etc/hosts` file on this host to all the other selected hosts.

Note: If DNS is being used, skip this step.

Note: Typically, `/etc/resolv.conf` is set up as part of OS installation for each host. However, if `/etc/resolv.conf` was not set up on all the hosts during OS installation, the **Copy a File to All Hosts** operation could be used at this time to copy `/etc/resolv.conf` from the Management Node to all the other nodes.

Show uname -a for All Hosts

(Linux) The **Show uname -a for All Hosts** selection shows the OS version on all the hosts. Review the results carefully to verify all the hosts have the expected OS version. In typical clusters, all hosts are running the same OS and kernel version.

If any hosts are identified with an incorrect OS version, the OS on those hosts should be corrected at this time and this sequence should be cancelled when prompted. After the OS versions have been corrected, repeat the preceding setup steps. There is no harm in repeating the steps for all the hosts.

Install/Upgrade OPA Software

(Host) The **Install/Upgrade OPA Software** selection installs the Intel® Omni-Path Fabric Host Software on all the hosts. By default, it looks in the current directory for the `IntelOPA-[Basic|IFS].DISTRO.VERSION.tgz` file. If the tarball is not found in the current directory, the installer application prompts for a directory name where this file can be found.



Note: An initial installation uninstalls any existing OFA Delta or IFS software. Initial installs must be performed when installing on a clean system or on a system that has stock OFA installed. To upgrade the fabric, refer to [Upgrade the Fabric Software](#) on page 105.

Perform the following steps to install the selected hosts:

1. The **Install/Upgrade Intel Software** selection displays the following prompt:

```
Performing Host Setup: Install/Upgrade Intel OPA Software
Do you want to use ./IntelOPA-[Basic|IFS].DISTR0.VERSION.tgz? [y]:
```

2. Press **ENTER** to accept the default (y).

The system prompts:

```
Would you like to do a fresh [i]ninstall, an [u]pgrade or [s]kip this step? [u]:
```

3. Type **i** and press **ENTER** to perform an initial installation.

The system prompts:

```
You have selected to perform an initial installation
This will uninstall any existing OPA software on the selected nodes
Are you sure you want to proceed? [n]:
```

4. Type **y** and press **ENTER** to proceed.

The system prompts:

```
/usr/sbin/opahostadmin -f /etc/sysconfig/opa/hosts -d . load
Executing load Test Suite (load) Day Mth DD HH:MM:SS timezone yyyy ...
.
.
.
Hit any key to continue (or ESC to abort)...
```

5. Press any key to proceed.

Once the installation is complete on the selected hosts, the **FastFabric OPA Host Setup Menu** displays.

The installation is complete.

Note: If any hosts fail to be installed, as shown in the following example:

```
TEST SUITE load: 1 Cases; 0 PASSED; 1 FAILED
TEST SUITE load FAILED
```

Use the [View opahostadmin result files](#) option to review the result files from the update. For more details, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.



Configure IPoIB IP Address

(Host) The **Configure IPoIB IP Address** selection creates the `ifcfg-ib0` files on each host. (Previous non-OFA releases created the `ifcfg-ib1` file.) The file is created with a statically assigned IPv4 address. The IPoIB IP address for each host is determined by the resolver (Linux* `host` command). If not found through the resolver, `/etc/hosts` on the given host is checked.

Build Test Apps and Copy to Hosts

(Host) The **Build Test Apps and Copy to Hosts** selection builds the MPI and/or SHMEM sample applications on the Management Node and copies the resulting object files to all the hosts. This is in preparation for execution of MPI and/or SHMEM performance tests and benchmarks in a later step.

Reboot Hosts

(Linux) The **Reboot Hosts** selection reboots all the selected hosts and ensures they fully reboot, as verified through ping over the management network. When the hosts come back up, they are running the newly installed Fabric software.

Refresh SSH Known Hosts

(Linux) The **Refresh SSH Known Hosts** runs the `opasetupssh -p -U -f /etc/sysconfig/opa/hosts` command to refresh the SSH known hosts list on this server for the Management Network. This may be used to update security for this host if hosts are replaced, reinstalled, renamed, or repaired.

Rebuild MPI Library and Tools

(Host) The **Rebuild MPI Library and Tools** rebuilds the MPI Library and related tools (such as `mpirun`), and installs the resulting rpms on all the hosts. This is performed using the `do_build` tool supplied with the MPI Source. When rebuilding MPI, `do_build` prompts you for selection of which MPI to rebuild, and provides choices as to which available compiler to use. Refer to *Intel® Omni-Path Host Fabric Interface Installation Guide* and *Intel® Omni-Path Fabric Host Software User Guide* for more information.

Run a Command on All Hosts

(Linux) If there are any other setup operations that need to be performed on all hosts, use the **Run a Command on All Hosts** option. Each time this is executed, a Linux* shell command may be specified to be executed against all selected hosts. You can also specify a sequence of commands separated by semicolons.

Note:

Intel recommends that you run the `date` command on all hosts to verify that the date and time are consistent. If needed, use the **Copy a File to All Hosts** option to copy the appropriate files to all hosts to enable and configure NTP.

Copy a File to All Hosts

(Linux) The **Copy a File to All Hosts** selection runs the `opascpall` command. A file on the local host may be specified to be copied to all selected hosts.



View opahostadmin Result Files

(All) The **View opahostadmin Result File** displays the `test.log` and `test.res` files that contain the results from prior `opahostadmin` runs, such as installing Fabric software or rebooting all hosts. You are also given the option to remove these files after viewing them.

If prior files are not removed, subsequent runs of `opachassisadmin`, `opahostadmin`, or `opaswitchadmin` from within the current directory continue to append to these files.



8.0 Install Host Software on the Remaining Hosts Using CLI Commands

This section provides step-by-step information to install the Intel® Omni-Path Fabric Host Software on the remaining hosts using CLI commands.

1. (Optional) Edit or review the `/etc/sysconfig/opa/opafastfabric.conf` file.

Review the complete file. In particular, review the following:

- `FF_IPOIB_SUFFIX`
- `FF_IPOIB_NETMASK`
- `FF_IPOIB_CONFIG`
- `FF_PRODUCT`
- `FF_PACKAGES`
- `FF_INSTALL_OPTIONS`
- `FF_UPGRADE_OPTIONS`

Refer to the "Configuration Files for FastFabric" section of the *Intel® Omni-Path Fabric Suite FastFabric User Guide* for more information.

Note: During setup of password-less SSH, FastFabric provides the opportunity to enter the host root password interactively when needed. Therefore, Intel recommends that you do not place it within the `opafastfabric.conf` file. If you are required to keep the root password for the hosts in the `opafastfabric.conf` file, Intel recommends that you change the `opafastfabric.conf` permissions to be `0x600` (root-only access).

2. (Optional) Create, edit, or review the `/etc/sysconfig/opa/hosts` file.

When placed in the editor for hosts, create the file with a list of the hosts names (the TCP/IP management network names), except the Management Node from which FastFabric is presently being run. Enter one host's name per line. For example:

```
host1
host2
```

Note: Do not list the Management Node itself (the node where FastFabric is currently running).

If additional Management Nodes are to be used, they may be listed at this time, and FastFabric can aid in their initial installation and verification.

For further details about the file format, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

3. Verify the hosts are pingable over the management network.

```
opapingall -p
```



If all hosts were found, continue to the next step. If some hosts were not found, exit the menu and review the following for those hosts that were not found:

- Host powered on and booted?
 - Host connected to management network?
 - Host management network IP address and network settings consistent with DNS or `/etc/hosts`?
 - Management node connected to the management network?
 - Management node IP address and network settings correct?
 - Management network itself up (including switches, routers, and others)?
 - Correct set of hosts listed in the hosts file? You may need to repeat the previous step to review and edit the file.
4. Set up secure password-less SSH, such that the Management Node can securely log into all the hosts as `root` through the management network, without requiring a password.

```
opasetupssh -S -p -i "" -f hostfile
```

5. (Optional) Copy the `/etc/hosts` file on this host to all the other selected hosts.

```
opascpall -p -f hostfile /etc/hosts /etc/hosts
```

Note: If DNS is being used, skip this step.

6. (Optional) Copy the `/etc/resolv.conf` file on this host to all the other selected hosts.

```
opascpall -p -f hostfile /etc/resolv.conf /etc/resolv.conf
```

7. Show the OS version on all the hosts.

```
opacmdall -T 60 -f hostfile 'uname -a'
```

8. Install the Intel® Omni-Path Fabric Host Software on all the hosts.

```
opahostadmin -f hostfile -d dir load
```

By default, it looks in the current directory for the `IntelOPA-Basic.DISTRO.VERSION.tgz` file.

Note: An initial installation uninstalls any existing OFA Delta or IFS software. Initial installs must be performed when installing on a clean system or on a system that has stock OFA installed. For additional details, refer to [Upgrade the Fabric Software](#) on page 105.

Note: If any hosts fail to be installed, as shown in the following example:

```
TEST SUITE load: 1 Cases; 0 PASSED; 1 FAILED
TEST SUITE load FAILED
```

Use the [View opahostadmin Result Files](#) option to review the result files from the update. For more details, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

9. Create the `ifcfg-ib0` files on each host.

```
opahostadmin -f hostfile configipoib
```



The file is created with a statically assigned IPv4 address. The IPoIB IP address for each host is determined by the resolver (Linux* host command). If not found through the resolver, `/etc/hosts` on the given host is checked.

10. Build the MPI and/or SHMEM sample applications on the Management Node and copy the resulting object files to all the hosts.

```
MPICH_PREFIX=path_to_mpi cd /usr/lib/opa/src/mpi_apps; make  
clobber quick opascall -t -p -f hostfile source_dir dest_dir
```

Note: This is in preparation for execution of MPI and/or SHMEM performance tests and benchmarks in a later step.

Note: This option is only available when using the Intel® Omni-Path Fabric Host Software packaging of OFA (OFA Delta).

11. Reboot all the selected hosts and ensure they fully reboot, as verified through ping over the management network.

```
opahostadmin -f hostfile reboot
```

When the hosts come back up, they are running the newly installed Fabric software.



9.0 Verify Intel® Omni-Path Fabric Host Software on the Remaining Servers Using the FastFabric TUI

This section provides step-by-step information to verify that Intel® Omni-Path Fabric Host Software is installed and running on the remaining servers using the FastFabric TUI. As a result of running this sequence, a `punchlist.csv` file is produced. This file provides a cumulative summary of tests that failed and may be provided to technicians for corrective action. The file can easily be imported into spreadsheets or other tools.

1. **(All)** If the **FastFabric OPA Host Verification/Admin Menu** is not displayed, type `opafastfabric` and press **Enter**.
2. **(All)** Press **4**.

Displays the **FastFabric OPA Host Verification/Admin Menu**.

```
FastFabric OPA Host Verification/Admin Menu
Host File: /etc/sysconfig/opa/allhosts

Validation:
0) Edit Config and Select/Edit Host File      [ Skip ]
1) Summary of Fabric Components               [ Skip ]
2) Verify Hosts Pingable, SSHable and Active [ Skip ]
3) Perform Single Host Verification           [ Skip ]
4) Verify OPA Fabric Status and Topology      [ Skip ]
5) Verify Hosts See Each Other                [ Skip ]
6) Verify Hosts Ping via IPoIB                [ Skip ]
7) Refresh SSH Known Hosts                   [ Skip ]
8) Check MPI Performance                     [ Skip ]
9) Check Overall Fabric Health                [ Skip ]
a) Start or Stop Bit Error Rate Cable Test   [ Skip ]
Admin:
b) Generate All Hosts Problem Report Info    [ Skip ]
c) Run a Command on All Hosts                [ Skip ]
Review:
d) View opahostadmin Result Files            [ Skip ]

P) Perform the Selected Actions               N) Select None
X) Return to Previous Menu (or ESC)
```

3. Select the items **0** through **8** in the **Validation** section of the menu.
4. Press **P**.

Edit Config and Select/Edit Host File

(All) The **Edit Config and Select/Edit Host File** section permits the hosts, ports, and FastFabric configuration files to be edited. When placed in the editor for `opafastfabric.conf`, review all the settings. Especially review the settings for the following:

- `FF_TOPOLOGY_FILE`
- `FF_IPOIB_SUFFIX`
- `FF_DEVIATION_ARGS`



- `ff_host_basename_to_ipoib`
- `ff_host_basename`

Refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide* Configuration Files section for more information about `opafastfabric.conf`.

Intel recommends that a FastFabric topology file is created as `/etc/sysconfig/opa/topology.0:0.xml` to describe the intended topology of the fabric. The file can also augment assorted fabric reports with customer-specific information, such as cable labels and additional details about nodes, SMs, links, ports, and cables. Refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for more information about topology verification files.

Review the following parameters which are used for overall fabric health checks:

- `FF_ANALYSIS_DIR`
- `FF_ALL_ANALYSIS`
- `FF_FABRIC_HEALTH`
- `FF_CHASSIS_CMDS`
- `FF_CHASSIS_HEALTH`
- `FF_ESM_CMDS`

`FF_ALL_ANALYSIS` should be updated to reflect the type of SM (esm or hostsm).

When placed in the editor for `ports`, review the file. For typical single-subnet clusters, the default of `0:0` may be used. This uses the first active port on the Management Node to access the fabric. For more information on configuring a cluster with multiple subnets, see [Multi-Subnet Fabrics](#) on page 90. For further details about the file format, refer to the "Selection of Ports" section in the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

When placed in the editor for `allhosts`, create the file with the Management Node's hosts name (the TCP/IP management network name, for example `mgmthost`) and include the hosts file previously created. Enter one file per line. For example:

```
mgmthost
include /etc/sysconfig/opa/hosts
```

For further details about the file format, refer to the "Selection of Hosts" section in the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

Summary of Fabric Components

(All) The **Summary of Fabric Components** selection provides a brief summary of the counts of components in the fabric, including how many switch chips, hosts, and links are in the fabric. It also indicates whether any degraded or omitted links were found. This can indicate a poorly seated or bad cable, incorrect fabric configuration, or security issues. Review the results against the expected configuration of the cluster.

