

# **Intel® Omni-Path Fabric Software**

## **Installation Guide**

---

***Rev. 8.1***

***November 2017***



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.

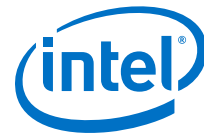
Copies of documents which have an order number and are referenced in this document may be obtained by calling 1-800-548-4725 or visit <http://www.intel.com/design/literature.htm>.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at [intel.com](http://intel.com).

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 © 2015–2017, Intel Corporation. All rights reserved.



## Revision History

For the latest documentation, go to: [www.intel.com/omnipath/FabricSoftwarePublications](http://www.intel.com/omnipath/FabricSoftwarePublications).

Date	Revision	Description
November 2017	8.1	Document has been updated with the following: <ul style="list-style-type: none"> <li>Made corrections in section <a href="#">OS RPMs Installation Prerequisites</a></li> <li>Made corrections in Appendix, <a href="#">Setting Up PXE Boot for Supported Operating Systems</a></li> </ul>
October 2017	8.0	Document has been updated with the following: <ul style="list-style-type: none"> <li>Updated and added information to the <a href="#">Fabric Design Prerequisites</a> section</li> <li>Added "Sandia-OpenSHMEM (hfi, gcc)" to the <a href="#">OPA-Basic Package</a> section</li> <li>Updated section <a href="#">OS RPMs Installation Prerequisites</a> to add RHEL* 7.4, SLES* 12.3 and an example of a install command</li> <li>Added <code>sandiashmem</code> to the "Supported Component (comp) Names" list in <a href="#">Using the ./INSTALL Command Line Options</a> section</li> <li>Updated <a href="#">Upgrade Prerequisites</a> to include link to <a href="#">OS RPMs Installation Prerequisites</a></li> <li>Removed outdated appendix, "HFI UEFI PXE Installation and Configuration"</li> <li>Added new Appendix, <a href="#">Setting Up PXE Boot for Supported Operating Systems</a></li> <li>Moved <a href="#">Upgrade the HFI UEFI Firmware</a> under <a href="#">Setting Up PXE Boot for Supported Operating Systems</a> and updated the section</li> <li>The Intel® Omni-Path Fabric Suite <i>FastFabricCommand Line Interface Reference Guide</i> has been merged into the <i>Intel® Omni-Path Fabric Host Software User Guide</i> and the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i>. In this document, all references have been updated appropriately. See the <a href="#">Intel® Omni-Path Documentation Library</a> on page 11 for details.</li> </ul>
August 2017	7.0	Document has been updated with the following: <ul style="list-style-type: none"> <li>Globally, updated the following filepaths: <ul style="list-style-type: none"> <li>from <code>/etc/sysconfig/opa</code> to <code>/etc/opa</code></li> </ul> </li> <li>Added high level installation flowchart to section <a href="#">Installation Overview</a> and updated the information</li> <li>Updated section <a href="#">Fabric Manager GUI Packages</a></li> <li>Updated section <a href="#">OS RPMs Installation Prerequisites</a></li> <li>Added section <a href="#">Download and Install NVIDIA* Software (Optional)</a></li> <li>Updated section <a href="#">Install the Intel® Omni-Path Software</a></li> <li>Updated section <a href="#">Install Using the Linux* Distribution Software</a></li> <li>Updated section <a href="#">Install Intel® Omni-Path Fabric Suite Fabric Manager GUI</a></li> </ul>
April 2017	6.0	Document has been updated with the following: <ul style="list-style-type: none"> <li>Document has been restructured and rewritten for usability.</li> <li>Globally, updated the following filepaths: <ul style="list-style-type: none"> <li>from <code>/usr/lib/opa-fm/samples/</code> to <code>/usr/share/opa-fm/samples/</code></li> <li>from <code>/usr/lib/opa-fm/etc/opafm*</code> to <code>/usr/lib/opa-fm/bin/opafm*</code></li> </ul> </li> </ul>
continued...		



Date	Revision	Description
		<ul style="list-style-type: none"> <li>— from /etc/sysconfig/opafm.xml to /etc/opa-fm/opafm.xml</li> <li>— from /etc/sysconfig/* to /etc/*</li> <li>— from /usr/lib/opa/ to /usr/share/opa/</li> <li>— from /usr/lib/opa/src/* to /usr/src/opa/*</li> <li>• Added <a href="#">Intel® Omni-Path Documentation Library</a> to Preface.</li> <li>• Added new section <a href="#">Intel® Omni-Path Architecture Overview</a>.</li> <li>• Added new section <a href="#">Install Using the Linux* Distribution Software</a>.</li> <li>• Added new section <a href="#">Performance Tuning Prerequisites</a>.</li> <li>• Deprecated opaxlattopology_cust in <a href="#">Fabric Setup Prerequisites</a>.</li> <li>• Added new section <a href="#">Update the Fabric Manager opafm Configuration File</a>.</li> <li>• Added new section <a href="#">Thermal Monitoring Module Verification</a>.</li> <li>• Updated Installation Checklist to include configuration and verification pointers.</li> </ul>
Dec 2016	5.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> <li>• Removed all information pertaining to RHEL* 7.1</li> <li>• Updated "Prepare Fabric Manager for Fabric Manager GUI"</li> <li>• Added new section "Multi-Rail Usage" to Chapter "Additional Installation and Setup Tasks"</li> <li>• Moved Appendix "Multi-Subnet Fabrics" under new Chapter "Additional Installation and Setup Tasks"</li> <li>• Updated "HFI UEFI PXE Installation and Configuration" section</li> <li>• Removed Appendix "hfdiags User Guide". Moved to <i>Intel® Omni-Path Fabric Host Software User Guide</i>.</li> </ul>
Aug 2016	4.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> <li>• Added information on UEFI firmware Prerequisites</li> <li>• Updated Install TUI menus</li> <li>• Updated ./INSTALL command information to include --user-space</li> <li>• Updated Fabric Manager GUI Installation information</li> <li>• Updated "HFI UEFI PXE Installation and Configuration" section</li> <li>• Updated "hfdiags User Guide" section</li> </ul>
May 2016	3.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> <li>• Updated OS RPMs Installation Prerequisites</li> <li>• Added MPI Compiler Prerequisites</li> <li>• Added information about minimum disk space requirement for the Fabric Manager GUI installation</li> <li>• Replaced "Client/Server Configuration to Boot Over Fabric" section with "HFI UEFI PXE Installation and Configuration"</li> </ul>
February 2016	2.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> <li>• Updated OS RPMs Installation Prerequisites</li> <li>• Added "Client/Server Configuration to Boot Over Fabric" appendix</li> </ul>
November 2015	1.0	Initial Release



## Contents

---

<b>Revision History.....</b>	<b>3</b>
<b>Preface.....</b>	<b>11</b>
Intended Audience.....	11
Intel® Omni-Path Documentation Library.....	11
Cluster Configurator for Intel® Omni-Path Fabric.....	14
Documentation Conventions.....	14
Laser Safety Information.....	15
Electrostatic Discharge Sensitivity (ESDS) Precautions.....	15
License Agreements.....	15
Technical Support.....	15
<b>1.0 Introduction.....</b>	<b>16</b>
1.1 Document Organization.....	16
<b>2.0 Overview.....</b>	<b>18</b>
2.1 Intel® Omni-Path Architecture Overview.....	18
2.1.1 Host Fabric Interface.....	20
2.1.2 Intel® OPA Switches.....	20
2.1.3 Intel® OPA Management.....	21
2.2 Intel® Omni-Path Software Overview.....	21
2.3 Installation Overview.....	23
2.4 Installation Packages.....	24
2.4.1 OPA-Basic Package.....	24
2.4.2 OPA-IFS Package.....	24
2.4.3 Fabric Manager GUI Packages.....	25
<b>3.0 Checklists.....</b>	<b>26</b>
3.1 OPA-Basic Checklists.....	26
3.1.1 OPA-Basic Pre-Installation Checklist.....	26
3.1.2 Install the OPA-Basic Software Checklist.....	26
3.1.3 Upgrade the OPA-Basic Software Checklist.....	27
3.2 OPA-IFS Checklists.....	27
3.2.1 OPA-IFS Pre-Installation Checklist.....	27
3.2.2 Install and Configure the OPA-IFS Software Checklist.....	27
3.2.3 Upgrade the OPA-IFS Software Checklist.....	28
<b>Part 1: Installing the Software.....</b>	<b>29</b>
<b>4.0 Installation Getting Started.....</b>	<b>30</b>
4.1 Pre-Installation Requirements.....	30
4.1.1 Fabric Design Prerequisites.....	30
4.1.2 Fabric Setup Prerequisites.....	31
4.1.3 OS RPMs Installation Prerequisites.....	33
4.1.4 MPI Compiler Prerequisites.....	38
4.1.5 Performance Tuning Prerequisites.....	38
4.2 Download the Intel® Omni-Path Software.....	39
4.3 Unpack the Tar File.....	39