If components are missing, or degraded or omitted links are found, they should be corrected. Subsequent steps aid in locating any such links.



Verify Hosts Pingable, SSHable, and Active

(All) The **Verify Hosts Pingable, SSHable, and Active** selection verifies each host and provides a concise summary of the bad hosts found. Interactive prompts allow you to select ping, SSH, and port active verification. After completion of this test, you have the option of using the resulting good hosts file for the remainder of the operations within this TUI session.

Perform Single Host Verification

(All) The **Perform Single Host Verification** uses the `opaverifyhosts` command to perform a single host test on all hosts. Prior to using this selection in the TUI you must have a copy of the `hostverify.sh` in the directory pointed to by `FF_HOSTVERIFY_DIR`. If the file does not exist in that directory copy the sample file `/usr/lib/opa/samples/hostverify.sh` to the directory pointed to by `FF_HOSTVERIFY_DIR`. When placed in the editor to review `hostverify.sh`, review the settings near the top and the list of TESTS selected, edit and save as needed. Refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for information on `opaverifyhosts`.

Verify OPA Fabric Status and Topology

(All) The **Verify OPA Fabric Status and Topology** selection runs the following checks:

- Perform a fabric error analysis.
- Clear error counters after generating the report.
- Perform a fabric link speed error analysis.
- Check for links that are configured to run slower than supported.
- Check links that are connected with mismatched speed potential.
- Verify the fabric topology.
- Verify all aspects of the topology including links, nodes, and SMs.
- Include unexpected devices in the punch list.

Verify Hosts See Each Other

(Host) The **Verify Hosts See Each Other** selection confirms that each host can see all the others through queries to the Subnet Administrator. This ensures all nodes are connected to the same fabric and can properly access the Subnet Administrator.

Verify Hosts Ping via IPoIB

(Host) The **Verify Hosts Ping via IPoIB** selection confirms that IPoIB is properly configured and running on all the hosts. This is accomplished through the Management Node pinging each host through IPoIB.

Note: This operation requires that IPoIB is enabled on the Management Node as well as each host selected for verification. Also, the management host must have IPoIB configured.

Refresh SSH Known Hosts

(Linux) The **Refresh SSH Known Hosts** selection refreshes the SSH `known_hosts` file on the Management Node to include the IPoIB hostnames of all the hosts.



Note: This operation requires that IPoIB is enabled on the Management Node as well as each host selected for verification.

Check MPI Performance

(Host) The **MPI Performance** selection does a quick check of PCIe and MPI performance through end-to-end latency and bandwidth tests.

1. When MPI Performance is selected, the following prompt is displayed:

```
Test Latency and Bandwidth deviation between all hosts? [y]:
```

2. Press **Enter** to select default (y).

The following prompt is displayed:

```
View Load on hosts prior to test? [y]:
```

3. Press **Enter** to select default (y).

This displays the results of pair-wise analysis of latency and bandwidth for the selected hosts and reports pairs outside an acceptable tolerance range. By default, performance is compared relative to other hosts in the fabric. It is assumed that all hosts selected for a given run have comparable fabric performance. Failing hosts are clearly indicated.

Intel recommends that you review the `FF_DEVIATION_ARGS` parameter in `opafastfabric.conf` and adjust it as appropriate for the cluster. The default is quite lenient to accommodate a wide range of cluster designs.

Note: This test identifies nodes whose performance is not consistent with others in the fabric. It is not intended as a benchmark of fabric latency and bandwidth. This test purposely uses techniques to reduce test runtime.

The results are also written to the `test.res` file, which may be viewed through the **View opahostadmin result files** option. Refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for more details.

If any hosts fail, carefully examine the failing hosts to verify the HFI models, PCIe slot used, BIOS settings, and any motherboard or BIOS settings related to devices on PCIe buses or slot speeds. Also verify the HFI and any riser cards are properly seated.

The bandwidth that is reported should also be checked against the PCIe speeds in the Performance Impact table below. If all pairs are not in the expected performance range, carefully examine all hosts to verify the HFI models, PCIe slot used, BIOS settings and any motherboard or BIOS settings related to devices on PCIe buses or slot speeds. Also verify the HFI and any riser cards are properly seated.

Table 1. Performance Impact

PCIe Speed	Fabric Speed	Typical Bandwidth
PCIe 8GT/s x16 (Gen3)	100 Gbit/sec	12.0 - 12.4 GB/sec
PCIe 8GT/s x8 (Gen3)	100 Gbit/sec	6.4 - 6.8 GB/sec
PCIe 5GT/s x16 (Gen2)	100 Gbit/sec	6.4 - 6.8 GB/sec
PCIe 5GT/s x8 (Gen2)	100 Gbit/sec	3.2 - 3.4 GB/sec



Note: 1GB/sec = 1,000,000,000 bytes/second

Check Overall Fabric Health

(All) The **Check Overall Fabric Health** selection permits the present fabric configuration to be baselined for use in future fabric health checks. Perform this check after configuring any additional Management Nodes and establishing a healthy fabric via successful execution of all the other tests discussed in this section. If desired, a baseline of an incomplete or unhealthy fabric may be taken for future comparison after making additions or corrections to the fabric. Refer to [Configure and Initialize Health Check Tools](#) on page 76 for more information.

Start or Stop Bit Error Rate Cable Test

(All) The **Start or Stop Bit Error Rate Cable Test** selection performs host and/or ISL cable testing. The test allows for starting and stopping an extended Bit Error Rate test. The system prompts to clear hardware counters.

Note: Clearing of hardware counters (-A option) is optional and may affect the PM and other tools. See "PM Running Counters to Support opareport" section in the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for more information.

Intel recommends that you run this test for 20-60 minutes for a thorough test. While the test is running, monitor the fabric for signal integrity or stability errors using `opatop`, `opareport`, and/or the Fabric Manager GUI. Once the desired test time has elapsed, return to this item in the menu and stop the test.

Generate all Hosts Problem Report Info

(Host) The **Generate all Hosts Problem Report Info** runs the `captureall` command to collect configuration and status information from all hosts and generates a single *.tgz file that can be sent to an Intel support representative.

Based on the answer to the following prompt, various levels of detail about the fabric are included in the capture.

```
Capture detail level (1-Normal, 2-Fabric, 3-Fabric+FDB, 4-Analysis):
```

The detail levels are:

- 1-Normal - Obtains local information from each host
- 2-Fabric - In addition to "Normal", also obtains basic fabric information by queries to the SM and fabric error analysis using `iba_report`.
- 3-Fabric+FDB - In addition to "Fabric", also obtains all the switch forwarding tables and OPA multicast membership lists from the SM.
- 4-Analysis - In addition to "Fabric+FDB", also obtains `all_analysis` results. If `all_analysis` has not yet been run, it is run as part of the capture.

Note: Detail levels 2-4 can be used when fabric operational problems occur. If the problem appears to be node-specific, detail level 1 should be sufficient. Detail levels 2-4 require an operational Fabric Manager. Typically your support representative requests a given detail level. If a given detail level takes excessively long or fails to be gathered, try a lower detail level.



For detail levels 2-4, the additional information is only gathered on the node running the `captureall` command. The information is gathered for every fabric specified in the `/etc/sysconfig/opa/ports` file.

Run a Command on All Hosts

(Linux) The **Run a command on all hosts** runs the `cmdall` command. A Linux* shell command may be specified to be executed against all selected hosts. You may also specify a sequence of commands separated by semicolons.

View opahostadmin Result Files

(All) The **View opahostadmin result files** permits viewing of the `test.log` and `test.res` files, which reflect the results from `opahostadmin` runs, such as those for installing fabric software or rebooting all hosts. You are also given the option to remove these files after viewing them.

If prior files are not removed, subsequent runs of `opachassisadmin`, `opahostadmin`, or `opaswitchadmin` from within the current directory continue to append to these files.



10.0 Verify Intel® Omni-Path Fabric Host Software on the Remaining Servers Using CLI Commands

This section provides step-by-step information to verify that Intel® Omni-Path Fabric Host Software is installed and running on the remaining servers using CLI commands. As a result of running this sequence, a `punchlist.csv` file is produced. This file provides a cumulative summary of tests that failed and may be provided to technicians for corrective action. The file can easily be imported into spreadsheets or other tools.

1. (Optional) Edit or review the `/etc/sysconfig/opa/opafastfabric.conf` file.

Review all of the settings. In particular, review the following:

- `FF_TOPOLOGY_FILE`
- `FF_IPOIB_SUFFIX`
- `FF_DEVIATION_ARGS`
- `ff_host_basename_to_ipoib`
- `ff_host_basename`

Refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide* Configuration Files section for more information about the `opafastfabric.conf` file.

Intel recommends that a FastFabric topology file is created as `/etc/sysconfig/opa/topology.0:0.xml` to describe the intended topology of the fabric. The file can also augment assorted fabric reports with customer-specific information, such as cable labels and additional details about nodes, SMs, links, ports, and cables. Refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for more information about topology verification files.

Review the following parameters which are used for overall fabric health checks:

- `FF_ANALYSIS_DIR`
- `FF_ALL_ANALYSIS`
- `FF_FABRIC_HEALTH`
- `FF_CHASSIS_CMDS`
- `FF_CHASSIS_HEALTH`
- `FF_ESM_CMDS`

`FF_ALL_ANALYSIS` should be updated to reflect the type of SM (esm or hostsm).

2. (Optional) Edit or review the `/etc/sysconfig/opa/ports` file.



Review the file. For typical single-subnet clusters, the default of 0:0 may be used. This uses the first active port on the Management Node to access the fabric. Enter one host's name per line. For further details about the file format, refer to the "Selection of Ports" section in the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

3. (Optional) Edit or review the `/etc/sysconfig/opa/allhosts` file.

Review the file. It should include the Management Node's hosts name (the TCP/IP management network name, for example `mgmthost`) and include the hosts file previously created. Enter one file per line. For example:

```
mgmthost
include /etc/sysconfig/opa/hosts
```

For further details about the file format, refer to the "Selection of Hosts" section in the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

4. Provide a brief summary of the counts of components in the fabric, including how many switch chips, hosts, and links are in the fabric.

```
opafabricinfo
```

It also indicates whether any degraded or omitted links were found. This can indicate a poorly seated or bad cable, incorrect fabric configuration, or security issues. Review the results against the expected configuration of the cluster.

If components are missing, or degraded or omitted links are found, they should be corrected. Subsequent steps aid in locating any such links.

5. (Optional) Verify each host is pingable. The subsequent `opafindgood` step is a superset of this step.

```
opapingall -p -f hostfile
```

If all hosts were found, continue to the next step. If some hosts were not found, exit the menu and review the following for those hosts that were not found:

- Host powered on and booted?
- Host connected to management network?
- Host management network IP address and network settings consistent with DNS or `/etc/hosts`?
- Management node connected to the management network?
- Management node IP address and network settings correct?
- Management network itself up (including switches, routers, and others)?
- Correct set of hosts listed in the hosts file? You may need to repeat the previous step to review and edit the file.

6. Verify each host is pingable, SSH-enabled, and active on the Intel® Omni-Path Fabric and produce a list of good hosts meeting all criteria.

```
opafindgood -R -A -Q -f hostfile
```

The following files are created in `opasorthosts` order with all duplicates removed in the `OPA_CONFIG_DIR/` directory:

- good
- alive



- running
- active
- bad
- quarantined

The resulting `good` file can then be used in as input for subsequent verification commands and to create `mpi_hosts` files for running `mpi_apps` and the HFI-SW cable test.

7. Perform a single host test on all hosts. Prior to using the `opaverifyhosts` command you must have a copy of the `hostverify.sh` in the directory pointed to by `FF_HOSTVERIFY_DIR`. If the file does not exist in that directory copy the sample file `/usr/lib/opa/samples/hostverify.sh` to the directory pointed to by `FF_HOSTVERIFY_DIR`. Review the single host test in `/usr/lib/opa/samples/hostverify.sh`. Review the settings near the top and the list of TESTS selected, edit and save as needed.

```
opaverifyhosts -k -c -u hostverify.res -T timelimit -f  
hostfile test
```

Refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for information on `opaverifyhosts`.

The results can be seen in the `$FF_RESULT_DIR/verifyhosts.res` file. A punch list of failures is also appended to the `$FF_RESULT_DIR/punchlist.csv` file. Only failures are shown on `stdout`.

8. Verify OPA Fabric status and topology.

```
opalinkanalysis -U -x snapshot_suffix all verifyall >  
$FF_RESULT_DIR/linkanalysis.res 2>&1
```

To clear error counters after generating the report, add `clearerrors` and optionally `clearhwerrors` options to the `opalinkanalysis` run.

Note: Clearing of hardware counters (`-A` option) is optional and may affect the PM and other tools See "PM Running Counters to Support `opareport`" section in the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for more information.

The following items are verified:

- Perform a fabric error analysis.
- Perform a fabric link speed error analysis.
- Check for links that are configured to run slower than supported.
- Check links that are connected with mismatched speed potential.
- Verify the fabric topology.
- Verify all aspects of the topology including links, nodes, and SMs.
- Include unexpected devices in the punchlist.

Refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for information on `opalinkanalysis`.

The results can be seen in the `$FF_RESULT_DIR/linkanalysis.res` file. A punch list of issues is appended to the `$FF_RESULT_DIR/punchlist.csv` file.



9. Verify that each host can see all the others through queries to the Subnet Administrator.

```
opahostadmin -f hostfile sacache
```

10. Verify that IPoIB is properly configured and running on all the hosts.

```
opahostadmin -f hostfile ipoibping
```

If this is unsuccessful, verify that the management host has IPoIB configured.

11. Refresh the SSH `known_hosts` file on the Management Node to include the IPoIB hostnames of all the hosts.

```
opasetupssh -p -U -f hostfile
```

12. Perform a quick check of PCIe and MPI performance through end-to-end latency and bandwidth tests.

```
opahostadmin -f hostfile mpiperfdeviation
```

This displays the results of pair-wise analysis of latency and bandwidth for the selected hosts and reports pairs outside an acceptable tolerance range. By default, performance is compared relative to other hosts in the fabric. It is assumed that all hosts selected for a given run have comparable fabric performance. Failing hosts are clearly indicated.

Intel recommends that you review the `FF_DEVIATION_ARGS` parameter in `opafastfabric.conf` and adjust it as appropriate for the cluster. The default is quite lenient to accommodate a wide range of cluster designs.

Note: This test identifies nodes whose performance is not consistent with others in the fabric. It is not intended as a benchmark of fabric latency and bandwidth. This test purposely uses techniques to reduce test runtime.

The results can be seen in the `test.res` file.

Refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for more details on `opahostadmin`.

If any hosts fail, carefully examine the failing hosts to verify the HFI models, PCIe slot used, BIOS settings, and any motherboard or BIOS settings related to devices on PCIe buses or slot speeds. Also verify the HFI and any riser cards are properly seated.

13. Baseline the present fabric configuration for use in future fabric health checks.

```
opaallanalysis -b
```

This should be performed after configuring any additional Management Nodes. Refer to [Configure and Initialize Health Check Tools](#) on page 76 for more information.

14. Perform host and/or ISL cable testing.

```
opacabletest -C -f hostfile stop start
```

The test allows for starting and stopping an extended Bit Error Rate test. The system prompts to clear hardware counters.

Note: Clearing of hardware counters (`-A` option) is optional and may affect the PM and other tools. See "PM Running Counters to Support `opareport`" section in the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for more information.



Intel recommends that you run this test for 20-60 minutes for a thorough test. While the test is running, monitor the fabric for signal integrity or stability errors using `opatop`, `opareport`, and/or the Fabric Manager GUI. Once the desired test time has elapsed, stop the test using the command:

```
opacabletest -f hostfile stop
```



11.0 Install Additional Management Nodes

If the fabric has more than one Management Node, use the procedures in this section to set up additional management nodes. Previous sections in this installation guide described how to perform basic software installation, setup, and verification on the nodes in a fabric. This section describes how to install and configure the management software itself.

Note: The following steps assume a symmetrical configuration where all Management Nodes have the same connectivity and capabilities. In asymmetrical configurations where the Management Nodes are not all connected to the same set of management networks and subnets, the files copied to each management node may need to be slightly different. For example, configuration files for fabric analysis may indicate different port numbers, or host files used for FastFabric, and MPI may need to list different hosts. For multiple-subnet configurations, refer to [Multi-Subnet Fabrics](#) on page 90.

Repeat the following steps on each additional Management Node:

1. **(All)** Upgrade the software to add additional components using the procedure documented in [Upgrade from IntelOPA-Basic to IntelOPA-IFS](#) on page 108. The Management Node must have at least Intel® Omni-Path Fabric Suite FastFabric, the Intel® Omni-Path Fabric Stack, and should have IPoIB installed and configured.

If this node is intended to also run a redundant fabric manager, install the Intel® Omni-Path Fabric Suite Fabric Manager. For MPI clusters, the Management Node should also include at least OFA openmpi or OFA mvapich2. If you plan to rebuild MPI, the Intel® Omni-Path Fabric Host Software Fabric Development package and MPI Source packages are also required.

Note: Do not uninstall or replace existing configuration files that were previously created, especially IPoIB-related configuration files.

2. **(All)** Copy the FastFabric configuration files from the initial Management Node. At a minimum, the following files should be copied:

```
/etc/sysconfig/opa/opafastfabric.conf
/etc/sysconfig/opa/ports
/etc/sysconfig/opa/topology*.xml
/etc/sysconfig/opa/hosts
/etc/sysconfig/opa/allhosts
/etc/sysconfig/opa/switches
/etc/sysconfig/opa/chassis
```

After copying the files, edit the `hosts` and `allhosts` files such that the file on each Management Node omits itself from the `hosts` files (but lists all other Management Nodes) and specifies itself in the `allhosts` file.

See *Intel® Omni-Path Fabric Suite FastFabric User Guide* Configuration Files section for a complete list of FastFabric configuration files.

3. **(All)** Copy the Fabric Manager configuration file (`/etc/sysconfig/opafm.xml`) from the initial Management Node, if the Fabric Manager is also going to be run. After copying the file, edit the file on each Management Node as needed.



Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for more information on how to configure the Fabric Manager.

4. **(Linux)** Perform the **Set Up Password-Less SSH/SCP** option in the **Host Setup via FastFabric** menu.
5. Perform the **Refresh SSH Known Hosts** option in the **Host Setup via FastFabric** menu.



12.0 Configure and Initialize Health Check Tools

This section describes how to use the Intel® Omni-Path Fabric Suite FastFabric health check tools. For more information, see the detailed discussion in the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*. The health check tools may be run on one or more Management Nodes within the cluster. Follow this procedure on each Management Node from which the health check tools are used.

1. **(All)** Edit `opafastfabric.conf` and review the following parameters:

- `FF_ANALYSIS_DIR`
- `FF_ALL_ANALYSIS`
- `FF_FABRIC_HEALTH`
- `FF_CHASSIS_CMDS`
- `FF_CHASSIS_HEALTH`
- `FF_ESM_CMDS`

`FF_ALL_ANALYSIS` should be updated to reflect the type of SM (esm or hsm).

Note: If you are running in a back-to-back configuration, `FF_ALL_ANALYSIS` should not contain `chassis`.

2. **(All)** Create `/etc/sysconfig/opa/esm_chassis`, listing the chassis that are running SMs if using Embedded SM(s) in the Intel® Omni-Path Fabric Chassis. Create the file with a list of the chassis names using the assigned TCP/IP Ethernet management port names. IP addresses can also be used; however, Intel recommends that you use names instead. For example:

```
Chassis1
Chassis2
```

For further details about the file format, refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*.

3. **(All)** Perform a health check using `opaallanalysis -e`. If any errors are encountered, resolve the errors and rerun `opaallanalysis -e` until a clean run occurs.
4. **(All)** Create a cluster configuration baseline using: `opaallanalysis -b`.
Note: This may also be done using the FastFabric TUI menu by selecting **Check Overall Fabric Health** and answering **y** to the question: Baseline present configuration? [n]:
5. **(All)** If required, schedule regular runs of `opaallanalysis` through cron or other mechanisms. Refer to the Linux* OS documentation for more information on cron. Also refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for more information about `opaallanalysis` and its automated use.



13.0 Run High Performance Linpack 2 Benchmark

This section describes how to gather MPI benchmark data. A set of common MPI benchmarks are installed as part of the installation process. A popular measure of overall performance is High Performance Linpack 2 (HPL2), which is used to rate systems on the Top 500 list.

The following steps describe how to make initial runs of HPL2 and gather preliminary baseline numbers. The defaults provided should perform within 10 – 20% of optimal HPL2 results for the cluster. Further tuning is beyond the scope of this document.

1. **(Host)** To run HPL2, first select a configuration file appropriate to your cluster. It is best to start with a small configuration to verify HPL has been properly compiled:

```
cd /usr/lib/opa/src/mpi_apps
./config_hpl2
```

This command configures a two process test run of HPL2.

2. **(Host)** Create the file `/usr/lib/opa/src/mpi_apps/mpi_hosts` listing the host names of all the hosts.

Note: Use `mpi_hosts.sample` as a starting point for creating the `mpi_hosts` file.

3. **(Host)** Run HPL2 with the command:

```
./run_hpl2
```

This initial run is a very small problem size to determine if the run is successful. Performance of this run is expected to be low.

If the initial run is successful, you are ready to move onto full scale HPL2 runs. Assorted sample HPL.dat files are provided in the `/usr/lib/opa/src/mpi_apps/hpl-config` folder. These files are a good starting point for most clusters and should get within 10 – 20% of the optimal performance for the cluster. The problem sizes used assume a cluster with 1GB of physical memory per processor. For example, for a 2 processor node, 2GB of node memory is assumed. For each cluster size, 4 files are provided:

- **t** – A very small test run (5000 problem size)
- **s** – A small problem size on the low end of optimal problem sizes
- **m** – A medium problem size
- **l** – A large problem size

The sample files can be selected using `config_hpl2`. The following command displays the pre-configured problem sizes available:

```
./config_hpl2
```



For example, to do a small run for a 256 processor cluster, that is, 128 nodes of dual CPU systems, perform the following:

1. Enter `./config_hpl2 256s`
2. Enter `./run_hpl2 256`

During these runs, use the `top` command on a node to monitor memory and CPU usage. The `xhpl` should use 98 – 99% of the CPU. If any other processes are taking more than 1 – 2%, review the host configuration and stop these extra processes if possible. HPL is very sensitive to swapping. If a lot of swapping is seen, and `xhpl` is dropping below 97% for long durations, this may indicate a problem size that is too large for the memory and OS configuration.

You can continue to tune HPL2 to refine performance. Parameters in `HPL.dat` can all affect HPL performance. In addition, the selection of compiler and BLAS Math library may also significantly affect performance. The new `HPL.dat` files may be placed in `/usr/lib/opa/src/mpi_apps/hpl-config`. Use `config_hpl` to select them and copy them to all nodes in the run. Alternately, `scpall` may be used to copy the file to all nodes. Refer to *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for more information on `scpall`.

To generate a completely custom HPL2 configuration, use `hpl_dat_gen`. It probes the current node for information on the number of cores and the amount of memory and prompts you for the desired characteristics of the run. For example, to create an HPL2 configuration that runs on 36 cores on each of 256 nodes, and uses about 50% of the memory on each node, do the following:

Enter `./hpl_dat_gen`

The system returns:

```
# of compute nodes [1]? 256
# of compute nodes [1]? 256
# of cores per node [72]? 36
# of RAM per node (in MB) [64156]?
Memory pressure (range between 0.1 and 0.9) [0.3]? 0.5
```

Enter `./run_hpl2 9216` to use this configuration. A sample of the generated file follows:

```
HPLinpack benchmark input file - 9216 processes, 0.5 memory size
Generated by hpl_dat_gen.sh.
HPL.out      output file name (if any)
6           device out (6=stdout,7=stderr,file)
1           # of problems sizes (N)
733488      Ns
1           # of NBs
168         NBs
0           PMAP process mapping (0=Row-,1=Column-major)
1           # of process grids (P x Q)
72          Ps
128         Qs
16.0        threshold
1           # of panel fact
1           PFACTs (0=left, 1=Crout, 2=Right)
1           # of recursive stopping criterium
4           NBMINs (>= 1)
1           # of panels in recursion
2           NDIVs
```



```
1      # of recursive panel fact.
2      RFACTs (0=left, 1=Crout, 2=Right)
1      # of broadcast
1      BCASTs (0=1rg,1=1rM,2=2rg,3=2rM,4=Lng,5=LnM)
1      # of lookahead depth
1      DEPTHs (>=0)
2      SWAP (0=bin-exch,1=long,2=mix)
168     swapping threshold
0      L1 in (0=transposed,1=no-transposed) form
0      U in (0=transposed,1=no-transposed) form
1      Equilibration (0=no,1=yes)
8      memory alignment in double (> 0)
```



14.0 Install Intel® Omni-Path Fabric Suite Fabric Manager GUI

This section provides the procedures to install the Fabric Manager GUI on both Windows* and Linux* platforms. Refer to the following sections:

- [Windows* Installation](#)
- [Linux* Installation](#) on page 84

14.1 Windows* Installation

14.1.1 System Requirements for a Windows* Environment

The minimum system requirements are as follows:

- Windows* operating system
 - Windows Vista* Service Pack 2 or higher
 - Windows* 7 Service Pack 1 or higher
 - Windows* 8.x
 - Windows Server* 2008 Service Pack 2 or higher
 - Windows Server* 2012
- x86 or x64 processor architecture
- Oracle* Java Runtime Environment (JRE) 1.7 or higher
- Ethernet card/local network access
- Minimum disk space requirement is 100MB. This amount includes the disk space needed for the application binaries and the application cache for topology information.
- 2GB or greater of RAM
- 1280x800 resolution (65K color depth)

14.1.2 Download the Intel® Omni-Path Fabric Suite Fabric Manager GUI

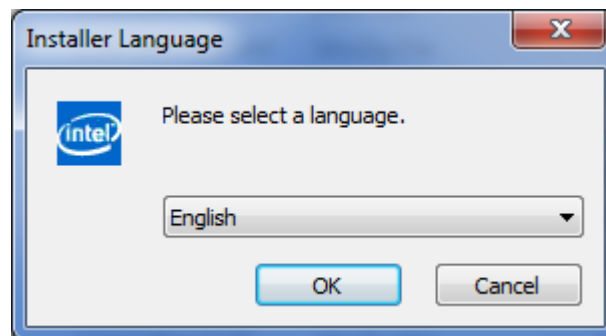
1. Go to <http://www.intel.com>.
2. Click the **Download Center** link.
3. In the **search downloads** box enter `Fabric Manager GUI`.
OR
In the **Find By Category** section:
 - a. Select **Network Connectivity**.



- b. Select **Intel® Omni-Path Architecture**.
- c. Select **Intel® Omni-Path Fabric Suite Fabric Manager**.