4.4 Download and Install NVIDIA* Software (Optional).....	40
<b>5.0 Install the Intel® Omni-Path Software.....</b>	<b>41</b>
5.1 Before You Begin.....	41
5.1.1 Using the ./INSTALL Command Line Options.....	41
5.2 Install Using the TUI Menus.....	45
5.3 Install Using CLI Commands.....	46
5.4 Install Using the Linux* Distribution Software.....	47
5.5 Configure IPoIB IPV6.....	49
<b>Part 2: Configuring the Software.....</b>	<b>51</b>
<b>6.0 Configuration Getting Started.....</b>	<b>52</b>
6.1 Pre-Configuration Requirements.....	52
6.1.1 Chassis Configuration Prerequisites.....	52
6.1.2 Switch Configuration Prerequisites.....	53
6.2 How to Use the FastFabric TUI.....	53
<b>7.0 Configure the Chassis.....</b>	<b>56</b>
7.1 Before You Begin.....	56
7.2 Configure the Chassis Using the FastFabric TUI Menu.....	57
7.3 Configure the Chassis Using the FastFabric CLI Commands.....	61
<b>8.0 Configure the Externally-Managed Switches.....</b>	<b>66</b>
8.1 Before You Begin.....	66
8.2 Configure the Externally-Managed Switches Using FastFabric TUI Menu.....	67
8.3 Configure the Externally-Managed Switches Using FastFabric CLI Commands.....	69
<b>9.0 Install Host Software on Remaining Servers.....</b>	<b>71</b>
9.1 Before You Begin.....	71
9.2 Install the Host Software on the Remaining Hosts Using the FastFabric TUI Menu.....	73
9.3 Verify the Host Software on the Remaining Servers Using the FastFabric TUI Menu.....	76
9.4 Install the Host Software on the Remaining Hosts Using the FastFabric CLI Commands..	79
9.5 Verify the Host Software on the Remaining Servers Using CLI Commands.....	81
<b>10.0 Set Up Additional Management Nodes.....</b>	<b>85</b>
10.1 Before You Begin.....	85
10.2 Configure Additional Management Nodes Using FastFabric TUI Menus.....	85
10.3 Configure Additional Management Nodes Using FastFabric CLI Commands.....	87
<b>11.0 Perform Initial Health Check.....</b>	<b>88</b>
11.1 Before You Begin.....	88
11.2 Configure and Initialize Health Check Tools Using FastFabric TUI Menu.....	89
11.3 Configure and Initialize Health Check Tools Using FastFabric CLI Commands.....	90
<b>12.0 Perform High Performance Linpack Benchmark .....</b>	<b>92</b>
12.1 Before You Begin.....	93
12.2 Run a Sample High Performance Linpack 2 Configuration.....	93
12.3 Generate and Run a Custom High Performance Linpack Configuration.....	94
<b>13.0 Install Intel® Omni-Path Fabric Suite Fabric Manager GUI.....</b>	<b>96</b>
13.1 Before You Begin.....	96
13.1.1 System Requirements for a Windows* Environment.....	96



13.1.2 System Requirements for a Linux* Environment.....	96
13.1.3 Download the Fabric Manager Software.....	97
13.2 Install the Fabric Manager GUI Software on Windows*.....	97
13.3 Install the Fabric Manager GUI Software on Linux*.....	99
13.4 Prepare Fabric Manager for Fabric Manager GUI.....	100
13.5 Start the Fabric Manager GUI on Windows*.....	101
13.6 Start the Fabric Manager GUI on Linux*.....	102
13.7 Uninstall Fabric Manager GUI.....	102
13.7.1 Uninstall Fabric Manager GUI Software on Windows*.....	103
13.7.1.1 Clear the Fabric Manager GUI Cache on Windows*.....	103
13.7.2 Uninstall Fabric Manager GUI Software on Linux*.....	104
13.7.2.1 Clear the Fabric Manager GUI Cache on Linux*.....	104
<b>14.0 Additional Installation and Setup Tasks.....</b>	<b>106</b>
14.1 Multi-Subnet Fabrics.....	106
14.1.1 Setting Up Primarily Independent Subnets.....	106
14.1.2 Setting Up Overlapping Subnets.....	108
14.2 Multi-Rail Usage.....	110
14.2.1 Multi-Rail Overview.....	110
14.2.2 Setting Up Dual Rails for a Single Subnet.....	112
14.2.3 Configuring Dual Rails for Dual Subnets.....	114
<b>15.0 Installation Verification and Additional Settings.....</b>	<b>117</b>
15.1 LED Link and Data Indicators.....	117
15.2 Thermal Monitoring Module Verification.....	117
15.3 Adapter and Other Settings.....	118
15.4 ARP Neighbor Table Setup for Large Clusters.....	118
15.5 SM Loop Test.....	119
<b>Part 3: Upgrading the Software.....</b>	<b>120</b>
<b>16.0 Upgrade Getting Started.....</b>	<b>121</b>
16.1 Upgrade Prerequisites.....	121
16.2 Download the Intel® Omni-Path Software.....	121
16.3 Unpack the Tar File.....	122
<b>17.0 Upgrade the Intel® Omni-Path Software.....</b>	<b>123</b>
17.1 Before You Begin.....	123
17.2 Upgrade Scenarios.....	123
17.3 Upgrade the OPA-Basic Software.....	123
17.4 Upgrade the OPA-IFS Software.....	124
17.4.1 Upgrade the Software on the Management Nodes.....	124
17.4.2 Update the Fabric Manager opafm Configuration File.....	125
17.4.3 Upgrade the Software on the Remaining Servers.....	126
17.5 Upgrade from OPA-Basic to OPA-IFS Software Package.....	128
<b>Appendix A UEFI Secure Boot Support.....</b>	<b>129</b>
A.1 UEFI Firmware Support.....	129
A.2 Linux* Support.....	129
<b>Appendix B Setting Up PXE Boot for Supported Operating Systems.....</b>	<b>130</b>
B.1 PXE Server.....	131



B.1.1 Configure PXE Server with RHEL* 7.4.....	131
B.1.1.1 Before You Begin.....	131
B.1.1.2 Configure Network Interface.....	131
B.1.1.3 Configure DHCP Server.....	132
B.1.1.4 Install a TFTP Server.....	134
B.1.1.5 Create Directory Structure for PXE Server .....	135
B.1.1.6 Configure the HTTP Server.....	135
B.1.1.7 Configure the Boot Loader.....	136
B.1.2 Configure PXE Server with SLES* 12 SP3.....	136
B.1.2.1 Before You Begin.....	136
B.1.2.2 Configure Network Interface.....	137
B.1.2.3 Configure DHCP Server.....	137
B.1.2.4 Install a TFTP Server.....	140
B.1.2.5 Create Directory Structure for PXE Server .....	140
B.1.2.6 Configure the HTTP Server.....	141
B.1.2.7 Configure the Boot Loader.....	141
B.2 Client Node.....	142
B.2.1 Before You Begin.....	143
B.2.2 Upgrade the HFI UEFI Firmware.....	143
B.2.3 Install RHEL* 7.4 over Intel® OP HFI.....	145
B.2.4 Install SLES* 12 SP3 over Intel® OP HFI.....	146





## Figures

1	Intel® OPA Fabric.....	19
2	Intel® OPA Building Blocks.....	20
3	Intel® OPA Fabric and Software Components.....	22



## **Tables**

1	Red Hat* Enterprise Linux* OS Distribution RPMs Matrix.....	34
2	SUSE* Linux* Enterprise Server OS Distribution RPMs Matrix.....	36
3	BIOS versus Boot Mode.....	142



## 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.

## Intel® Omni-Path Documentation Library

Intel® Omni-Path publications are available at the following URLs:

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

Use the tasks listed in this table to find the corresponding Intel® Omni-Path document.

Task	Document Title	Description
<b>Key:</b> Shading indicates the URL to use for accessing the particular document.		
• Intel® Omni-Path Switches Installation, User, and Reference Guides:	<a href="http://www.intel.com/omnipath/SwitchPublications">http://www.intel.com/omnipath/SwitchPublications</a>	
• Intel® Omni-Path Software Installation, User, and Reference Guides (includes HFI documents):	<a href="http://www.intel.com/omnipath/FabricSoftwarePublications">http://www.intel.com/omnipath/FabricSoftwarePublications</a> (no shading)	
• Drivers and Software (including Release Notes):	<a href="http://www.intel.com/omnipath/Downloads">http://www.intel.com/omnipath/Downloads</a>	
<i>continued...</i>		



Task	Document Title	Description
Setting up an Intel® OPA cluster	<i>Intel® Omni-Path Fabric Setup Guide</i> (Old title: <i>Intel® Omni-Path Fabric Staging Guide</i> )	Provides a high level overview of the steps required to stage a customer-based installation of the Intel® Omni-Path Fabric. Procedures and key reference documents, such as Intel® Omni-Path user guides and installation guides are provided to clarify the process. Additional commands and BKM's are defined to facilitate the installation process and troubleshooting.
Installing hardware	<i>Intel® Omni-Path Fabric Switches Hardware Installation Guide</i>	Describes the hardware installation and initial configuration tasks for the Intel® Omni-Path Switches 100 Series. This includes: Intel® Omni-Path Edge Switches 100 Series, 24 and 48-port configurable Edge switches, and Intel® Omni-Path Director Class Switches 100 Series.
	<i>Intel® Omni-Path Host Fabric Interface Installation Guide</i>	Contains instructions for installing the HFI in an Intel® OPA cluster. A cluster is defined as a collection of nodes, each attached to a fabric through the Intel interconnect. The Intel® HFI utilizes Intel® Omni-Path switches and cabling.
Installing host software Installing HFI firmware Installing switch firmware (externally-managed switches)	<i>Intel® Omni-Path Fabric Software Installation Guide</i>	Describes using a Text-based User Interface (TUI) to guide you through the installation process. You have the option of using command line interface (CLI) commands to perform the installation or install using the Linux* distribution software.
Managing a switch using Chassis Viewer GUI Installing switch firmware (managed switches)	<i>Intel® Omni-Path Fabric Switches GUI User Guide</i>	Describes the Intel® Omni-Path Fabric Chassis Viewer graphical user interface (GUI). It provides task-oriented procedures for configuring and managing the Intel® Omni-Path Switch family. Help: GUI online help.
Managing a switch using the CLI Installing switch firmware (managed switches)	<i>Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide</i>	Describes the command line interface (CLI) task information for the Intel® Omni-Path Switch family. Help: -help for each CLI.
Managing a fabric using FastFabric	<i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> (Merged with: <i>Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide</i> )	Provides instructions for using the set of fabric management tools designed to simplify and optimize common fabric management tasks. The management tools consist of TUI menus and command line interface (CLI) commands. Help: -help and man pages for each CLI. Also, all host CLI commands can be accessed as console help in the Fabric Manager GUI.
Managing a fabric using Fabric Manager	<i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i>	The Fabric Manager uses a well defined management protocol to communicate with management agents in every Intel® Omni-Path Host Fabric Interface (HFI) and switch. Through these interfaces the Fabric Manager is able to discover, configure, and monitor the fabric.
	<i>Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide</i>	Provides an intuitive, scalable dashboard and set of analysis tools for graphically monitoring fabric status and configuration. It is a user-friendly alternative to traditional command-line tools for day-to-day monitoring of fabric health. Help: Fabric Manager GUI Online Help.
Configuring and administering Intel® HFI and IPoIB driver	<i>Intel® Omni-Path Fabric Host Software User Guide</i>	Describes how to set up and administer the Host Fabric Interface (HFI) after the software has been installed. The audience for this document includes both cluster
continued...		