14.1.3 Use the Installation Wizard to Install the Fabric Manager GUI

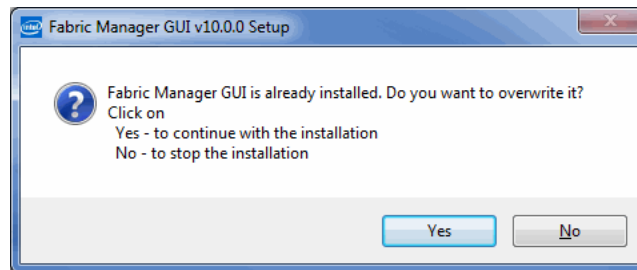
1. Double-click the IntelOPA-FMGUI.windows-x.x.x.x.x.exe file on the desktop where:
x.x.x.x.x is the version number of the Fabric Manager GUI application being installed.
2. At the User Account Control prompt: Do you want to allow the following program to make changes to this computer? Click **Yes**.
3. Select the language to be used for the installation screens and then click **OK**.



- If this is a first-time installation, the Intel® Omni-Path **Fabric Manager GUI Setup** window opens and you see the following:



- If the Fabric Manager GUI is already installed on your system, the installer prompts you with the following dialog:

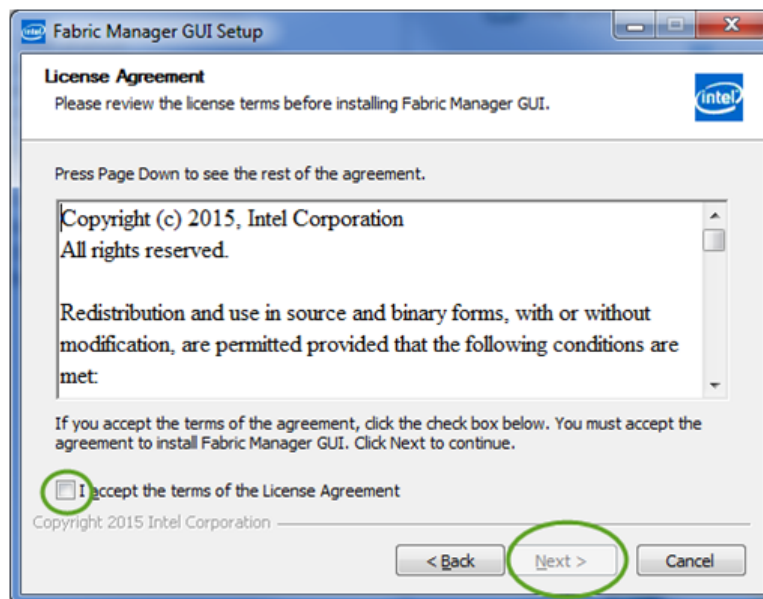


Select **Yes** to continue; the Fabric Manager GUI uninstaller starts; follow the instructions on the screen to uninstall the previous version of the application. After the uninstall is finished, the following screen displays.



4. Click **Next**.

The **License Agreement** window displays.





5. Accept the License Agreement by checking the **I accept the terms of the License Agreement** checkbox and then click **Next**.

The **Choose Install Location** window displays.

Note: Intel recommends that you use the default file location.

6. Click **Install**.

The **Installing** window displays briefly.

When the installation is complete, the **Completing Fabric Manager GUI Setup** window displays.

7. Click **Finish**.

The Fabric Manager GUI is installed.

14.1.4 Changing Database Location

A customer can change the database location after the installation by following these steps.

1. Create file `settings.xml` under the folder `<app_data_path>\Intel\FabricManagerGUI` where `<app_data_path>` is a user's application data folder. Under Windows 7/8, it's `C:\Users\<user_name>\AppData\Roaming`
2. Add the following contents:

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE properties SYSTEM "http://java.sun.com/dtd/properties.dtd">
<properties>
  <comment>FM GUI Application configuration settings</comment>
  <entry key="db.connection.url">jdbc:hsqldb:file:c:\temp\db\mydbname;

  hsqldb.result_max_memory_rows=1000;hsqldb.default_table_type=cached;
  hsqldb.log_size=50;</entry>
</properties>
```

Change `c:\temp\db\mydbname` to the location for the database file.

Note: Omitting the other options will cause database performance issues in HSQLDB. Please only change the database file path.

3. Launch FM GUI, which will then use the new location for database storage.

14.2 Linux* Installation

14.2.1 System Requirements for a Linux* Environment

The minimum system requirements for the Linux* installation are as follows:

- Linux* operating system
 - Red Hat* Enterprise Linux* 5.5 or higher
 - SUSE* Linux* Enterprise Server 10 or higher
- x86 or x64 processor architecture
- Oracle* Java* Runtime Environment (JRE) 1.7 or higher
- X Window System



- Ethernet card/local network access
- Minimum disk space requirement is 100MB. This amount includes the disk space needed for the application binaries and the application cache for topology information.
- 2GB or greater of RAM
- 1280x800 resolution (65K color depth)

14.2.2 Download the Intel® Omni-Path Fabric Suite Fabric Manager GUI

1. Go to <http://www.intel.com>.
2. Click the **Download Center** link.
3. In the **search downloads** box enter Fabric Manager GUI.
OR
In the **Find By Category** section:
 - a. Select **Network Connectivity**.
 - b. Select **Intel® Omni-Path Architecture**.
 - c. Select **Intel® Omni-Path Fabric Suite Fabric Manager**.

14.2.3 Use RPM to Install the Fabric Manager GUI

1. Log in to the server where the Fabric Manager GUI will be installed, as a user with root access.
2. Open a Terminal window in X Windows.
3. Change directories to the directory where you downloaded the .rpm file.
4. Enter the following command:


```
rpm -iv IntelOPA-FMGUI-linux-x.x.x.x.x.noarch.rpm
```

 where x.x.x.x.x is the version number of the Fabric Manager GUI application being installed.
5. Wait for the successful completion of the command.

The Fabric Manager GUI is installed.

14.2.4 Changing Database Location

A customer can change the database location after the installation by following these steps.

1. Create file `settings.xml` under the folder `~\.Intel\FabricManagerGUI`
2. Add the following contents:

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE properties SYSTEM "http://java.sun.com/dtd/properties.dtd">
<properties>
  <comment>FM GUI Application configuration settings</comment>
  <entry key="db.connection.url">jdbc:hsqldb:file:/var/temp/db/mydbname;
```



```
hsqldb.result_max_memory_rows=1000;hsqldb.default_table_type=cached;
      hsqldb.log_size=50;</entry>
</properties>
```

Change `/var/temp/db/mydbname` to the location for the database file.

Note: Omitting the other options will cause database performance issues in HSQLDB. Please only change the database file path.

3. Launch FM GUI, which will then use the new location for database storage.

14.3 Prepare Fabric Manager for Fabric Manager GUI

Fabric Manager GUI requires a running fabric with FE available. Before running Fabric Manager GUI, ensure that Fabric Manager has the proper configuration and is running normally.

Perform the following:

1. Change the `<SubnetPrefix>` for `fm0` to `0xfe80000000001000` in `opafm.xml`.

```
<!-- Shared Instance config, applies to all components: SM, PM and FE -->
<Shared>
  <!-- Fm.Shared.Start controls overall startup of the Instance. -->
  <!-- If 0, none of the components in the Instance are started. -->
  <!-- If 1, instance is enabled and Fm.Sm.Start, Fm.Pm.Start, etc -->
  <!-- control startup of each manager. The default for each manager -->
  <!-- is defined by Common.Sm.Start, Common.Pm.Start, etc -->
  <!-- ESM does not support Start via XML configuration. Use CLI commands -->
  <Start>1</Start>
  <!-- <StartupRetries>5</StartupRetries> -->
  <!-- <StartupStableWait>10</StartupStableWait> -->

  <!-- Name, Hfi, Port, and PortGUID are ignored for ESM since they -->
  <!-- are automatically set -->
  <Name>fm0</Name> <!-- also for logging with _sm, _fe, _pm appended -->
  <Hfi>1</Hfi> <!-- local HFI to use for FM instance, 1=1st HFI -->
  <Port>1</Port> <!-- local HFI port to use for FM instance, 1=1st Port -->
  <PortGUID>0x0000000000000000</PortGUID> <!-- local port to use for FM -->
  <SubnetPrefix>0xfe80000000001000</SubnetPrefix> <!-- should be unique -->
```

Refer to [Multi-Rail Usage](#) on page 94 for additional instructions on setting up multiple rails in a single or multi-subnet.

2. Ensure that the FE is enabled in `opafm.xml`:

In the `<FE>` section, "Start" is 1.

```
<Fe>
  <Start>1</Start><!-- default FE startup for all instances -->
```

3. Ensure that FE has proper security setup:

If you want to run FE without SSL, ensure that the `SslSecurityEnable` is 0. And when you run Fabric Manager GUI, set up FE host as unsecured (uncheck the "Secure" checkbox):

```
<!-- OpenSSL FE network security parameters -->
  <SslSecurityEnable>0</SslSecurityEnable>
```



If you want to run FE with SSL, ensure that you set `SslSecurityEnable` to 1 and follow steps in *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* in sections and appendices titled "Out-of-Band Security," "Core-Level Public Key Infrastructure (PKI) Best Practices Guidelines," "Advanced-Level Public Key Infrastructure Best Practices Guidelines," and "SSL Key Creation for Fabric Manager GUI" to set up keys and `opafm.xml` properly. When you run Fabric Manager GUI, set up FE host as "Secure" with the key files described in "SSL Key Creation for Fabric Manager GUI."

4. Ensure that PM is enabled in `opafm.xml`:

In the `<Pm>` section, "Start" is 1.

```
<Pm>
  <!-- ESM does not support Start via XML configuration. Use CLI commands -->
  <Start>1</Start> <!-- default PM startup for all instances -->
```

5. If you made any changes to `opafm.xml`, ensure that you restart `opafm` so that your changes take effect.

14.4 Start the Fabric Manager GUI Application

14.4.1 Windows* Procedure

The following procedure provides the steps to start the Intel® Omni-Path Fabric Suite Fabric Manager GUI application.

1. From the Windows* **Start** menu, select **All Programs**.
2. Select **Intel > Omni-Path > Fabric Manager GUI**

The Fabric Manager GUI application starts.

14.4.2 Linux* Procedure

The following procedure starts the Intel® Omni-Path Fabric Suite Fabric Manager GUI application from a terminal window on X Windows.

1. Open a terminal window in X Windows.
2. Type the following command:

```
fmgui
```

The Fabric Manager GUI application starts.

Alternatively, the following steps start the Fabric Manager GUI application from the K Desktop Environment (KDE).

1. From the K Menu, select **Applications**.
2. Select the **Fabric** folder.
3. Select **Fabric Manager GUI**.

The Fabric Manager GUI application starts.

14.5 Configure Startup Options

Note: Intel recommends that you accept the default settings.



Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager GUI Online Help* for procedures to set user preferences.

14.6 Uninstall the Intel® Omni-Path Fabric Suite Fabric Manager GUI

The Fabric Manager GUI must be closed for the uninstall to be successful. The uninstall program does not warn you if the application is open. Warnings may be received at the end of the uninstall process stating that certain files have not been removed or it might just fail.

14.6.1 Windows* Procedure

1. From the Windows* **Start** menu, select **Control Panel**.
2. Under the **Programs** category, select **Uninstall a program**.
3. In the program list, select **Intel Fabric Manager GUI**.
4. Click the **Uninstall** button at the top of the list.
5. Follow the instructions on the uninstall window.

The Fabric Manager GUI is uninstalled.

14.6.2 Linux* Procedure

1. Log in to the server where Fabric Manager GUI is installed as a user with root access.
2. Open a Terminal window in X Windows.
3. Type the following command:

```
# rpm -e IntelOPA-FMGUI-linux-x.x.x.x.x.noarch.rpm
```

where `x.x.x.x.x` is the version number of the Fabric Manager GUI application being uninstalled.
4. Wait for the successful completion of the command.

The Fabric Manager GUI is uninstalled.

14.7 Clearing the Fabric Manager GUI Cache

The Fabric Manager GUI uses local disk storage to keep a cache of fabric information. The size of this cache depends on the size of the fabric(s) configured in the Fabric Manager GUI and the refresh rate used for performance monitoring. You can reclaim this disk storage by using the Clear FM GUI Cache. Be aware that clearing the cache also clears subnet configurations. Also, this option is only available to those users who have used the Fabric Manager GUI.

14.7.1 Cleanup Procedure for Windows*

1. Select Windows **Start** menu.
2. Select **All Programs**.
3. Select **Intel**.
4. Select **Omni-Path**.



5. Select **Clear FM GUI Cache**.

14.7.2 Cleanup Procedure for Linux*

1. In KDE or GNOME, select **Applications**.
2. Select **Fabric**; if not available, try the **System** category.
3. Select **Clear FM GUI Cache**.

Alternatively, use the following steps to clear the FM GUI cache from a terminal window on X Window:

1. Open a Terminal window in X Windows.
2. Enter the following commands:

```
cd ~/.Intel/FabricManagerGUI  
fmguiclear.sh
```

The Fabric Manager GUI cache is cleared.



15.0 Additional Installation and Setup Tasks

This chapter describes additional installation and setup tasks that, depending on your configuration, you may need to perform during or after software installation.

15.1 Multi-Subnet Fabrics

This section describes differences in the installation procedure that are relevant for multi-subnet fabrics. Intel® Omni-Path Fabric Suite FastFabric supports management of both single-subnet fabric and multi-subnet fabrics.

When operating a multi-subnet fabric, a subnet manager (SM) is required for each subnet. An SM may be run within switches within each subnet, or a host-based SM may be run. A host-based SM can manage multiple subnets (assuming the host server is connected to more than one subnet).

A number of combinations are possible for multi-subnet fabrics:

1. **All subnets are completely independent (except for any interconnecting routers):** If a separate FastFabric node is being used per subnet, and servers are not installed in more than one subnet, the individual subnets can be treated completely separately. In this case, follow all the previous FastFabric instructions for each fabric.
2. **The subnets are primarily independent:** If the only components common to more than one subnet are the FastFabric nodes (and possibly SM nodes) and no servers are installed in more than one subnet, refer to the following instructions for [Primarily Independent Subnets](#).
3. **The subnets are overlapping:** If multiple components are common to more than one subnet, such as FastFabric node(s), servers, and others, refer to the following instructions for [Overlapping Subnets](#).

15.1.1 Primarily Independent Subnets

If the FastFabric node (and possible SM nodes) is the only common server between subnets, FastFabric may be used to assist in server installation and fabric operation. Follow the installation instructions outlined in [Install the Intel® Omni-Path Software](#) with the following adjustments:

From [Design the Fabric](#), design the cabling such that the FastFabric node is connected to each subnet it manages. The FastFabric node must also have a management network path to all the nodes in all the subnets that it manages. As part of the design, consider where routes between subnets are wanted between storage routers, IPoIB routers, and other devices.

[Set Up the Fabric](#) can be performed as per the instructions. When installing the IFS software on the Fabric Management Node, IPoIB must be configured such that each subnet is an independent IPoIB network interface, typically with different IP subnets. Refer to the *Intel® Omni-Path Fabric Host Software User Guide* for more information on configuring IPoIB.



[Configure Intel® Omni-Path Chassis](#) can be performed as per the instructions. When creating the chassis file, list all Intel internally-managed switches in all subnets. If required, additional files may also be created per subnet that list only the Intel chassis in each subnet. When editing the ports file, list all the Fabric Management Node ports that access the managed fabrics. If required, additional files may be created per subnet that list only the Fabric Management Node port connected to the given managed fabric.

[Configure Chassis Fabric Manager \(FM\)](#) can be performed as per the instructions. At least one subnet manager is required per subnet. Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for more information on how to configure a host SM node to manage more than one subnet.

[Configure Firmware on the Externally Managed Intel® Omni-Path Switches](#) can be performed as per the instructions. When editing the ports file, list all the Fabric Management Node ports that access the managed fabrics. If required, additional files may be created per subnet that list only the Fabric Management Node port connected to the given managed fabric. If required, the switches file may specify a hfi:port per switch. However, if hfi:port is not specified, all the hfi:port specified in the ports file are searched to locate the given switch's Node GUID.

[Install Host Software on the Remaining Hosts Using the FastFabric TUI](#) can be performed as per the instructions. When creating the `hosts` file, list the hosts in all subnets except the Fabric management node where FastFabric is being run. If required, additional files may be created per subnet that list the hosts in each subnet, except the Fabric Management Node.

[Verify Intel® Omni-Path Fabric Host Software on the Remaining Servers Using the FastFabric TUI](#) has the following adjustments from the instructions:

- **(All)** Create the `allhosts` file as per the instructions. Next, create additional files per subnet that list all the hosts in each subnet including the Fabric Management Node. When editing the ports file, list all the Fabric Management Node ports that access the managed fabrics. If required, additional files may be created per subnet that list only the Fabric Management Node port connected to the given managed fabric.
- **(All)** [Summary of Fabric Components](#) can be performed as per the instructions.
- **(All)** [Verify Hosts Pingable, SSHable, and Active](#) can be performed as per the instructions.
- **(All)** [Verify OPA Fabric Status and Topology](#) can be performed as per the instructions.
- **(Host)** [Verify Hosts See Each Other](#) can be run for each subnet by using the `allhosts` files specific to each subnet (i.e., those listing only hosts in a single subnet).
- **(Host)** [Verify Hosts Ping via IPoIB](#) may be run per the instructions.
- **(Host)** [Check MPI Performance](#) can be run for each subnet by using the `allhosts` files specific to each subnet (i.e., those listing only the hosts in a single subnet).

[Install Additional Management Nodes](#) can be performed as per the instructions. When copying FastFabric configuration files to the additional Fabric Management Nodes, be sure to also copy the additional `hosts`, `chassis`, and `allhosts` files that were created per subnet.



Note: In asymmetrical configurations where the Fabric Management Nodes are not all connected to the same set of subnets, the files copied to each management node may need to be slightly different. For example, configuration files for `opafabricanalysis` may indicate different port numbers or host files used for FastFabric, and MPI may need to list different hosts.

[Configure and Initialize Health Check Tools](#) can be performed as per the instructions. Additionally, make sure the `/etc/sysconfig/opa/ports` file lists each of the Fabric Management Node local HFIs and ports that are connected to a unique subnet. When running `opareports`, `opafabricinfo`, `opafabricanalysis`, or `opaallanalysis`, the default is to use the `ports` file. If required, the `-p` and `-t` options or the `PORTS/PORTS_FILE` environment variables may be used to specify all the HFIs and ports on the Fabric Management Node such that all subnets are checked. Similarly, the `esm_chassis` and `chassis` files used must list all relevant Intel chassis in all subnets.

[Run High Performance Linpack 2 Benchmark](#) can be run for each subnet by creating `mpi_hosts` files specific to each subnet, that is, only listing hosts in a single subnet.

[Upgrade the Management Node](#) can be performed as per the instructions.

15.1.2 Overlapping Subnets

If multiple components are common between subnets (in addition to the Fabric Management Nodes), FastFabric may be used to assist in server installation and fabric operation. Follow the installation instructions outlined in [Install the Intel® Omni-Path Software](#) with the following adjustments:

From [Design the Fabric](#), design the cabling such that the FastFabric node is connected to each subnet it manages. The FastFabric node must also have a management network path to all the nodes in all the subnets it manages. As part of the design, consider where routes between subnets are required, between routers, IPoIB routers, and other devices.

[Set Up the Fabric](#) can be performed as per the instructions. When installing the IFS software on the Fabric Management Node, IPoIB must be configured such that each subnet is an independent IPoIB network interface, typically with different IP subnets. Refer to the *Intel® Omni-Path Fabric Host Software User Guide* for more information on configuring IPoIB.

[Configure Intel® Omni-Path Chassis](#) can be performed as per the instructions. When creating the chassis file, list all Intel internally-managed switches in all subnets. If required, additional files may be created per subnet that list only the Intel chassis in each subnet. When editing the ports file, list all the Fabric Management Node ports that access the managed fabrics. If required, additional files may be created per subnet that list only the Fabric Management Node port connected to the given managed fabric.

[Configure Chassis Fabric Manager \(FM\)](#) can be performed as per the instructions. At least one subnet manager is required per subnet. Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for more information on how to configure a host-based SM node to manage more than one subnet.

[Configure Firmware on the Externally Managed Intel® Omni-Path Switches](#) can be performed as per the instructions. When editing the ports file, list all the Fabric Management Node ports that access the managed fabrics. If required, additional files



may be created per subnet that list only the Fabric Management Node port connected to the given managed fabric. If required, the switches file may specify a `hfi:port` per switch. However, if `hfi:port` is not specified, all the `hfi:port` entries specified in the ports file are searched to locate the given switch's node GUID.

[Install Host Software on the Remaining Hosts Using the FastFabric TUI](#) can be performed as per the instructions. When creating the `hosts` file, list all the hosts in all subnets except the Fabric Management Node where FastFabric is being run. If required, additional files may also be created per subnet that list the hosts in each subnet, except the Fabric Management Node.

For hosts that are connected to more than one subnet, IPoIB must be configured such that each subnet is an independent IPoIB network interface, typically with different IP subnets. Refer to the *Intel® Omni-Path Fabric Host Software User Guide* for more information on configuring IPoIB.

[Verify Intel® Omni-Path Fabric Host Software on the Remaining Servers Using the FastFabric TUI](#) has the following adjustments from the instructions:

- **(All)** Create the `allhosts` file per the instructions. Next, create additional files per subnet that list all the hosts in each subnet including the Fabric Management Node. When editing the ports file, list all the Fabric Management Node ports which access the managed fabrics. If required, additional files may also be created per subnet that list only the Fabric Management Node port connected to the given managed fabric.
- **(All)** [Summary of Fabric Components](#) can be performed per the instructions.
- **(All)** [Verify Hosts Pingable, SSHable, and Active](#) can be performed as per the instructions.
- **(All)** [Verify OPA Fabric Status and Topology](#) can be performed as per the instructions.
- **(Host)** [Verify Hosts See Each Other](#) can be run for each subnet by using the `allhosts` files specific to each subnet, that is, those only listing hosts in a single subnet.
- **(Host)** [Verify Hosts Ping via IPoIB](#) may be run per the instructions.
- **(Linux)** [Refresh SSH Known Hosts](#) may be run per the instructions.
- **(Host)** [Check MPI Performance](#) can be run for each subnet by using the `allhosts` files specific to each subnet, that is, those listing only the hosts in a single subnet. This is currently not available for OFED.

[Install Additional Management Nodes](#) can be performed as per the instructions. When copying FastFabric configuration files to the additional Fabric Management nodes, be sure to also copy the additional `hosts`, `chassis`, and `allhosts` files created per subnet.

Note: In asymmetrical configurations, where the Fabric Management Nodes are not all connected to the same set of subnets, the files copied to each management node may need to be slightly different. For example, configuration files for `opafabricanalysis` indicating different port numbers or host files used for FastFabric and MPI may need to list different hosts.

[Configure and Initialize Health Check Tools](#) can be performed per the instructions. In addition, make sure the `/etc/sysconfig/opa/ports` file lists the Fabric Management Node local HFIs and ports that are connected to a unique subnet. When

running `opareport`, `opafabricinfo`, `opafabricanalysis`, or `opaallanalysis`, the default is to use the ports file. If required, the `-p` and `-t` options or the `PORTS/PORTS_FILE` environment variable may be used to specify all the HFIs and ports on the Fabric Management Node, such that all subnets are checked. Similarly, the `esm_chassis` and `chassis` files used must list all relevant Intel chassis in all subnets.

[Run High Performance Linpack 2 Benchmark](#) can be run for each subnet by creating `mpi_hosts` files specific to each subnet, that is, only listing hosts in a single subnet.

[Upgrade the Management Node](#) can be performed per the instructions.

15.2 Multi-Rail Usage

This section provides an overview and instructions for configuring the common multi-rail scenarios.

15.2.1 Multi-Rail Overview

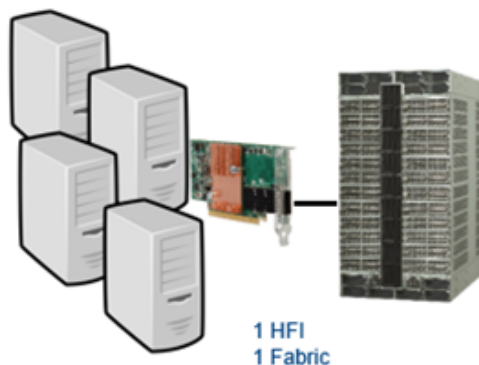
A multi-rail configuration provides load balancing and failover capabilities, adding a higher degree of fabric redundancy. If one HFI or an entire subnet fails, traffic can be moved to the remaining switches.

The multi-rail feature can be applied to a single subnet or multiple subnets. By enabling multi-rail, a process can use multiple network interface cards (HFIs) to transfer messages.

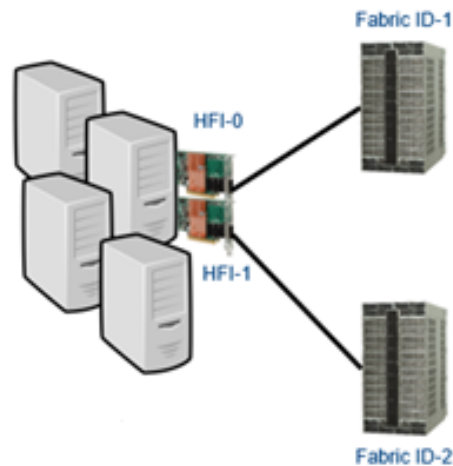
Note: Subnets can also be referred to as *planes* or *fabrics*. Rails are also referred to as *HFIs*.

Three basic scenarios include:

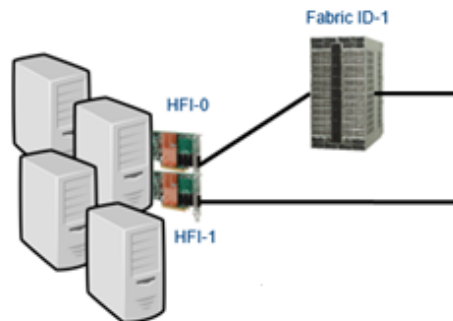
- Single-rail in a single subnet: This scenario, shown in the following figure, consists of one HFI in a server connected to one subnet. This is the default configuration during installation. This configuration provides the performance required by most applications in use today.



- Dual-rail in dual subnets: This scenario, shown in the following figure, consists of two HFIs in the same server connected to separate subnets. This configuration provides improved MPI message rate, latency, and bandwidth to the node as well as flexibility for configuring failover and load-balancing.



- Dual-rail in a single subnet: This scenario, shown in the following figure, consists of two HFIs in the same server connected to the same subnet. This configuration also provides improved MPI message rate, latency, and bandwidth to the node, but only basic HFI failover capabilities if configured.



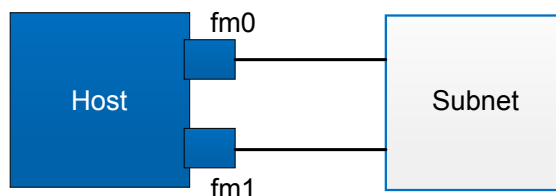
Note: Other multi-rail scenarios can be configured up to the supported number of FM instances. These scenarios are out of scope at this time.

15.2.2 Setting Up Single or Dual Rails for a Single Subnet

Support for single rails or dual rails in a single subnet is the default scenario expected during installation. The driver will detect that there are one or two HFIs connected to the host on the same fabric. To initiate these configurations, you need to change the SubnetPrefix for `fm0` as described in the steps below.

PREREQUISITES:

- HFIs have been installed in the host servers.
- HFIs have been cabled as shown below:



Note: In the figure above, fm0 and fm1 refer to the installed HFIs.

- The single subnet has already been configured during software installation.