Task	Document Title	Description
Running MPI applications on Intel® OPA		administrators and Message-Passing Interface (MPI) application programmers, who have different but overlapping interests in the details of the technology.
Writing and running middleware that uses Intel® OPA	<i>Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide</i>	Provides a reference for programmers working with the Intel® PSM2 Application Programming Interface (API). The Performance Scaled Messaging 2 API (PSM2 API) is a low-level user-level communications interface.
Optimizing system performance	<i>Intel® Omni-Path Fabric Performance Tuning User Guide</i>	Describes BIOS settings and parameters that have been shown to ensure best performance, or make performance more consistent, on Intel® Omni-Path Architecture. If you are interested in benchmarking the performance of your system, these tips may help you obtain better performance.
Designing an IP or storage router on Intel® OPA	<i>Intel® Omni-Path IP and Storage Router Design Guide</i>	Describes how to install, configure, and administer an IPoIB router solution (Linux* IP or LNet) for inter-operating between Intel® Omni-Path and a legacy InfiniBand* fabric.
Building a Lustre* Server using Intel® OPA	<i>Building Lustre* Servers with Intel® Omni-Path Architecture Application Note</i>	Describes the steps to build and test a Lustre* system (MGS, MDT, MDS, OSS, OST, client) from the HPDD master branch on a x86_64, RHEL*/CentOS* 7.1 machine.
Building Containers for Intel® OPA fabrics	<i>Building Containers for Intel® Omni-Path Fabrics using Docker* and Singularity* Application Note</i>	Provides basic information for building and running Docker* and Singularity* containers on Linux*-based computer platforms that incorporate Intel® Omni-Path networking technology.
Writing management applications that interface with Intel® OPA	<i>Intel® Omni-Path Management API Programmer's Guide</i>	Contains a reference for programmers working with the Intel® Omni-Path Architecture Management (Intel OPAMGT) Application Programming Interface (API). The Intel OPAMGT API is a C-API permitting in-band and out-of-band queries of the FM's Subnet Administrator and Performance Administrator.
Learning about new release features, open issues, and resolved issues for a particular release	<i>Intel® Omni-Path Fabric Software Release Notes</i>	
	<i>Intel® Omni-Path Fabric Manager GUI Release Notes</i>	
	<i>Intel® Omni-Path Fabric Switches Release Notes</i> (includes managed and externally-managed switches)	

## How to Search the Intel® Omni-Path Documentation Set

Many PDF readers, such as Adobe\* Reader and Foxit\* Reader, allow you to search across multiple PDFs in a folder.

Follow these steps:

1. Download and unzip all the Intel® Omni-Path PDFs into a single folder.
2. Open your PDF reader and use **CTRL-SHIFT-F** to open the Advanced Search window.
3. Select **All PDF documents in...**
4. Select **Browse for Location** in the dropdown menu and navigate to the folder containing the PDFs.
5. Enter the string you are looking for and click **Search**.

Use advanced features to further refine your search criteria. Refer to your PDF reader Help for details.



## 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 15 for more information.  
For more information, visit [www.intel.com](http://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.



## Laser Safety Information

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

A rectangular label with a thick black border. The text inside is centered and reads:

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

## Electrostatic Discharge Sensitivity (ESDS) Precautions

The assemblies used in the switch chassis are ESD sensitive. Observe ESD handling procedures when handling any assembly used in the switch chassis.

## 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 <http://www.intel.com/omnipath/support> for additional detail.



## 1.0 Introduction

---

This guide provides instructions for installing the Intel® Omni-Path Architecture (Intel® OPA) software and configuring the system for the first time. It also provides instructions for upgrading the software.

For details about the other documents for the Intel® Omni-Path product line, refer to [Intel® Omni-Path Documentation Library](#) in this document.

You install the software using one of the following methods:

- Guided installation using Text User Interface (TUI) menus (recommended)
- Command Line Interface (CLI) commands
- Linux\* Distribution Software

**Note:** The default Intel® OPA configuration during installation is for a single-subnet fabric.

To set up other fabric configurations, please contact your Intel Support Personnel. For additional information, refer to [Multi-Subnet Fabrics](#) on page 106.

## 1.1 Document Organization

This manual is organized as follows:

- Introduction
- Overview
- Checklists

Part 1: Installing the Software

- Installation Getting Started
- Install Intel® Omni-Path Software

Part 2: Configuring the Software

- Configuration Getting Started
- Configure the Chassis
- Configure the Externally Managed Switches
- Install Host Software on Remaining Servers
- Set Up Additional Management Nodes
- Perform Initial Health Check
- Perform High Performance Linpack Benchmark
- Install Intel® Omni-Path Fabric Suite Fabric Manager GUI
- Additional Installation and Setup Tasks
- Installation Verification and Additional Settings





Part 3: Upgrading the Software

- Upgrade Getting Started
- Upgrading the Intel® Omni-Path Software

Appendices

- UEFI Secure Boot Support
- HFI UEFI PXE Installation and Configuration



## 2.0 Overview

---

This section provides an overview of the Intel® Omni-Path Architecture and software installation process.

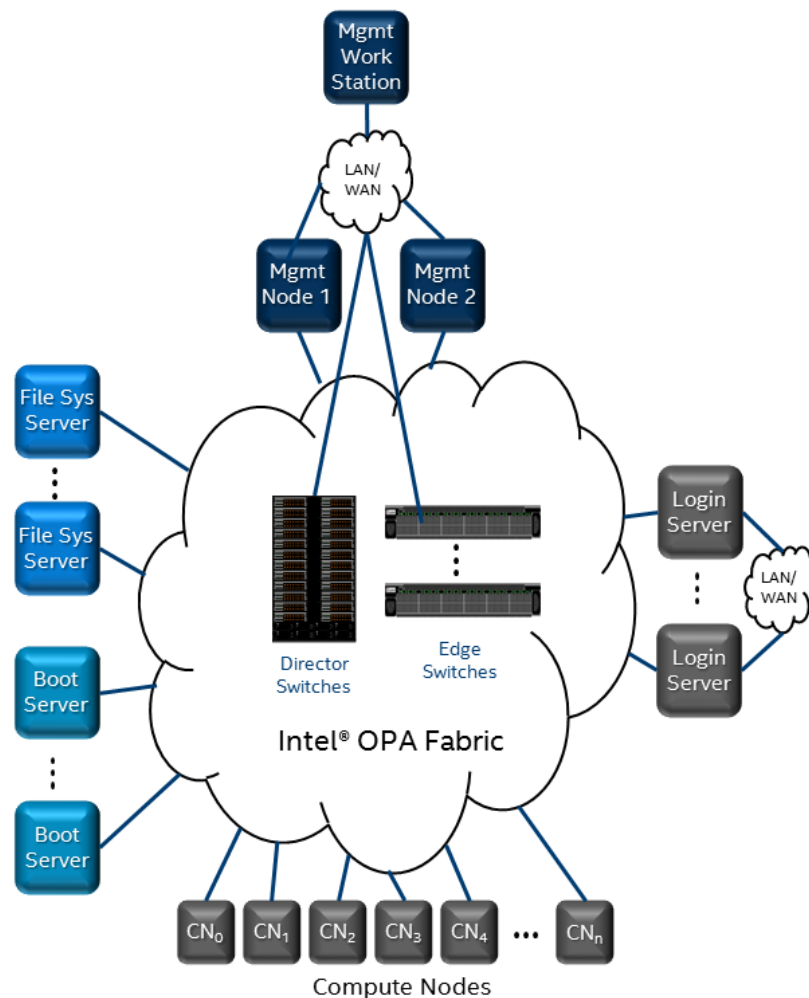
### 2.1 Intel® Omni-Path Architecture Overview

The Intel® Omni-Path Architecture (Intel® OPA) interconnect fabric design enables a broad class of multiple node computational applications requiring scalable, tightly-coupled processing, memory, and storage resources. Options for close "on-package" integration between Intel® OPA family devices, Intel® Xeon® Processors, and Intel® Xeon Phi™ Processors, enable significant system level packaging and network efficiency improvements. When coupled with open standard APIs developed by the OpenFabrics Alliance\* (OFA) Open Fabrics Interface (OFI) workgroup, host fabric interfaces (HFIs) and switches in the Intel® OPA family systems are optimized to provide the low latency, high bandwidth, and high message rate needed by large scale High Performance Computing (HPC) applications.

Intel® OPA provides innovations for a multi-generation, scalable fabric, including link layer reliability, extended fabric addressing, and optimizations for many-core processors. High performance datacenter needs are also a core Intel® OPA focus, including link level traffic flow optimization to minimize datacenter-wide jitter for high priority packets, robust partitioning support, quality of service support, and a centralized fabric management system.

The following figure shows a sample Intel® OPA-based fabric, consisting of different types of nodes and servers.

Figure 1. Intel® OPA Fabric



To enable the largest scale systems in both HPC and the datacenter, fabric reliability is enhanced by combining the link level retry typically found in HPC fabrics with the conventional end-to-end retry used in traditional networks. Layer 2 network addressing is extended for systems with over ten million endpoints, thereby enabling use on the largest scale datacenters for years to come.

To enable support for a breadth of topologies, Intel® OPA provides mechanisms for packets to change virtual lanes as they progress through the fabric. In addition, higher priority packets are able to preempt lower priority packets to provide more predictable system performance, especially when multiple applications are running simultaneously. Finally, fabric partitioning is provided to isolate traffic between jobs or between users.

The software ecosystem is built around OPA software and includes four key APIs.