Perform the following steps to set the SubnetPrefix:

1. On the Host, open the `/etc/sysconfig/opafm.xml` file for editing.
2. Search on "<Shared>" to review the fm0 settings.

An example of the shared instance for fm0 with key items in bold is shown below.

```

<!-- Shared Instance config, applies to all components: SM, PM and FE -->
<Shared>
  <!-- Fm.Shared.Start controls overall startup of the Instance. -->
  <!-- If 0, none of the components in the Instance are started. -->
  <!-- If 1, instance is enabled and Fm.Sm.Start, Fm.Pm.Start, etc -->
  <!-- control startup of each manager. The default for each manager -->
  <!-- is defined by Common.Sm.Start, Common.Pm.Start, etc -->
  <!-- ESM does not support Start via XML configuration. Use CLI commands -->
  <Start>1</Start>
  <!-- <StartupRetries>5</StartupRetries> -->
  <!-- <StartupStableWait>10</StartupStableWait> -->

  <!-- Name, Hfi, Port, and PortGUID are ignored for ESM since they -->
  <!-- are automatically set -->
  <Name>fm0</Name> <!-- also for logging with _sm, _fe, _pm appended -->
  <Hfi>1</Hfi> <!-- local HFI to use for FM instance, 1=1st HFI -->
  <Port>1</Port> <!-- local HFI port to use for FM instance, 1=1st Port -->
  <PortGUID>0x0000000000000000</PortGUID> <!-- local port to use for FM -->
  <SubnetPrefix>0xfe80000000000000</SubnetPrefix> <!-- should be unique -->

  <!-- Overrides of the Common.Shared parameters if desired -->
  <!-- ESM does not support LogFile -->
  <!-- <LogFile>/var/log/fm0_log</LogFile> --> <!-- log for this instance -->
</Shared>

```

3. Change the <SubnetPrefix> for fm0 to **0xfe80000000001000**.

An example of the change is shown below.

```

<!-- Shared Instance config, applies to all components: SM, PM and FE -->
<Shared>
  <!-- Fm.Shared.Start controls overall startup of the Instance. -->
  <!-- If 0, none of the components in the Instance are started. -->
  <!-- If 1, instance is enabled and Fm.Sm.Start, Fm.Pm.Start, etc -->
  <!-- control startup of each manager. The default for each manager -->
  <!-- is defined by Common.Sm.Start, Common.Pm.Start, etc -->
  <!-- ESM does not support Start via XML configuration. Use CLI commands -->
  <Start>1</Start>
  <!-- <StartupRetries>5</StartupRetries> -->
  <!-- <StartupStableWait>10</StartupStableWait> -->

```




```
<!-- Name, Hfi, Port, and PortGUID are ignored for ESM since they -->
<!-- are automatically set -->
<Name>fm0</Name> <!-- also for logging with _sm, _fe, _pm appended -->
<Hfi>1</Hfi> <!-- local HFI to use for FM instance, 1=1st HFI -->
<Port>1</Port> <!-- local HFI port to use for FM instance, 1=1st Port -->
<PortGUID>0x0000000000000000</PortGUID> <!-- local port to use for FM -->
<SubnetPrefix>0xfe80000000001000</SubnetPrefix> <!-- should be unique -->

<!-- Overrides of the Common.Shared parameters if desired -->
<!-- ESM does not support LogFile -->
<!-- <LogFile>/var/log/fm0_log</LogFile> --> <!-- log for this instance -->
</Shared>
```

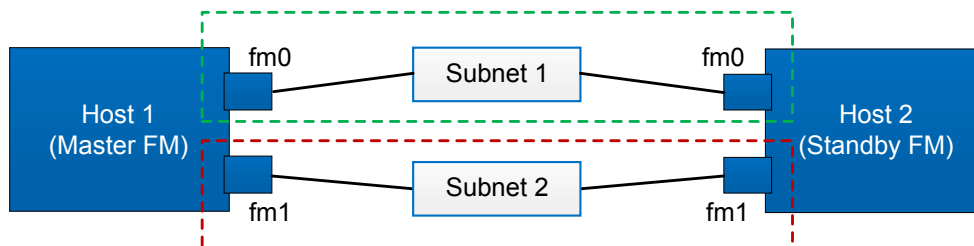
4. Save the opafm.xml file.
 5. Restart the Fabric Manager using `systemctl restart opafm`.
- Host 1 starts up as the master Fabric Manager.

15.2.3 Configuring Dual Rails for Dual Subnets

A common multi-rail, multi-subnet scenario is to set up dual rails for dual subnets where you to configure four HFIs in two Hosts to two different subnets to provide failover and load-balancing functions.

PREREQUISITES:

- HFIs have been installed in the host servers.
- HFIs have been cabled as shown below:



Note: In the figure above, `fm0` and `fm1` refer to the installed HFIs.

- Two subnets have already been configured during software installation.

Perform the following steps to set up the master Fabric Manager on Host 1 and standby Fabric Manager on Host 2:

1. On Host 1, open the `/etc/sysconfig/opafm.xml` file for editing.
2. Search on "`<Shared>`" to review the `fm0` settings.

An example of the shared instance for `fm0` with key items in bold is shown below.

```
<!-- Shared Instance config, applies to all components: SM, PM and FE -->
<Shared>
  <!-- Fm.Shared.Start controls overall startup of the Instance. -->
  <!-- If 0, none of the components in the Instance are started. -->
  <!-- If 1, instance is enabled and Fm.Sm.Start, Fm.Pm.Start, etc -->
  <!-- control startup of each manager. The default for each manager -->
  <!-- is defined by Common.Sm.Start, Common.Pm.Start, etc -->
  <!-- ESM does not support Start via XML configuration. Use CLI commands -->
  <Start>1</Start>
```



```
<!-- <StartupRetries>5</StartupRetries> -->
<!-- <StartupStableWait>10</StartupStableWait> -->

<!-- Name, Hfi, Port, and PortGUID are ignored for ESM since they -->
<!-- are automatically set -->
<Name>fm0</Name> <!-- also for logging with _sm, _fe, _pm appended -->
<Hfi>1</Hfi> <!-- local HFI to use for FM instance, 1=1st HFI -->
<Port>1</Port> <!-- local HFI port to use for FM instance, 1=1st Port -->
<PortGUID>0x0000000000000000</PortGUID> <!-- local port to use for FM -->
<SubnetPrefix>0xfe80000000000000</SubnetPrefix> <!-- should be unique -->

<!-- Overrides of the Common.Shared parameters if desired -->
<!-- ESM does not support LogFile -->
<!-- <LogFile>/var/log/fm0_log</LogFile> --> <!-- log for this instance -->
</Shared>
```

3. Change the `<SubnetPrefix>` for fm0 to **0xfe80000000001000**.

An example of the change is shown below.

```
<!-- Shared Instance config, applies to all components: SM, PM and FE -->
<Shared>
  <!-- Fm.Shared.Start controls overall startup of the Instance. -->
  <!-- If 0, none of the components in the Instance are started. -->
  <!-- If 1, instance is enabled and Fm.Sm.Start, Fm.Pm.Start, etc -->
  <!-- control startup of each manager. The default for each manager -->
  <!-- is defined by Common.Sm.Start, Common.Pm.Start, etc -->
  <!-- ESM does not support Start via XML configuration. Use CLI commands -->
  <Start>1</Start>
  <!-- <StartupRetries>5</StartupRetries> -->
  <!-- <StartupStableWait>10</StartupStableWait> -->

  <!-- Name, Hfi, Port, and PortGUID are ignored for ESM since they -->
  <!-- are automatically set -->
  <Name>fm0</Name> <!-- also for logging with _sm, _fe, _pm appended -->
  <Hfi>1</Hfi> <!-- local HFI to use for FM instance, 1=1st HFI -->
  <Port>1</Port> <!-- local HFI port to use for FM instance, 1=1st Port -->
  <PortGUID>0x0000000000000000</PortGUID> <!-- local port to use for FM -->
  <SubnetPrefix>0xfe80000000001000</SubnetPrefix> <!-- should be unique -->

  <!-- Overrides of the Common.Shared parameters if desired -->
  <!-- ESM does not support LogFile -->
  <!-- <LogFile>/var/log/fm0_log</LogFile> --> <!-- log for this instance -->
</Shared>
```

4. Find the next occurrence of `<Shared>`.

An example of the shared instance for fm1 with key items in bold is shown below.

```
<!-- Shared Instance config, applies to all components: SM, PM and FE -->
<Shared>
  <Start>0</Start> <!-- Overall Instance Startup, see fm0 for more info -->
  <Name>fm1</Name> <!-- also for logging with _sm, _fe, _pm appended -->
  <Hfi>1</Hfi> <!-- local HFI to use for FM instance, 1=1st HFI -->
  <Port>2</Port> <!-- local HFI port to use for FM instance, 1=1st Port -->
  <PortGUID>0x0000000000000000</PortGUID> <!-- local port to use for FM -->
  <SubnetPrefix>0xfe80000000001001</SubnetPrefix> <!-- should be unique -->
  <!-- Overrides of the Common.Shared or Fm.Shared parameters if desired -->
  <!-- <LogFile>/var/log/fm1_log</LogFile> --> <!-- log for this instance -->
  <!-- <StartupRetries>5</StartupRetries> -->
  <!-- <StartupStableWait>10</StartupStableWait> -->
</Shared>
```

5. Edit fm1 settings as shown below:

- Start = 1 (Initializes second HFI)



- HFI = 2
 - Port = 1
6. Verify that the SubnetPrefix is the second, unique subnet ID set up during software installation.

An example of the final settings for fm1 is shown below.

```
<!-- Shared Instance config, applies to all components: SM, PM and FE -->
<Shared>
  <Start>1</Start> <!-- Overall Instance Startup, see fm0 for more info -->
  <Name>fm1</Name> <!-- also for logging with _sm, _fe, _pm appended -->
  <Hfi>2</Hfi> <!-- local HFI to use for FM instance, 1=1st HFI -->
  <Port>1</Port> <!-- local HFI port to use for FM instance, 1=1st Port -->
  <PortGUID>0x0000000000000000</PortGUID> <!-- local port to use for FM -->
  <SubnetPrefix>0xfe80000000001001</SubnetPrefix> <!-- should be unique -->
  <!-- Overrides of the Common.Shared or Fm.Shared parameters if desired -->
  <!-- <LogFile>/var/log/fm1_log</LogFile> --> <!-- log for this instance -->
  <!-- <StartupRetries>5</StartupRetries> -->
  <!-- <StartupStableWait>10</StartupStableWait> -->
</Shared>
```

7. Save the opafm.xml file.
8. Restart the Fabric Manager on Host 1 using `systemctl restart opafm`.
Host 1 starts up as the master Fabric Manager.
9. Run `service opafm status` to verify that the two instances of the FM are running.
10. Copy the opafm.xml file to Host 2.
11. Restart the Fabric Manager on Host 2 using `systemctl restart opafm`.
Host 2 starts up in standby mode.



16.0 Installation Verification and Additional Settings

This section provides instructions for verifying that the software has been properly installed, the Intel® Omni-Path Fabric drivers are loaded, and that the fabric is active and ready to use. Information on the Intel HFIs and Performance tuning is also provided.

16.1 LED Link and Data Indicators

For information on the LEDs function as beaconing, link, and data indicators refer to *Intel® Omni-Path Fabric Switches Hardware Installation Guide* and/or *Intel® Omni-Path Host Fabric Interface Installation Guide*.

16.2 Adapter and Other Settings

The following settings can be adjusted for better performance.

- **Ensure that an MTU of 8k bytes, is used with the Intel® Omni-Path Host Fabric Interface.** 8K MTU is enabled in the Intel® Omni-Path driver by default. To change this setting for the driver, refer to the *Intel® Omni-Path Fabric Performance Tuning User Guide*.
- **Use a PCIe Max Read Request size of at least 512 bytes with the Intel® Omni-Path Host Fabric Interface.** The Intel® Omni-Path Host Fabric Interface can support sizes from 128 bytes to 4096 bytes in powers of two. This value is typically set in the BIOS.
- **Use a PCIe MaxPayload size of 256, where available, with the Intel® Omni-Path Host Fabric Interface.** The Intel® Omni-Path Host Fabric Interface can support 128, 256, or 512 bytes. This value is typically set by the BIOS as the minimum value supported both by the PCIe card and the PCIe root complex.

Check the PCIe bus width. If slots have a smaller electrical width than mechanical width, lower than expected performance may occur. Use the following command to check PCIe Bus link speed and width.

```
lspci -vv -d :24f0 | grep LnkSta:
```

16.3 ARP Neighbor Table Setup for Large Clusters

On large clusters or subnets, the ARP neighbor table may overflow and produce a neighbor table overflow message to `/var/log/messages` along with other effects such as ping failing. The Intel® Omni-Path Fabric Suite (IFS) includes a script, enabled by default during installation, that automatically tunes the ARP Neighbor Table when invoked. The script, `opa-arptbl-tuneup`, is run once by `opa.service` when a node starts or restarts, but it can also be run manually.

To run `opa-arptbl-tuneup` manually, it must execute at the root. The syntax for the parameter options is as follows:



- `opa-arptbl-tuneup start` - adjust kernel ARP table size
- `opa-arptbl-tuneup stop` - restore previous configuration
- `opa-arptbl-tuneup status` - check if original table size was changed
- `opa-arptbl-tuneup restart` - stop then start
- `opa-arptbl-tuneup force-reload` - stop then start
- `opa-arptbl-tuneup --help` - usage information for the script

There are two ARP/Neighbor Tables in the kernel, one for IPv4 networks and one for IPv6 networks. The operating system uses the table for the particular network in use. There are three threshold parameters for each table: `gc_thresh1`, `gc_thresh2`, and `gc_thresh3`. You can check the present threshold level 1 by entering the command:

```
cat /proc/sys/net/ipv4/neigh/default/gc_thresh1
```

You may repeat for `gc_thresh2` and `gc_thresh3`.

16.4 SM Loop Test

The SM loop test is a diagnostic test provided in the Fabric Manager. As part of this test, the Subnet Manager stress tests inter-switch links (ISLs) by continuously passing traffic through them. Other tools, like FastFabric, can be used to monitor the links for signal integrity issues or other errors. The advantage of the loop test is that it provides a guaranteed way to test all of the ISLs in the fabric, without the need for a large number of end hosts or applications. For information on the SM Loop Test and how to use the test, refer to the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide*.



17.0 Upgrade the Management Node

This section provides step-by-step directions to upgrade an Fabric Management Node from a previous Intel® Omni-Path Fabric Suite (IFS) software version to the latest IFS software version.

17.1 Before You Upgrade

Prior to upgrading the IFS software, ensure the following items have been completed:

- Review the Release Notes for a list of compatible software.
- Uninstall all versions of third-party IB stacks.
- Back up the following configuration files in case the upgrade fails:
 - /etc/sysconfig/opafm.xml
 - /etc/sysconfig/opa/*
 - /var/usr/lib/opa/analysis/baseline/*
 - Refer to the OS documentation for a list of any other OS-specific files that should be included in any backups.

17.2 Intel® Omni-Path Fabric Suite Upgrade

Upgrading the IFS software on a Fabric Management Node with an existing IFS software is accomplished using the `IntelOPA-IFS.DISTRO.VERSION.tgz` package file.

1. Download and extract the installation packages following the [Download and Extract Installation Packages](#) section.
2. Change directory to `IntelOPA-IFS.DISTRO.VERSION` directory.

```
cd IntelOPA-IFS.DISTRO.VERSION
```
3. Start the TUI installation.

```
./INSTALL
```

The **Intel OPA VERSION Software** main menu is displayed.

```
Intel OPA VERSION Software

1) Install/Uninstall Software
2) Reconfigure OFA IP over IB
3) Reconfigure Driver Autostart
4) Generate Supporting Information for Problem Report
6) FastFabric (Host/Chassis/Switch Setup/Admin)

X) Exit
```

4. Select 1) Install/Uninstall Software.



Screen 1 of 2 of the **Intel OPA Install Menu** is displayed.

```
Intel OPA Install (VERSION release) Menu

Please Select Install Action (screen 1 of 2):
0) OFA OPA Stack          [ Upgrade ] [Available] VERSION
1) OFA IBACM              [ Upgrade ] [Available] VERSION
2) Intel HFI Components [ Upgrade ] [Available] VERSION
3) OPA Tools              [ Upgrade ] [Available] VERSION
4) OFA OPA Development [ Upgrade ] [Available] VERSION
5) FastFabric            [ Upgrade ] [Available] VERSION
6) OFA IP over IB        [ Upgrade ] [Available] VERSION
7) OPA FM                [ Upgrade ] [Available] VERSION
8) MVAPICH2 (hfi,gcc)    [ Upgrade ] [Available] VERSION
9) MVAPICH2 (hfi,Intel) [ Upgrade ] [Available] VERSION
a) OpenMPI (hfi,gcc)     [ Upgrade ] [Available] VERSION
b) OpenMPI (hfi,Intel)  [ Upgrade ] [Available] VERSION
c) GASNet (hfi,gcc)     [ Upgrade ] [Available] VERSION
d) OpenSHMEM (hfi,gcc)  [ Upgrade ] [Available] VERSION

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

5. Review the items to be upgraded; the default value is in brackets (Upgrade or Up To Date). To change a value, type the alphanumeric character associated with the item.
6. Press **N** to go to the next screen.

Screen 2 of 2 of the **Intel OPA Install Menu** is displayed.

```
Intel OPA Install (VERSION release) Menu

Please Select Install Action (screen 2 of 2):
0) MVAPICH2(verbs,gcc) [Don't Install] [Available] VERSION
1) OpenMPI (verbs,gcc) [Don't Install] [Available] VERSION
2) MPI Source          [ Upgrade ] [Available] VERSION
3) Pre-Boot Components [Don't Install] [Available] VERSION
4) OFA Debug Info      [Don't Install] [Available] VERSION

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

7. Review the items to be upgraded; the default value is in brackets (Upgrade, Up To Date, or Don't Install).
- To change a value, type the alphanumeric character associated with the item.
8. Press **P** to perform the selected actions from the two screens.
9. For each system prompt, select the default by pressing **Enter**.

The **Intel OPA Autostart Menu** is displayed.

```
Intel OPA Autostart (VERSION release) Menu

Please Select Autostart Option:
0) OFA OPA Stack () [Enable ]
1) OFA IBACM (ibacm) [Disable]
2) Intel HFI Components () [Enable ]
3) OFA IP over IB () [Enable ]
4) OPA FM (opafm) [Disable]
```



```
P) Perform the autostart changes
S) Autostart All                      R) Autostart None
X) Return to Previous Menu (or ESC)
```

10. Review the items for Autostart; the default value is in brackets (Enable or Disable). To change a value, type the alphanumeric character associated with the item.

Intel recommends leaving all of the Autostart selections as default.

11. Press **P** to perform the selected actions from the screen.

The system prompts:

```
Hit any key to continue...
```

12. Press any key to proceed with the installation.

The system prompts:

```
A System Reboot is recommended to activate the software changes
Hit any key to continue...
```

13. Press any key to complete the installation.

The installation completes and returns to the main menu.

14. Press **X** to exit.

15. Reboot the server.

```
reboot
```




18.0 Upgrade the Fabric Software

This section describes how to upgrade the Intel® Omni-Path Fabric software to a new version.

Note: On each Fabric Management Node, perform an upgrade installation of the IFS software using the procedure documented in [Upgrade the Management Node](#) on page 102.

1. **(All)** Perform steps 1 and 2 in the [Install Host Software on the Remaining Hosts Using the FastFabric TUI](#) on page 54.
2. Select the following items to be performed in the menu:
 - (Optional) The 1) Edit Config and Select/Edit Hosts Files option permits you to edit the `opafastfabric.conf` and `hosts` files.
 - The 5) Install/Upgrade OPA Software option upgrades the Intel® Omni-Path Fabric software on all the selected hosts. By default, it looks in the current directory for the `FF_PRODUCT.$FF_PRODUCT_VERSION.tgz` file. If it is not found in the current directory, you are prompted to supply a directory name where this file can be found.

Note: The upgrade installation updates any existing OFA Delta software, and is only valid for hosts that already have a previous version of OFA Delta software installed.

3. Press **P** to perform the upgrade.

If	Then
1) Edit Config and Select/Edit Hosts Files was selected	Continue with the next step.
1) Edit Config and Select/Edit Hosts Files was not selected	Skip to Step 6 .

4. Press any key for the following system prompt:

```
About to: vi /etc/sysconfig/opa/opafastfabric.conf
Hit any key to continue (or ESC to abort)...
```

The `opafastfabric.conf` file is opened in the editor. Review all the settings. See *Intel® Omni-Path Fabric Suite FastFabric User Guide* Configuration Files section for more information about the `opafastfabric.conf` file.

5. When finished updating the `opafastfabric.conf` file, quit out of the editor.

The system prompts:

```
You should select a file which OMITS this host
Select Host File to Use/Edit [/etc/sysconfig/opa/hosts]:
```

6. Select a `hosts` file that lists all the hosts except the Fabric Management nodes. If necessary, create a new file at this time, potentially based on the existing `/etc/sysconfig/opa/hosts` file.



Note: Do not list any of the Fabric Management Nodes in the host file (for example, the nodes that have FastFabric installed).

Note: The file may list the Management Network or IPoIB hostnames for the selected hosts.

The `host` file is opened in the editor, review all the information. See *Intel® Omni-Path Fabric Suite FastFabric User Guide* Configuration Files section for more information about the `host` file.

7. When finished updating the `hosts` file, quit out of the editor.

The system prompts:

```
Do you want to edit/review/change the files? [y]:
```

8. Press **N** to continue.

The system prompts:

```
Performing Host Setup: Install/Upgrade OPA Software  
Hit any key to continue...
```

9. Press any key to continue.

The system displays prompts that require your response as it is going through the installation.

10. Select the default by pressing **Enter**.

The **Intel OPA Autostart Menu** displays.

```
Intel OPA Autostart (VERSION release) Menu  
  
Please Select Autostart Option:  
0) OPA OPA Stack () [Enable ]  
1) OPA IBACM (ibacm) [Disable]  
2) Intel HFI Components () [Enable ]  
3) OPA IP over IB () [Enable ]  
4) OPA FM (opafm) [Disable]  
  
P) Perform the autostart changes  
S) Autostart All R) Autostart None  
X) Return to Previous Menu (or ESC)
```

11. Press **P** to proceed.

When the selected hosts have completed rebooting, the **FastFabric OPA Host Setup Menu** is displayed. The upgrade is complete.

If any hosts fail to be updated, use the `View opahostadmin result files` option to review the result files from the update. Refer to the *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide* for more details.

Note: When using the Intel® Omni-Path Fabric software, you can use FastFabric to upgrade the Intel® Omni-Path Fabric software stack on the remaining hosts.

12. **(Linux)** If any other setup operations need to be performed on all hosts, use the `Run a command on all hosts` menu option. This option executes the specified Linux* shell command against all selected hosts. It can also execute a sequence of commands separated by semicolons.

Note: Check the relevant release notes for the new Intel® Omni-Path Fabric Host Software release being installed for any additional required steps.



13. **(Linux)** The option `Reboot Hosts` reboots all the selected hosts and ensures they go down and come back up properly, as verified through ping over the management network. When the hosts come back up, they are running the installed Intel® Omni-Path Fabric Host Software.
14. Perform the verification steps in one of the following procedures:
 - [Verify Intel® Omni-Path Fabric Host Software on the Remaining Servers Using the FastFabric TUI](#)
 - [Verify Intel® Omni-Path Fabric Host Software on the Remaining Servers Using CLI Commands](#)



19.0 Upgrade from IntelOPA-Basic to IntelOPA-IFS

This section describes how to upgrade the Intel® Omni-Path Software (Intel® OP Software) from IntelOPA-Basic to IntelOPA-IFS using the Install TUI menus. The IFS package installs FastFabric and OPA_FM software.

This installation procedure presumes that you are logged in and have completed all pre-installation requirements.

1. Download and extract the `IntelOPA-IFS.DISTRO.VERSION.tgz` package file as described in [Download and Extract Installation Packages](#).
2. Change to the `IntelOPA-IFS.DISTRO.VERSION` directory with the command:

```
cd IntelOPA-IFS.DISTRO.VERSION
```

3. Start the Install TUI:

```
./INSTALL
```

The **Intel OPA VERSION Software** main menu displays.

```
Intel OPA VERSION Software

1) Install/Uninstall Software
2) Reconfigure OFA IP over IB
3) Reconfigure Driver Autostart
4) Generate Supporting Information for Problem Report
6) FastFabric (Host/Chassis/Switch Setup/Admin)

X) Exit
```

4. Select 1) Install/Uninstall Software.

Screen 1 of 2 of the **Intel OPA Install Menu** displays.

```
Intel OPA Install (VERSION release) Menu

Please Select Install Action (screen 1 of 2):
0) OFA OPA Stack      [ Up To Date ] [Available] VERSION
1) OFA IBACM          [ Up To Date ] [Available] VERSION
2) Intel HFI Components [ Up To Date ] [Available] VERSION
3) OPA Tools          [ Up To Date ] [Available] VERSION
4) OFA OPA Development [ Up To Date ] [Available] VERSION
5) FastFabric         [ Install   ] [Available] VERSION
6) OFA IP over IB      [ Up To Date ] [Available] VERSION
7) OPA FM             [ Install   ] [Available] VERSION
8) MVAPICH2 (hfi,gcc) [ Up To Date ] [Available] VERSION
9) MVAPICH2 (hfi,Intel) [ Up To Date ] [Available] VERSION
a) OpenMPI (hfi,gcc)   [ Up To Date ] [Available] VERSION
b) OpenMPI (hfi,Intel) [ Up To Date ] [Available] VERSION
c) GASNet (hfi,gcc)   [ Up To Date ] [Available] VERSION
d) OpenSHMEM (hfi,gcc) [ Up To Date ] [Available] VERSION

N) Next Screen
```



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

5. Review the items to ensure that FastFabric and OPA_FM show Install and all other selections show Up To Date.
6. Press **P** to perform the FastFabric and OPA_FM installation.
7. Follow the system prompts. For each prompt, select the default by pressing **Enter**.

The **Intel OPA Autostart Menu** displays.

```
Intel OPA Autostart (VERSION release) Menu

Please Select Autostart Option:
0) OFA OPA Stack ()                [Enable ]
1) OFA IBACM (ibacm)               [Disable]
2) Intel HFI Components ()         [Enable ]
3) OFA IP over IB ()               [Enable ]
4) OPA FM (opafm)                  [Disable]

P) Perform the autostart changes
S) Autostart All                   R) Autostart None
X) Return to Previous Menu (or ESC)
```

8. Review the items for Autostart. The default value is in brackets (Enable or Disable). To change a value, enter the alphanumeric character associated with the item.

Intel recommends leaving all of the Autostart selections set to the default values.

9. Press **P** to perform the selected actions from the screen.

The system prompts:

```
Hit any key to continue...
```

10. Press any key to proceed with the installation.

The system prompts:

```
A System Reboot is recommended to activate the software changes
Hit any key to continue...
```

11. Press any key to proceed with the installation.

The system prompts:

```
Rebuilding boot image with "/usr/bin/dracut -f"...done.
```

The system returns to a system prompt.

12. Press **X** to exit.
13. Reboot the server.



Appendix A HFI UEFI PXE Installation and Configuration

The following information describes a mechanism for configuring a server to boot over Intel® Omni-Path using the Intel® Omni-Path Host Fabric Interface (HFI) Unified Extensible Firmware Interface (UEFI) firmware.