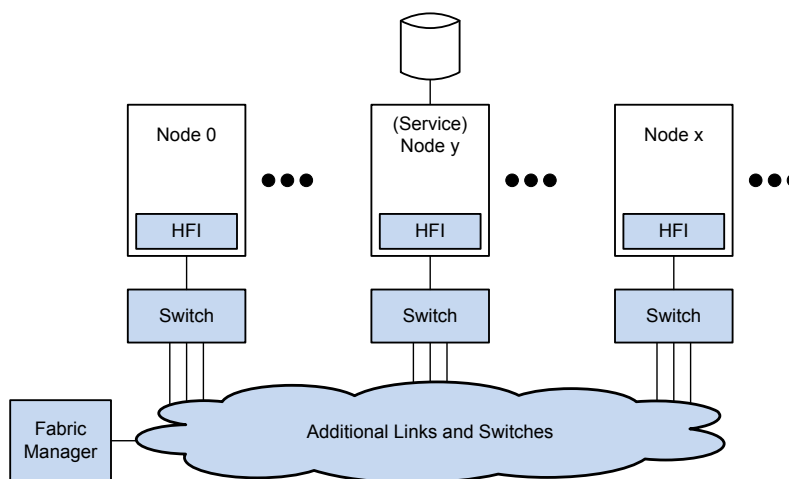
1. The OPA OFI represents a long term direction for high performance user level and kernel level network APIs.

2. The Performance Scaled Messaging 2 (PSM2) API provides HPC-focused transports and an evolutionary software path from the Intel® True Scale Fabric.
3. OFA Verbs provides support for existing remote direct memory access (RDMA) applications and includes extensions to support Intel® OPA fabric management.
4. Sockets is supported via OFA IPoFabric (also called IPoIB) and rSockets interfaces. This permits many existing applications to immediately run on Intel® Omni-Path as well as provide TCP/IP features such as IP routing and network bonding.

Higher level communication libraries, such as the Message Passing Interface (MPI), and Partitioned Global Address Space (PGAS) libraries, are layered on top of these low level OFA APIs. This permits existing HPC applications to immediately take advantage of advanced Intel® Omni-Path features.

Intel® Omni-Path Architecture 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. These building blocks are shown in the following figure.

**Figure 2. Intel® OPA Building Blocks**



### 2.1.1 Host Fabric Interface

Each host is connected to the fabric through a Host Fabric Interface (HFI) adapter. The HFI translates instructions between the host processor and the fabric. The HFI includes the logic necessary to implement the physical and link layers of the fabric architecture, so that a node can attach to a fabric and send and receive packets to other servers or devices. HFIs also include specialized logic for executing and accelerating upper layer protocols.

### 2.1.2 Intel® OPA Switches

Intel® OPA switches are OSI Layer 2 (link layer) devices, and act as packet forwarding mechanisms within a single Intel® OPA fabric. Intel® OPA switches are responsible for implementing Quality of Service (QoS) features, such as virtual lanes, congestion management, and adaptive routing. Switches are centrally managed by the Intel® Omni-Path Fabric Suite Fabric Manager software, and each switch includes a



management agent to handle management transactions. Central management means that switch configurations are programmed by the FM software, including managing the forwarding tables to implement specific fabric topologies, configuring the QoS and security parameters, and providing alternate routes for adaptive routing. As such, all OPA switches must include management agents to communicate with the Intel® OPA Fabric Manager.

### 2.1.3 Intel® OPA Management

The Intel® OPA fabric is centrally managed and supports redundant Fabric Managers that manage every device (server and switch) in the fabric through management agents associated with those devices. The Primary Fabric Manager is an Intel® OPA fabric software component selected during the fabric initialization process.

The Primary Fabric Manager is responsible for:

1. Discovering the fabric's topology.
2. Setting up Fabric addressing and other necessary values needed for operating the fabric.
3. Creating and populating the Switch forwarding tables.
4. Maintaining the Fabric Management Database.
5. Monitoring fabric utilization, performance, and statistics rates.

The fabric is managed by sending management packets over the fabric. These packets are sent *in-band* (that is, over the same wires as regular network packets) using dedicated buffers on a specific virtual lane (VL15). End-to-end reliability protocols are used to detect lost packets.

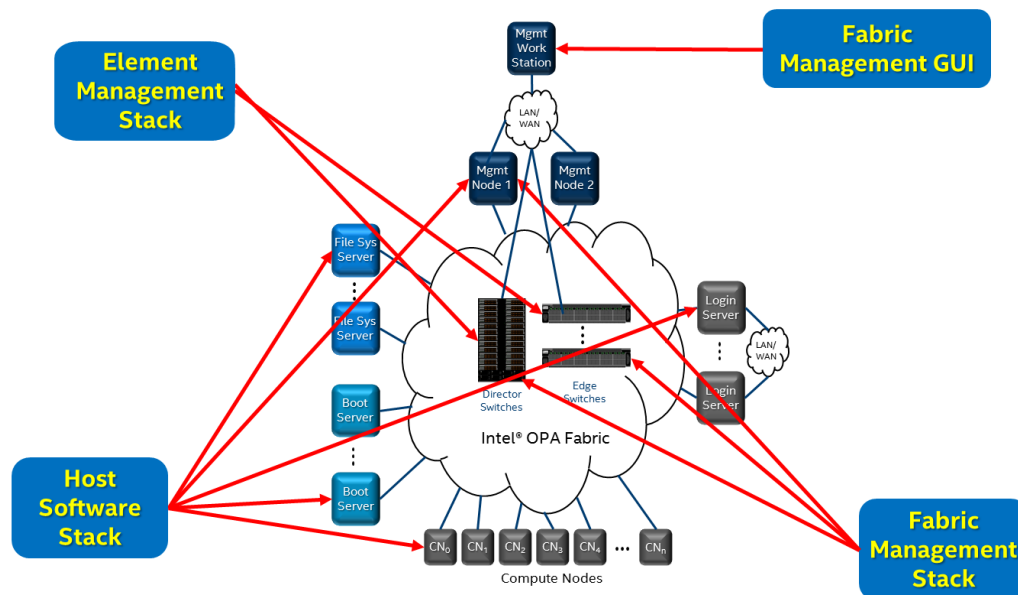
## 2.2 Intel® Omni-Path Software Overview

For software applications, Intel® OPA maintains consistency and compatibility with existing Intel® True Scale Fabric and InfiniBand\* APIs utilizing the open source OpenFabrics Alliance\* (OFA) software stack on Linux\* distribution releases.

### Software Components

The key software components and their usage models are shown in the following figure and described in the following paragraphs.

Figure 3. Intel® OPA Fabric and Software Components



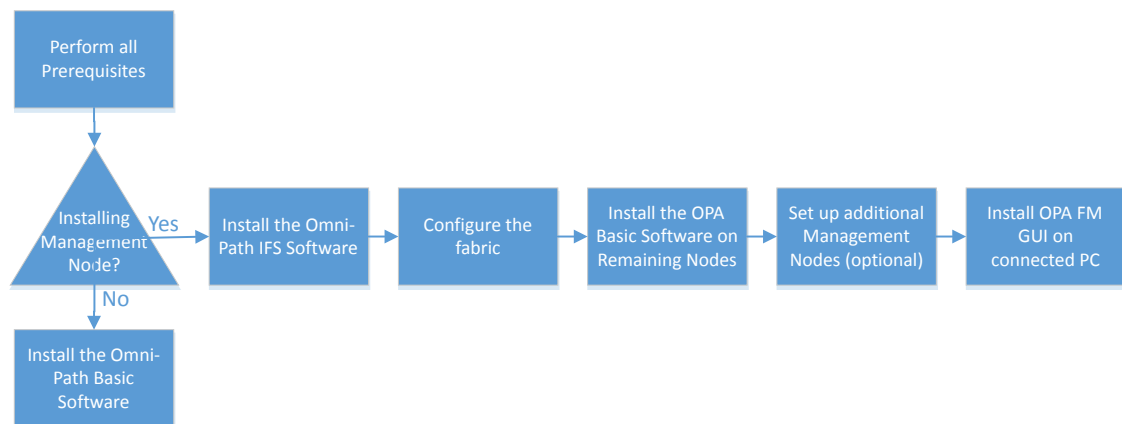
Software Component Descriptions	
<b>Element Management Stack</b>	
<ul style="list-style-type: none"> <li>Runs on an embedded Intel processor included in managed Intel® OP Edge Switch 100 Series and Intel® Omni-Path Director Class Switch 100 Series switches.</li> <li>Provides system management capabilities, including signal integrity, thermal monitoring, and voltage monitoring, among others.</li> <li>Accessed via Ethernet* port using command line interface (CLI) or graphical user interface (GUI).</li> </ul>	
User documents:	
<ul style="list-style-type: none"> <li>Intel® Omni-Path Fabric Switches GUI User Guide</li> <li>Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide</li> </ul>	
<b>Host Software Stack</b>	
<ul style="list-style-type: none"> <li>Runs on all Intel® OPA-connected host nodes and supports compute, management, and I/O nodes.</li> <li>Provides a rich set of APIs including OFI, PSM2, sockets, and OFA verbs.</li> <li>Provides high performance, highly scalable MPI implementation via OFA, PSM2, and an extensive set of upper layer protocols.</li> <li>Includes Boot over Fabric mechanism for configuring a server to boot over Intel® Omni-Path using the Intel® OP HFI Unified Extensible Firmware Interface (UEFI) firmware.</li> </ul>	
User documents:	
<ul style="list-style-type: none"> <li>Intel® Omni-Path Fabric Host Software User Guide</li> </ul>	
<i>continued...</i>	



Software Component Descriptions
<ul style="list-style-type: none"> <li>Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide</li> </ul>
<b>Fabric Management Stack</b> <ul style="list-style-type: none"> <li>Runs on Intel® OPA-connected management nodes or embedded Intel processor on the switch.</li> <li>Initializes, configures, and monitors the fabric routing, QoS, security, and performance.</li> <li>Includes a toolkit for configuration, monitoring, diagnostics, and repair.</li> </ul> User documents: <ul style="list-style-type: none"> <li>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</li> <li>Intel® Omni-Path Fabric Suite FastFabric User Guide</li> </ul>
<b>Fabric Management GUI</b> <ul style="list-style-type: none"> <li>Runs on laptop or workstation with a local screen and keyboard.</li> <li>Provides interactive GUI access to Fabric Management features such as configuration, monitoring, diagnostics, and element management drill down.</li> </ul> User documents: <ul style="list-style-type: none"> <li>Intel® Omni-Path Fabric Suite Fabric Manager GUI Online Help</li> <li>Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide</li> </ul>

## 2.3 Installation Overview

The standard software installation process takes you through installing the Intel® Omni-Path Fabric (OPA-Basic) or Intel® Omni-Path Fabric Suite (OPA-IFS) software package, configuring the system, and verifying the system settings.



Intel recommends that you install the OPA-IFS software on the Management Node using the Install TUI, and then use 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 either the FastFabric TUI or a provisioning or diskless boot mechanism.

**Note:** If you are using a provisioning system, consult the documentation that comes with the provisioning system.



## 2.4 Installation Packages

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

### 2.4.1 OPA-Basic Package

The `IntelOPABasic.DISTRO.VERSION.tgz` installation package (where `DISTRO.VERSION` is the OS distribution and software version) 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.

The installation package 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)
- Sandia-OpenSHMEM (hfi, gcc)
- MVAPICH2 (verbs, gcc)
- OpenMPI (verbs, gcc)
- MPI Source
- Pre-Boot components
- OFA Debug Info

**Note:** A separate Intel® Omni-Path Fabric Host Software installation package is available 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.

### 2.4.2 OPA-IFS Package

The `IntelOPA-IFS.DISTRO.VERSION.tgz` installation package (where `DISTRO.VERSION` is the OS distribution and software version) provides the Intel® Omni-Path Fabric Host Software package along with special features including the Intel® Omni-Path Fabric Suite FastFabric toolset and the Intel® Omni-Path Fabric Suite Fabric Manager.





The installation package includes the following components:

- OPA-Basic package  
For the list of components, refer to [OPA-Basic Package](#).
- Intel® Omni-Path Fabric Suite FastFabric  
Refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide* for more details.
- Intel® Omni-Path Fabric Suite Fabric Manager  
Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for more details.

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

### 2.4.3 Fabric Manager GUI Packages

The following packages are used to install the Intel® Omni-Path Fabric Suite Fabric Manager GUI

- Linux\* installation package:  
`IntelOPA-FMGUI-linux-x.x.x.x.x.noarch.rpm`
- Windows\* installation package:  
`IntelOPA-FMGUI.windows-x.x.x.x.x.exe`

The Intel® Omni-Path Fabric Suite Fabric Manager GUI provides a set of features for viewing and monitoring the fabric or multiple fabrics. The application is installed on a computer outside of the fabric.



## 3.0 Checklists

This section includes checklists to help you track tasks during fabric installation, configuration, and upgrade.

### 3.1 OPA-Basic Checklists

#### 3.1.1 OPA-Basic 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 is 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"><li>Root user command prompt ends in "#" or "\$". <i>Note:</i> A space must appear after "#" or "\$".</li><li>Fancy and colored prompts must be disabled.</li><li>TCL and Expect packages installed on all Fabric Management Nodes.</li></ul> 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"><li>SSH server enabled</li><li>All servers configured with the same root password</li></ul>	
6.	Ensure that there is a TCP/IP Host Name Resolution. <ul style="list-style-type: none"><li>If using <code>/etc/hosts</code>, update the <code>/etc/hosts</code> file on the Fabric Management Node.</li><li>If using DNS, all Management Network and IPoIB host names are added to DNS.</li><li>The <code>/etc/resolv.conf</code> file is configured on Fabric Management Node.</li></ul>	
7.	Ensure that an NTP server is set up.	
8.	Ensure HFI Node Description Names are assigned.	

#### 3.1.2 Install the OPA-Basic Software Checklist

Step	Description	Complete
1.	Complete the steps in the <a href="#">OPA-Basic Pre-Installation Checklist</a> .	
2.	Download and Extract the OPA-Basic Software Package per <a href="#">Installation Getting Started</a> .	
3.	<a href="#">Install the Intel® Omni-Path Software</a> .	



### 3.1.3 Upgrade the OPA-Basic Software Checklist

Step	Description	Complete
1.	Complete the steps in the <a href="#">Upgrade Prerequisites</a> on page 121.	
2.	Download and Unpack the new OPA-Basic Software Package per <a href="#">Download the Intel® Omni-Path Software</a> on page 39 and <a href="#">Unpack the Tar File</a> on page 39 respectfully.	
3.	Upgrade the OPA-Basic Software on each compute node per <a href="#">Upgrade the OPA-Basic Software</a> .	

## 3.2 OPA-IFS Checklists

### 3.2.1 OPA-IFS 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.	Ensure that all switch ports with management allowed are connected to the Management node.	
4.	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.	
5.	Ensure that the required operating system is installed on each server with the following options: <ul style="list-style-type: none"> <li>Root user command prompt ends in "#" or "\$". <i>Note:</i> A space must appear after "#" or "\$".</li> <li>Fancy and colored prompts must be disabled.</li> <li>TCL and Expect packages installed on all Fabric Management Nodes.</li> </ul> Refer to the <i>Intel® Omni-Path Fabric Software Release Notes</i> for supported operating systems.	
6.	Ensure the capability of remote login as root enabled. <ul style="list-style-type: none"> <li>SSH server enabled</li> <li>All servers configured with the same root password</li> </ul>	
7.	Ensure that there is a TCP/IP Host Name Resolution. <ul style="list-style-type: none"> <li>If using <code>/etc/hosts</code>, update the <code>/etc/hosts</code> file on the Fabric Management Node.</li> <li>If using DNS, all Management Network and IPoIB host names are added to DNS.</li> <li>The <code>/etc/resolv.conf</code> file is configured on Fabric Management Node.</li> </ul>	
8.	Ensure that an NTP server is set up.	
9.	Ensure HFI Node Description Names are assigned.	

### 3.2.2 Install and Configure the OPA-IFS Software Checklist

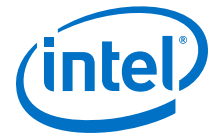
Step	Description	Complete
1.	Complete the steps in the <a href="#">OPA-IFS Pre-Installation Checklist</a> .	
2.	Download and extract the OPA-IFS Software Package per <a href="#">Installation Getting Started</a> .	
3.	<a href="#">Install the Intel® Omni-Path Software</a> .	
4.	<a href="#">Configure the Chassis</a> .	
5.	<a href="#">Configure the Externally-Managed Switches</a> .	
<b>continued...</b>		



Step	Description	Complete
6.	Install the Host Software on the Remaining Hosts Using the FastFabric TUI Menu.	
7.	Set Up Additional Management Nodes (optional).	
8.	Install Intel® Omni-Path Fabric Suite Fabric Manager GUI (optional).	
9.	Verify the Host Software on the Remaining Servers Using the FastFabric TUI Menu.	
10.	Configure and Initialize Health Check Tools Using FastFabric CLI Commands.	
11.	Run a Sample High Performance Linpack 2 Configuration.	

### 3.2.3 Upgrade the OPA-IFS Software Checklist

Step	Description	Complete
1.	Complete the steps in the <a href="#">OPA-IFS Pre-Installation Checklist</a> .	
2.	Download and extract the new OPA-IFS Software Packages per <a href="#">Upgrade Getting Started</a> .	
3.	<a href="#">Upgrade the Software on the Management Nodes</a> .	
4.	<a href="#">Upgrade the Software on the Remaining Servers</a> .	
5.	Verify the Host Software on the Remaining Servers Using the FastFabric TUI Menu.	



# **Part 1: Installing the Software**



## 4.0 Installation Getting Started

---

This section provides instructions and information for getting started with the Intel® Omni-Path Software installation.

### 4.1 Pre-Installation Requirements

This section provides the information and procedures needed prior to installing the fabric software. Typically, the Site Implementation Engineer performs the setup tasks described in this section to ensure that the fabric is ready for the software installation.

#### 4.1.1 Fabric Design Prerequisites

Ensure that the following requirements are met prior to installing the software.

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 2.5 GB of physical memory for each Fabric Manager instance.  
When managing a cluster of 10000 nodes or more, 5 GB of memory per Fabric Manager instance is required.  
For very large fabrics, with more than 16000 nodes, 15 GB 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. Minimum swap disk space should be at least 5 GB + 34 GB (for PM short-term history)
  - All servers with HFIs installed should have Unified Extensible Firmware Interface (UEFI) firmware installed.  
Refer to the [Client Node](#) on page 142 for UEFI BIOS versus Boot Mode information.  
*Note:* Intel® Omni-Path Architecture Option ROM support requires UEFI firmware; and, the HFI drivers are based on UEFI instead of Legacy-BIOS.
  - Intel recommends, but does not require the following:
    - Using Intel® Xeon® Processor dual-socket server.



- Processor SKUs that favor single threaded performance over maximum cores (i.e. higher base frequency).
- 32GB or more of ECC memory (ideally fully populating DIMM channels on processor to insure maximum memory bandwidth)
- 256GB or more of storage (ideally RAIDed for resiliency)
- Plan the cabling of the fabric and create a cable planning spreadsheet using the sample .cvs files installed into `/usr/share/opa/samples/` on the management node. After a plan is established, `opaxlattopology` or `opaxlattformology_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 **opaxlattopology\_cust** sections of the *Intel® Omni-Path Fabric Suite FastFabric User 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.

### 4.1.2 Fabric Setup Prerequisites

Ensure that the following requirements are met prior to installing and setting up the fabric.

*Note:* For information about the configuration files used by FastFabric, refer to *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

1. 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` (deprecated), `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. 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. 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 "\$" with a space after it.
- Fancy and colored prompts are disabled.
- TCL and Expect packages are 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 each Management Node.