The following table shows the BIOS versus Boot Mode information.

BIOS	Boot Mode	UEFI Driver Is Run By BIOS?	PCIe 8GT/s (Gen3) Transition	Platform Configuration Data for AOC support	Boot Options	Board Description	Subsystem/SubVendor ID
UEFI	UEFI BIOS	Yes	Achieved pre boot	Provided by UEFI driver	Boot over fabric Boot local Boot from Ethernet	Customized by UEFI driver	Customized by UEFI driver
UEFI	Legacy BIOS	Yes (BIOS that runs UEFI driver in this mode)	Achieved pre boot	Provided by driver post boot (with limitations)	Boot local Boot from Ethernet	Default board description	Customized by UEFI driver
UEFI	Legacy BIOS	No (BIOS does not run UEFI driver in this mode – e.g. Intel UEFI BIOS)	Achieved by host driver post boot	Provided by driver post boot (with limitations)	Boot local Boot from Ethernet	Default board description	Default subsystem values
Legacy	Legacy	No	Achieved by host driver post boot	Provided by driver post boot (with limitations)	Boot local Boot from Ethernet	Default board description	Default subsystem values

The following procedures are required to configure and boot a server over Intel® Omni-Path using the HFI UEFI firmware. There are also examples of recommended PXE (Preboot eXecution Environment) configurations, however, it does not cover all possible PXE configurations. The examples given are targeted for servers running Red Hat® Enterprise Linux® version 7 (RHEL® 7). For detailed descriptions of other PXE server configurations, please refer to your OS vendor documentation.

A.1 Client Installation

The HFI UEFI firmware is available as part of the hfi1-uefi.x86_64 RPM, released as part of the Intel® Omni-Path IFS package.



To install or upgrade the HFI UEFI firmware, follow the example shown below.

```
# rpm -i hfil-uefi-X.X-XX.x86_64.rpm
# cd /opt/opa/bios_images/
# ll
total 820
-rw-r--r-- 1 root root 531680 May 10 18:03 HfiPcieGen3_0xVV.efi
-rw-r--r-- 1 root root 31744 May 10 18:03 HfiPcieGen3Loader_0xVV.rom
-rw-r--r-- 1 root root 19530 May 10 18:03 License_UEFI_Option_ROM
-rw-r--r-- 1 root root 252298 May 10 18:03 License_UEFI_Option_ROM.pdf
# modprobe hfil
# hfil_eprom -w -o HfiPcieGen3Loader_0xVV.rom -b HfiPcieGen3_0x1a.efi
Device ID: 0x000015ef, 32 Mbits
Erasing oprom partition
Writing oprom partition
Erasing bulk partition
Writing bulk partition
# hfil_eprom -V -o
Device ID: 0x000015ef, 32 Mbits
Reading oprom partition
oprom partition version: 0x0000001a
# hfil_eprom -V -b
Device ID: 0x000015ef, 32 Mbits
Reading bulk partition
bulk partition version: 0x0000001a
# reboot
```

A.2 Server Configuration

A.2.1 Intel® Omni-Path Software Installation

The IFS package contains the HFI driver and the required IP over IB software stack. It also optionally installs a Fabric Manager (FM). The FM should be installed for the Intel® Omni-Path Boot client to be fully initialized.

In order to boot over Intel® Omni-Path, the following requirements must be met:

1. The Intel® Omni-Path Fabric Manager software must be running.
Use the command: `service opafm start`
2. The HFI driver (hfil) and IP over Infiniband modules must be running. Use the commands:
`modprobe hfil`
`modprobe ib_ipoib`

A.2.2 DHCP Server

Configure DHCP server for Dynamic IP addressing

Follow these steps to install a DHCP server.

1. Install DHCP package using the command: `yum install dhcp`



2. Add DHCP configuration file. The following is an example: /etc/dhcp/dhcpd.conf

```
option space pxelinux;
option pxelinux.magic code 208 = string;
option pxelinux.configfile code 209 = text;
option pxelinux.pathprefix code 210 = text;
option pxelinux.reboottime code 211 = unsigned integer 32;
option architecture-type code 93 = unsigned integer 16;

allow booting;
allow bootp;

subnet 192.168.100.0 netmask 255.255.255.0 {
    range 192.168.100.2 192.168.100.230;
    option subnet-mask 255.255.255.0;
    option broadcast-address 192.168.100.255;
    option routers 192.168.100.1;

    class "pxeclients" {
        match if substring (option vendor-class-identifier, 0, 9) = "PXEClient";
        next-server 192.168.100.1;

        if option architecture-type = 00:07 {
            filename "uefi/bootx64.efi";
        } else {
            filename "pxelinux/pxelinux.0";
        }
    }
}
```

3. Start DHCP server using the command: `service dhcp start`

Configure DHCP server for fixed IP addressing

1. Using the GUID from the client HFI in this example 001175010165abb2
2. 2) Next configure the dhcpd.conf file as the following
 - a. a. Affix the standard ff:00:00:00:00:02:00:00:01:57:

```
[root@phs1fnivd14u25 dhcp]# cat /etc/dhcp/dhcpd.conf
option space pxelinux;
option pxelinux.magic code 208 = string;
option pxelinux.configfile code 209 = text;
option pxelinux.pathprefix code 210 = text;
option pxelinux.reboottime code 211 = unsigned integer 32;
option architecture-type code 93 = unsigned integer 16;

allow booting;
allow bootp;

subnet 192.168.100.0 netmask 255.255.255.0 {
    range 192.168.100.20 192.168.100.230;
    option subnet-mask 255.255.255.0;
    option broadcast-address 192.168.100.255;
    option routers 192.168.100.1;

    host phs1fnivd14u23 {
        #dynamic;
        uid ff:00:00:00:00:02:00:00:01:57:00:11:75:01:01:65:ab:b2;
        fixed-address 192.168.100.15;
        supersede dhcp-client-identifier =
        ff:00:00:00:00:02:00:00:01:57:00:11:75:01:01:65:ab:b2;
    }

    class "pxeclients" {
        match if substring (option vendor-class-identifier, 0, 9)
```




```
=
"PXEClient";
    next-server 192.168.100.1;
    if option architecture-type = 00:07 {
        filename "uefi/bootx64.efi";
    } else {
        filename "pxelinux/pxelinux.0";
    }
}
}
```

3. Restart the DHCP
4. The output from `/var/log/messages` the dhcp server will look something like the following example:

```
Jun 30 03:16:13 phs1fnivd14u25 systemd: Started DHCPv4 Server Daemon.
Jun 30 03:16:26 phs1fnivd14u25 dhcpd: DHCPDISCOVER from
ff:00:00:00:00:02:00:00:01:57:00:11:75:01:01:65:ab:b2 via ib0
Jun 30 03:16:26 phs1fnivd14u25 dhcpd: DHCPOFFER on 192.168.100.15 to
ff:00:00:00:00:02:00:00:01:57:00:11:75:01:01:65:ab:b2 via ib0
Jun 30 03:16:29 phs1fnivd14u25 dhcpd: DHCPREQUEST for 192.168.100.15
(192.168.100.1) from ff:00:00:00:00:02:00:00:01:57:00:11:75:01:01:65:ab:b2
via ib0
Jun 30 03:16:29 phs1fnivd14u25 dhcpd: DHCPACK on 192.168.100.15 to
ff:00:00:00:00:02:00:00:01:57:00:11:75:01:01:65:ab:b2 via ib0
```

A.2.3 Network Interface

Before IP packets can be sent to and from the PXE server, an Intel® Omni-Path network interface must be configured. To configure an Intel® Omni-Path network interface, perform the following steps:

1. From the IFS install menu, select (Re)configure OFA IP over IB
2. Fill in the requested information based on the DHCP server configured in [DHCP Server](#) on page 111.
3. Once the installation has completed, run the following command, where `ib0` is the name of the network interface configured in step 2 above:

```
ifup ib0
```

A.2.4 TFTP Server

Follow these steps to install a TFTP server.

1. Install TFTP package using the command: `yum install tftp-server`
2. Install XINETD package using the command: `yum install xinetd`
3. Add TFTP configuration file. An example is shown below: `/etc/xinetd.d/tftp`

```
service tftp
{
    socket_type          = dgram
    protocol             = udp
    wait                 = yes
    user                 = root
    server                = /usr/sbin/in.tftpd
    server_args           = -s /var/lib/tftpboot
    disable               = no
    per_source            = 11
```



```
cps                = 100 2
flags              = IPv4
}
```

Note: The `disable` field is by default `yes` and should be changed to `no` before starting the XINETD server.

4. Start XINETD server using the command: `service xinetd start`

A.2.5 PXE Server

The PXE Server is contained within a directory structure indicated by the `server_args` argument of the TFTP configuration file.

1. Create base directory structure using the command: `mkdir -p /var/lib/tftpboot/uefi`

A.2.5.1 Boot Loader

This section describes configuring a PXE server to use the Elilo boot loader. For configuration details of other boot loaders, please refer to your OS vendor documentation.

1. Download the latest source files for Elilo from <https://sourceforge.net/projects/elilo/files/elilo/>.
2. Extract copy the following files to the directory structure created above:
`elilo-3.16-x86_64.efi` renamed to `-> bootx64.efi`
`elilo-3.16-source/examples/textmenu_chooser/textmenu-message.msg`
3. Add kernel and initramfs files to the directory structure created above.
4. Add Elilo configuration file. An example is shown below using `vmlinuz` and `initramfs.img`: `/var/lib/tftpboot/uefi/elilo.conf`

```
chooser=textmenu
message=textmenu-message.msg
prompt
delay=20

default=rhel7
image=vmlinuz
label=rhel7
description = "RHEL 7"
initrd=initramfs.img
read-only
```



Appendix B UEFI Secure Boot Support

Intel® Omni-Path software includes support for UEFI Secure Boot, both in the UEFI firmware environment and in Linux.

B.1 UEFI Firmware Support

UEFI Firmware Support

All UEFI drivers shipped as part of the Intel® Omni-Path Host Software are signed to support UEFI Secure Boot using the Microsoft UEFI CA signing process.

Limitations

PXE Booting over the Intel® Omni-Path fabric is not supported when UEFI Secure Boot mode is enabled.

B.2 Linux Support

Kernel modules shipped with major Linux distributions RHEL 7.3 and SLES 12.2 are signed using the keys for those distributions.

Kernel modules shipped as part of the Intel® Omni-Path IFS installation suite are signed using an Intel key. The Intel public key is automatically added to the Machine Owner Key (MOK) list in the UEFI firmware environment with no action required by the user. See [HFI UEFI PXE Installation and Configuration](#) for UEFI firmware requirements.

The following MOK utility can be used to list the currently enrolled keys in the MOK list keyring:

```
mokutil --list-enrolled
```

For more information about UEFI Secure Boot support in RHEL, see the RHEL documentation at https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/7/html/System_Administrators_Guide/sec-UEFI_Secure_Boot.html.

For more information about UEFI Secure Boot support in SLES, see the SLES documentation at https://www.suse.com/documentation/sled11/book_sle_admin/data/sec_uefi_secboot.html.

Limitations

When UEFI Secure Boot mode is enabled, the following Intel® Omni-Path Linux utilities are not available:

- hfil_eprom
- hfidiags



Appendix C Software Installation Checklist

This section includes checklists to help you track tasks during fabric setup, installation and verification. Check off each step as it is performed. Refer to [Install the Intel® Omni-Path Software](#) for a more detailed explanation of each step.

C.1 Pre-Installation Checklist

Step	Description	Complete
1.	Ensure that hardware is installed, cabled, and powered. Refer to the <i>Intel® Omni-Path Fabric Switches Hardware Installation Guide</i> .	
2.	Ensure that an HFI is installed in each server. Refer to the <i>Intel® Omni-Path Host Fabric Interface Installation Guide</i> .	
3.	The hardware configuration should be reviewed to ensure everything was installed and installed properly, according to the plan. Refer to the local hardware configuration plan.	
4.	Ensure that the required operating system is installed on each server with the following options: <ul style="list-style-type: none">Root user command prompt ends in "#" or "\$". <i>Note:</i> There must be a space after "#" or "\$".Fancy and colored prompts must be disabled.TCL and Expect packages installed on all Fabric Management Nodes. Refer to the <i>Intel® Omni-Path Fabric Software Release Notes</i> for supported operating systems.	
5.	Ensure the capability of remote login as root enabled. <ul style="list-style-type: none">SSH server enabledAll servers configured with the same root password	
6.	Ensure that there is a TCP/IP Host Name Resolution. <ul style="list-style-type: none">If using /etc/hosts , update the /etc/hosts file on the Fabric Management Node.If using DNS, all Management Network and IPoIB hostnames added to DNS./etc/resolv.conf file configured on Fabric Management Node.	
7.	Ensure that an NTP server is set up.	
8.	Use the installation checklist for your type of installation: <ul style="list-style-type: none">Install the Intel® Omni-Path Fabric Software on a Cluster	

C.2 Install the Intel® Omni-Path Fabric Software on a Cluster

Step	Description	Complete
1.	Complete the steps in the Pre-Installation Checklist .	
2.	Download and Extract Installation Packages .	
3.	Install the Intel® Omni-Path Software .	
4.	Configure Intel® Omni-Path Chassis .	
5.	Install Intel® Omni-Path Fabric Suite Fabric Manager GUI .	
6.	Configure Firmware on the Externally Managed Intel® Omni-Path Switches .	
continued...		



Step	Description	Complete
7.	Install Host Software on the Remaining Hosts Using the FastFabric TUI.	
8.	Verify Intel® Omni-Path Fabric Host Software on the Remaining Servers Using the FastFabric TUI.	
9.	Install Additional Management Nodes.	
10.	Configure and Initialize Health Check Tools.	
11.	Run High Performance Linpack 2 Benchmark.	