*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. 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 you are using the `/etc/hosts` file approach and not using Domain Name System (DNS):

- 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
```

If you are 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 the Host Software on the Remaining Hosts Using the FastFabric TUI Menu](#) procedures to copy the `/etc/resolv.conf` file from the Management Node to all the other nodes.
7. 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 the Chassis](#) on page 56.
8. 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 the OPA-IFS package.

### 4.1.3 OS RPMs Installation Prerequisites

Ensure that the following requirements are met prior to installing the software.

In addition to normal OS installation options, there are OS RPMs that must be installed before installing the Intel® Omni-Path software. Verify that all required OS RPMs are installed using one of the following tables:

- [Red Hat\\* Enterprise Linux\\* OS Distribution RPMs Matrix](#)



- [SUSE\\* Linux\\* Enterprise Server OS Distribution RPMs Matrix](#)

*Note:* Use the version distributed with your OS distribution.

**Table 1. Red Hat\* Enterprise Linux\* OS Distribution RPMs Matrix**

OS Distribution RPMs		RHEL* 6.7	RHEL* 7.3	RHEL* 7.4
<b>IB/OPA Centric</b>				
	libibmad		x	x
	libibumad		x	x
	libibumad-devel		x	
	libibverbs	x	x	x
	librdmacm	x	x	x
	libibcm	x	x	x
	ibacm		x	x
	qperf	x	x	x
	perftest	x	x	x
	rdma	x	x	
	infinipath-psm	x	x	x
	opensm-devel	x		
	opensm-libs	x	x	x
	rdma-core			x
	rdma-core-devel			x
<b>OPA Centric</b>				
	libhfi1		x	
<b>System Centric</b>				
	expat	x	x	x
	elfutils-libelf-devel	x	x	x
	libstdc++-devel	x	x	x
	gcc-gfortran	x	x	x
	atlas	x	x	x
	tcl	x	x	x
	expect	x	x	x
	tcsh	x	x	x
	sysfsutils	x	x	x
	pciutils	x	x	x
	bc (command line calculator for floating point math)	x	x	x
	rpm-build	x	x	x
	redhat-rpm-config	x	x	x
<b>continued...</b>				



OS Distribution RPMs		RHEL* 6.7	RHEL* 7.3	RHEL* 7.4
	kernel-devel	x	x	x
	c-ares	x		
	json-c	x	x	x
	json-c-devel	x		
	coreutils		x	x
	elfutils-libelf		x	x
	glibc		x	x
	libgfortran		x	x
	libgomp		x	x
	libquadmath		x	x
	libstdc++		x	x
	ncurses-libs		x	x
	numactl-libs		x	x
	openssl-libs		x	x
	zlib		x	x
	papi		x	x
<b>Other/Generally Installed</b>				
	chkconfig			
	bash		x	x
	systemd		x	x
	perl		x	x
	perl-Getopt-Long		x	x
	perl-Socket		x	x
	irqbalance		x	x
	kernel		x	x
	kmod		x	x
	libgcc		x	x
	pkgconfig		x	x
	python		x	x
<b>Build Requirements</b>				
	libibverbs-devel		x	x
	libibmad-devel		x	x
	librdmacm-devel		x	x
	ibacm-devel		x	x
	openssl-devel (1.0.1 or higher)		x	x
	libuuid-devel		x	x
<b>continued...</b>				



OS Distribution RPMs		RHEL* 6.7	RHEL* 7.3	RHEL* 7.4
	expat-devel		x	x
	infinipath-psm-devel		x	x
	valgrind-devel		x	x
	json-c-devel	x	x	x
	libpfm		x	x
	numactl-devel			x

**Table 2. SUSE\* Linux\* Enterprise Server OS Distribution RPMs Matrix**

OS Distribution RPMs		SLES* 12 SP2	SLES* 12 SP3
<b>IB/OPA Centric</b>			
	libibmad5	x	x
	libibumad-devel	x	
	libibverbs1	x	x
	librdmacm1	x	x
	libibcm1	x	x
	ibacm	x	x
	qperf	x	x
	perftest	x	x
	rdma	x	
	opensm-devel	x	x
	opensm-libs3	x	x
	libpsm_infinipath1	x	x
	libibumad3	x	x
	rdma-core		x
	rdma-core-devel		x
<b>OPA Centric</b>			
	libhfi1verbs-rdmapv2	x	
<b>System Centric</b>			
	bc (command line calculator for floating point math)	x	x
	rpm-build	x	x
	kernel-devel	x	x
	libjson-c	x	x
	libexpat1	x	x
	libelf-devel	x	x
	gcc-fortran	x	x
	libudev-devel	x	x
<b>continued...</b>			



OS Distribution RPMs		SLES* 12 SP2	SLES* 12 SP3
	coreutils	x	
	libelf1	x	x
	libgfortran3	x	x
	libgomp1	x	x
	libncurses5	x	x
	libnuma1	x	x
	libquadmath0	x	x
	libz1	x	x
	libopenssl1_0_0	x	x
	tcl	x	x
	expect	x	x
	tcsh	x	x
<b>HFI ASIC-lite support</b>			
	infiniband-diags	x	
	libipathverbs-rdmasv2	x	
<b>Other/Generally Installed</b>			
	bash	x	x
	glibc	x	x
	grep	x	x
	irqbalance	x	x
	kmod	x	x
	libedit0	x	x
	perl	x	x
	perl-base	x	x
	pkg-config	x	x
	python-base	x	x
	systemd	x	x
	libgcc_s1	x	x
	perl-base	x	x
	libstdc++6	x	x
	udev	x	x
<b>Build Requirements</b>			
	libibverbs-devel	x	
	librdmacm-devel	x	
	ibacm-devel	x	
	libopenssl-devel (1.0.1 or higher)	x	x
<b>continued...</b>			



OS Distribution RPMs		SLES* 12 SP2	SLES* 12 SP3
	libuuid-devel	x	x
	libexpat-devel	x	x
	libibumad-devel	x	
	infinipath-psm-devel	x	
	valgrind-devel	x	x
	libjson-c-devel	x	x
	libnuma-devel		x

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

The following is only an example and is showing an install of a group of OS RPMs on a RHEL OS server. Not all required OS RPMs are included in this example and some OS RPMs in this example might not be needed. Intel recommends building your own list of OS RPMs for installation.

```
# yum install libibmad libibverbs librdmacm libibcm qperf perftest
rdma infinipath-psm expat elfutils-libelf-devel libstdc++-devel
gcc-gfortran atlas tcl expect tcsh sysfsutils pciutils bc libibumad
libibumad-devel libibumad libibumad-devel libibverbs-devel libibmad-devel
librdmacm-devel libibcm-devel openssl-devel libuuid-devel expat-devel
infinipath-psm-devel valgrind-devel libgnome libibverbs*
opensm-libs libhfi1 papi ncurses-devel hwloc hwloc-gui
```

#### 4.1.4 MPI Compiler Prerequisites

Ensure that the following requirements are met prior to installing the software.

Two of the MPIs that are downloaded with the OPA-Basic Software package are listed in the Intel OPA Install Menu, MVAPICH2 and OpenMPI. These MPIs were built with Intel compilers.

- To use the MPIs for compiling and running your applications and benchmarks, you need to install a version of the Intel® Parallel Studio XE 2016 Composer Edition that includes the C++, and/or Fortran compilers according to your needs.
- Intel recommends, at a minimum, to 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>.

#### 4.1.5 Performance Tuning Prerequisites

Intel recommends that you pre-configure servers and settings to tune fabric performance to meet the needs of your system. These tasks can be performed before or after the installation. Refer to the *Intel® Omni-Path Fabric Performance Tuning User Guide* which describes BIOS settings and parameters that have been shown to improve performance, or make performance more consistent, on Intel® Omni-Path Architecture. If you are interested in benchmarking the performance of your system, these tips may help you obtain better performance.



## 4.2 Download the Intel® Omni-Path Software

If the OS you installed did not include the OPA RPMs, download the software package from an Intel web page or other Intel secured location using the following procedures.

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` and press **Enter**.
3. In the Description list select the "Intel® Omni-Path Fabric Software (Including Intel® Omni-Path Host Fabric Interface Driver)" for the version you are wanting to install .

*Note:* The latest version of each type of each download type is showing in the list. To show previous versions select "Show more" at the bottom of the list.

4. In the "Available Downloads" list, select the file(s) you need for the OS you have installed on your fabric.
  - For OPA-Basic software, select:

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

- For OPA-IFS software, select:

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

*Note:* `DISTRO.VERSION` refers to the distribution and CPU.

5. Review the Intel Software License Agreement.
6. Click "I accept the terms in the license agreement."
7. Save the download to your hard drive.

## 4.3 Unpack the Tar File

You unpack the tar file using the following procedure.

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.
  - For OPA-Basic, use:

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

- For OPA-IFS, use:

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



## 4.4 Download and Install NVIDIA\* Software (Optional)

If you are installing the Intel® Omni-Path Software onto a machine that has a non-standard kernel, the software may need to recompile the Intel® Omni-Path Host Fabric Interface (HFI) driver. If this is required while installing the GPUDirect\* components, you will need the NVIDIA\* driver source code and symbols available in order to compile the HFI GPUDirect\* enabled driver.

Use the following instructions to download and build the NVIDIA\* driver:

1. Download the NVIDIA\* driver source from:

<http://www.nvidia.com/Download/index.aspx>

It is also available in the CUDA\* Toolkit package that can be downloaded from:

<https://developer.nvidia.com/cuda-downloads>

*Note:* Users of CUDA\* applications will need the CUDA\* runtime as well, which is also part of the CUDA\* Toolkit package. Intel recommends the full CUDA\* Toolkit, but acknowledges that some administrators may only want the driver itself.

2. Extract the driver source and build the driver module as per the instructions at:

<http://docs.nvidia.com/cuda/gpudirect-rdma/index.html#linking-kernel-module-against-nvidia-ko>

3. Install the NVIDIA\* driver module if it's not already installed.

4. Run the command `export NVIDIA_GPU_DIRECT=<path to NVIDIA driver src build directory>` so the Intel® Omni-Path Software installer knows where to find the NVIDIA\* driver source when re-building the HFI driver.

*Note:* The `<path to NVIDIA driver src build directory>` is the directory where the `Module.symvers` file is present after the NVIDIA\* driver is built.

Additionally for any runtime applications that need CUDA\* runtime support you will need to have CUDA\* runtime installed for these applications. Intel recommends downloading the entire CUDA\* Toolkit and installing it.

Download and Installation instructions from NVIDIA\* can be found here:

<http://docs.nvidia.com/cuda/cuda-installation-guide-linux/>





## 5.0 Install the Intel® Omni-Path Software

---

This section provides information and procedures to install the Intel® Omni-Path Software on the Management Node or on a host node in the fabric.

You install the software using one of the following methods:

- [Install using TUI menus \(recommended\).](#)
- [Install using CLI commands.](#)
- [Install Using the Linux\\* Distribution Software](#) on page 47

Intel recommends that you install the OPA-IFS software on the Management Node using the Install TUI, and then use 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 either the FastFabric TUI or a provisioning or diskless boot mechanism.

**Note:** If you are using a provisioning system, consult the documentation that comes with the provisioning system.

Use the [Checklists](#) on page 26 to track your installation of the software.

### 5.1 Before You Begin

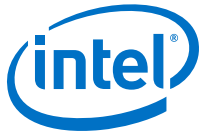
Before starting the installation, perform the following:

- Refer to the Release Notes for the list of compatible operating systems.
- Be sure you have completed all [Pre-Installation Requirements](#) on page 30.
- You have downloaded and extracted the software package per [Installation Getting Started](#) on page 30.
- If you are installing via Linux\* Distribution (Distro), you have completed the [OS RPMs Installation Prerequisites](#) on page 33.
- If you are using a customized installation via the Install CLI command, prepare your command line options. Refer to [Using the ./INSTALL Command Line Options](#).
- For IPoIB IPV4, gather your list of the IP addresses and netmasks for each interface you are going to set up.

#### 5.1.1 Using the ./INSTALL Command Line Options

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`

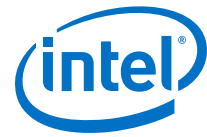


## Syntax

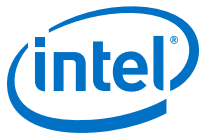
```
./INSTALL [-r root] [-v|-vv] -R osver -B osver  
[-a|-n|-U|-u|-s|-O|-N|-i comp|-e comp] [-E comp] [-D comp]  
[--user-space] [--user_configure_options 'options'] [--kernel_configure_options  
'options']  
[--prefix dir] [--without-depcheck] [--rebuild] [--force] [--answer keyword=value]
```

## Options

No option selected	Displays the Intel OPA Software TUI.
<code>-r root</code>	<p>Specifies an alternate root directory. The default is <code>/</code>.</p> <p><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.</p> <p><i>Note:</i> Intel® Omni-Path Fabric Suite FastFabric use is not permitted in this environment.</p>
<code>-v</code>	Provides verbose logging. Logs to the <code>/var/log/opa.log</code> file.
<code>-vv</code>	Provides very verbose debug logging. Logs to the <code>/var/log/opa.log</code> file.
<code>-R osver</code>	Force install for specific OS kernel version, rather than running kernel.
<code>-B osver</code>	Run build for all components targetting specific OS kernel version.
<code>-a</code>	<p>Installs all Upper Layer Protocols (ULP) and drivers with the default options, including UEFI firmware.</p> <p><i>Note:</i> This rebuilds the RPMs except for CUDA* which is not rebuilt unless <code>-G</code> is specified.</p>
<code>-n</code>	Installs all ULPs and drivers with the default options, but does not change the autostart options.
<code>-U</code>	Upgrades/reinstalls all presently installed ULPs and drivers with the default options, and does not change the autostart options.
<code>-u</code>	Uninstalls all ULPs and drivers with the default options.



<code>-s</code>	Enables autostart for all installed software.
<code>-O</code>	Keeps the current modified rpm configuration file.
<code>-N</code>	Uses a new default rpm configuration file.
<code>-i comp</code>	Installs the given component with the default options. This option can appear multiple times on a command line.
<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>--user-space</code>	Skips 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>--rebuild</code>	Forces a rebuild of OFED_DELTA srpms.
<code>--force</code>	Forces the installation, even if the distributions do not match. Use of this option can result in undefined behaviors.



`--answer keyword=value`

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** Allows 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.

**IrqBalance** Set `IrqBalance` to `Exact`?

`-C`

Shows the list of supported component names.

`-V`

Outputs the version number of the software.

`-G`

Install `GPUDirect*` components.

*Note:* Only `RHEL* 7.3`, `RHEL* 7.4`, and `SLES* 12 SP2` support `GPUDirect*`.

`--user_queries`

Permits non-root users to query the fabric. This is the default.

`--no_user_queries`

Specifies non-root users cannot query the fabric.

### Examples

```
./INSTALL
./INSTALL -n
./INSTALL -C
./INSTALL -V
./INSTALL -a -G
```



## 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, openmpi_gcc_cuda_hfi, gasnet, openshmem, sandiashmem,
mvapich2, openmpi, delta_mpirsrc,
hfil_uefi, delta_debug
```

### Supported Component (comp) Name Aliases:

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

## 5.2 Install Using the TUI Menus

You can install both the Intel® Omni-Path Software Basic and IFS software packages using the Intel OPA Software menu.

### Assumptions

- You are logged in.
- You have your IPV4 list of IP addresses and netmasks for each interface you are going to set up.

### Procedures

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

Step	Task/Prompt	Action
1.	At the command prompt, change directory to the location of the installation software package:	<ul style="list-style-type: none"> <li>For Basic, type the following and press <b>Enter</b>: <code>cd IntelOPA-Basic.DISTRO.VERSION</code></li> <li>For IFS, type the following and press <b>Enter</b>: <code>cd IntelOPA-IFS.DISTRO.VERSION</code></li> </ul> where <i>DISTRO.VERSION</i> is the distribution and CPU.
2.	At the command prompt, start the install script.	Type <code>./INSTALL</code> and press <b>Enter</b> .
3.	Select 1) Install/Uninstall Software.	Type <b>1</b> .
4.	Review the items to be installed.	Accept the defaults. Type <b>N</b> to go to the next page. <b>NOTE:</b> If you need to change any item, enter the alphanumeric character associated with the item to toggle between Install or Don't Install.
5.	Start the installation.	Type <b>P</b> to perform the actions.
6.	Preparing OFA <i>VERSION</i> release for Install... Rebuild OFA SRPMs (a=all, p=prompt per SRPM, n=only as needed?) [n]:	Press <b>Enter</b> to accept the default. <b>NOTE:</b> The system will display prompts that require your response throughout the installation.
7.	For each system prompt...	Accept the defaults by pressing <b>Enter</b> to continue.
8.	When the Intel OPA Autostart Menu displays, review the items.	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.

*continued...*



Step	Task/Prompt	Action
		<b>NOTE:</b> If you need to change any item, enter the alphanumeric character associated with the item to toggle between <i>Enable</i> or <i>Disable</i> .
9.	Run the OPA Autostart operations.	Type <b>P</b> .
10.	For each system prompt, "Hit any key to continue..."	Press any key. <b>NOTE:</b> When the installation completes, you are returned to the main menu.
<b>Set up IPoIB IPV4 Configuration</b>		
11.	Select 2) Reconfigure OFA IP over IB.	Type <b>2</b> .
12.	Configure OFA IP over IB IPV4 addresses now? [n]:	Type <b>y</b> .
13.	How many OFA IP over IB interfaces would you like to configure? [1]:	Type the number of interfaces and press <b>Enter</b> .
14.	For each interface, provide the following information: <ul style="list-style-type: none"><li>• OFA IP over IB interface name</li><li>• IPV4 address in dot notation (or dhcp)</li><li>• IPV4 netmask in dot notation</li></ul>	Provide the required information and press <b>Enter</b> . <b>NOTE:</b> After setup is complete, you are returned to the main menu.
15.	Reboot the server.	Type <i>reboot</i> and press <b>Enter</b> .
	<b>End Task</b>	

### Next Steps

- If you need to set up IPoIB IPV6, go to [Configure IPoIB IPV6](#).
- If you are ready to configure your IFS software, go to [Configure the Chassis](#).

## 5.3 Install Using CLI Commands

You can install both the Intel® Omni-Path Software Basic and IFS software packages using the `./INSTALL`.

The `./INSTALL` command has many options including installing single components, and enabling and disabling autostart of components. This section provides instructions for the default installation, but you can append the install command with specific options for a more customized installation. For more information on the `./INSTALL` syntax and options, refer to [Using the ./INSTALL Command Line Options](#).

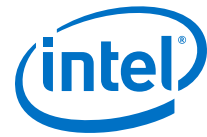
### Assumptions

- You are logged in.
- You have your IPV4 list of IP addresses and netmasks for each interface you are going to set up.

### Procedures

Perform the following steps to install the default Intel® OP Software configuration:

Step	Task/Prompt	Action
1.	At the command prompt, change directory to the location of the installation software package:	<ul style="list-style-type: none"><li>• For Basic, type the following and press <b>Enter</b>: <code>cd IntelOPA-Basic.DISTRO.VERSION</code></li></ul>
<b>continued...</b>		



Step	Task/Prompt	Action
		<ul style="list-style-type: none"> <li>For IFS, type the following and press <b>Enter</b>:  <code>cd IntelOPA-IFS.DISTRO.VERSION</code>            where <i>DISTRO.VERSION</i> is the distribution and CPU.         </li> </ul>
2.	At the command prompt, start the install script.	Type <code>./INSTALL -n</code> and press <b>Enter</b> .
3.	At the command prompt, reboot the server.	Type <code>reboot</code> and press <b>Enter</b> .
<b>Set up IPoIB IPV4 Configuration</b>		
4.	Manually edit or create the <code>/etc/sysconfig/network-scripts/ifcfg-ibX</code> file.	<p><b>Note:</b> Use the OS distribution-supplied instructions for setting up network interfaces.</p> <p>For example:</p> <pre> DEVICE=ib0 BOOTPROTO=static IPADDR=1.1.1.1 BROADCAST=1.255.255.255 NETWORK=1.0.0.1 NETMASK=255.0.0.1 ONBOOT=yes CONNECTED_MODE=yes MTU=65520           </pre>
5.	Create the <code>ifcfg-ibx</code> configuration file for each host using the host IP addresses from the <code>/etc/hosts</code> file.	Type <code>opahostadmin -f /etc/opa/hosts configipoib</code> and press <b>Enter</b> .
	<b>End Task</b>	

### Next Steps

- If you need to set up IPoIB IPV6, go to [Configure IPoIB IPV6](#).
- If you are ready to configure your IFS software, go to [Configure the Chassis](#).

## 5.4 Install Using the Linux\* Distribution Software

You can install the Intel® Omni-Path Software package that is included with the OS distribution.

### Assumptions

- You are logged in.
- You have your IPV4 list of IP addresses and netmasks for each interface you are going to set up.

### Procedures

Perform the following steps to install the default Intel® OP Software configuration:

Step	Task/Prompt	Action
<b>Install OPA-Basic Software</b>		
1.	At the command prompt, enter the installation command for <code>opa-basic-tools</code> .	
	<ul style="list-style-type: none"> <li>For RHEL* 7.3, ...</li> </ul>	Type <code>yum install -y opa-basic-tools</code> and press <b>Enter</b> .
	<ul style="list-style-type: none"> <li>For SLES* 12sp2, ...</li> </ul>	Type <code>zypper install -y opa-basic-tools</code> and press <b>Enter</b> .
<b>continued...</b>		



Step	Task/Prompt	Action
2.	At the command prompt, reboot the server.	Type <b>reboot</b> and press <b>Enter</b> .
3.	Check your link using <code>opainfo</code>	Type <b>opainfo</b> and press <b>Enter</b> . Example output:  <pre> hfi1_0:1                               PortGID: 0xfe80000000000000:001175010163f931   PortState:      Active     LinkSpeed      Act: 25Gb      En: 25Gb     LinkWidth      Act: 4         En: 4     LinkWidthDnGrd ActTx: 4  Rx: 4 En: 3,4     LCRC           Act: 14-bit    En: 14-bit,16-bit, 48-bit           Mgmt: True     LID: 0x00000010-0x00000010      SM LID: 0x0000000c SL: 0   QSFP: AOC      ,      5m FINISAR CORP      P/N FCBN425QB1C05    Rev A   Xmit Data:      0 MB Pkts:             251   Recv Data:      0 MB Pkts:             251   Link Quality: 5 (Excellent) </pre>
4.	Install the <code>rdma-core</code> rpm.	
	• For RHEL* 7.3, ...	Type <code>yum install -y rdma-core</code> and press <b>Enter</b> .
	• For SLES* 12sp2, ...	Type <code>zypper install -y rdma-core</code> and press <b>Enter</b> .
5.	On all compute nodes: install the PSM2 library.	
	• For RHEL* 7.3, ...	Type <code>yum install -y libpsm2</code> and press <b>Enter</b> .
	• For SLES* 12sp2, ...	Type <code>zypper install -y libpsm2-2</code> and press <b>Enter</b> .
<b>Install Intel® Omni-Path Fabric Suite Components on the Management Node</b>		
6.	Install FastFabric.	
	• For RHEL* 7.3, ...	Type <code>yum install -y opa-fastfabric</code> and press <b>Enter</b> .
	• For SLES* 12sp2, ...	Type <code>zypper install -y opa-fastfabric</code> and press <b>Enter</b> .
7.	For RHEL* 7.3 only: Install the <code>opa-address-resolution</code> rpm.	Type <code>yum install -y opa-address-resolution</code> and press <b>Enter</b> .
8.	Install Fabric Manager.	
	• For RHEL* 7.3, ...	Type <code>yum install -y opa-fm</code> and press <b>Enter</b> .
	• For SLES* 12sp2, ...	Type <code>zypper install -y opa-fm</code> and press <b>Enter</b> .
9.	Start the Fabric Manager.	Type <code>systemctl start opafm</code> and press <b>Enter</b> .
<b>Set up IPoIB IPV4 Configuration</b>		
10.	Manually edit or create the <code>ifcfg-ibX</code> file.	<i>Note:</i> Use the OS distribution-supplied instructions for setting up network interfaces.
	• For RHEL* 7.3, ...	Type <code>cat /etc/network-scripts/ifcfg-ib0</code> and press <b>Enter</b> . Example output:  <pre> DEVICE=ib0 BOOTPROTO=static IPADDR=10.228.200.173 </pre>

*continued...*





Step	Task/Prompt	Action
		<pre>BROADCAST=10.228.203.255 NETWORK=10.228.200.0 NETMASK=255.255.252.0 ONBOOT=yes CONNECTED_MODE=yes MTU=65520</pre>
	<ul style="list-style-type: none"> <li>For SLES* 12sp2, ...</li> </ul>	Type <code>cat /etc/network/ifcfg-ib0</code> and press <b>Enter</b> . Example output: <pre>DEVICE=ib0 BOOTPROTO=static IPADDR=10.228.200.173 BROADCAST=10.228.203.255 NETWORK=10.228.200.0 NETMASK=255.255.252.0 ONBOOT=yes CONNECTED_MODE=yes MTU=65520</pre>
11.	Bring up the ib0 interface.	Type <code>ifup ib0</code> and press <b>Enter</b> .
12.	Perform a test ping.	Type <code>ping &lt;remote IPoIB address&gt;</code> and press <b>Enter</b> . For example: <pre>ping 10.228.200.161 PING 10.228.200.161 (10.228.200.161) 56(84) bytes of data: 64 bytes from 10.228.200.161: icmp_seq=1 ttl=64 time=0.863 ms</pre>
<b>(Optional) Install the Fabric Manager GUI</b>		
13.	On one node in the fabric: install the Fabric Manager GUI.	<i>Note:</i> Intel recommends not to install the Fabric Manager GUI on the Management Node where the Fabric Manager is being used.
	<ul style="list-style-type: none"> <li>For RHEL* 7.3, ...</li> </ul>	Type <code>yum install -y opa-fmgui</code> and press <b>Enter</b> .
	<ul style="list-style-type: none"> <li>For SLES* 12sp2, ...</li> </ul>	Type <code>zypper install -y opa-fmgui</code> and press <b>Enter</b> .
<b>End Task</b>		

### Next Steps

- If you need to set up IPoIB IPV6, go to [Configure IPoIB IPV6](#).
- If you are ready to configure your IFS software, go to [Configuring the Software](#).

## 5.5 Configure IPoIB IPV6

You configure IPoIB for IPV6 on the management node. Instructions are dependent on the particular OS installed on the node.

### Using Red Hat\* Enterprise Linux\*

To configure IPoIB for IPV6, perform the following:

- 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.

### Using SUSE\* Linux\* Enterprise Server

To configure IPoIB for IPv6, perform the following:

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

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

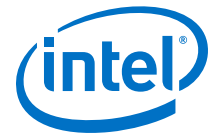
IPv6 address should look like the following example:

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

2. Restart the IPoIB network interfaces.

### Next Steps

- At this time, you should tune Verb and IPoFabric performance. Refer to the *Intel® Omni-Path Fabric Performance Tuning User Guide* for instructions.
- If you are ready to configure your IFS software, go to [Configure the Chassis](#).



## **Part 2: Configuring the Software**



## 6.0 Configuration Getting Started

---

This section provides instructions and information for getting started with the Intel® Omni-Path Fabric Suite configuration.

You configure the fabric components using one of the following methods:

- Using the FastFabric TUI menus (recommended).
- Using the CLI commands.

Use the [Checklists](#) on page 26 to track your configuration and verification of the software.

### 6.1 Pre-Configuration Requirements

This section provides the information and procedures needed prior to configuring and verifying the fabric software.

#### 6.1.1 Chassis Configuration Prerequisites

Ensure that the following requirements are met prior to configuring the chassis.

*Note:*

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

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.

### 6.1.2 Switch Configuration Prerequisites

Ensure that the following requirements are met prior to configuring the externally-managed switch.

1. **(Switch)** Select a unique name to be used for each switch. This name is configured as the Switch Description for the switch.

*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. 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 How to Use the FastFabric TUI

The FastFabric TUI menus are set up for ease of use. The submenus are designed to present operations in the order they would typically be used during an installation.

*Note:* All FastFabric TUI menu alpha-based options are case-insensitive.

### Selecting Menu Items and Performing Operations

1. From the Intel FastFabric OPA Tools main menu, select the target menu item (0-4).

```
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
```

**Note:** Menu item 5 provides access to the Fabric Performance Monitor (opatop).

The target menu is displayed as shown in the example below:

```
FastFabric OPA Chassis Setup/Admin Menu
Chassis File: /etc/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) Set Up Chassis Basic Configuration     [ Skip ]
4) Set Up 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)
```

2. Type the key corresponding to the target menu item (0-9, a-d) to toggle the Skip/Perform selection.

More than one item may be selected.

3. Type P to perform the operations that were selected.

**Notes:**

- If more than one menu item is selected, the operations are performed in the order shown in the menu. This is the typical order desired during fabric setup.
- If you want to perform operations in a different order, you must select the first target menu item, type P to perform the operation, then repeat this process for the next menu item operation to be performed, and so on.

4. Type N to clear all selected items.
5. Type X or press Esc to exit this menu and return to the Main Menu.

### Aborting Operations

While multiple menu items are performing, you have an opportunity to abort individual operations as they come up. After each operation completes and before the next operation begins, you are prompted as shown below:

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

- Press Esc to stop the sequence of operations return to the previous menu.

Any unperformed operations are still highlighted in the menu. To complete the selected operations, type P.



- Press any other key to perform the next selected menu item being performed.  
This prompt is also shown after the last selected item completes, providing an opportunity to review the results before the screen is cleared to display the menu.

### Submenu Configuration Files

On each FastFabric submenu, item 0 permits a different file to be selected and edited (using the editor selected by the EDITOR environment variable). It also permits reviewing and editing of the `opafastfabric.conf` file. The `opafastfabric.conf` file guides the overall configuration of FastFabric and describes cluster-specific attributes of how FastFabric operates.

At the top of each FastFabric submenu screen beneath the title, the directory and configuration file containing the components on which to operate are shown.

In the example below, the configuration file is noted in bold.

```
FastFabric OPA Host Setup Menu
Host File: /etc/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 ]
```

**Note:** During the execution of each menu selection, the actual FastFabric command line tool being used is shown. This can be used as an educational aid to learn the command line tools.

The example snippet below shows how the CLI is displayed in the TUI execution.

```
Performing Chassis Admin: Verify Chassis via Ethernet Ping
Executing: /usr/sbin/opapingall -C -p -F /etc/opa/chassis
```



## 7.0 Configure the Chassis

---

This section provides information and procedures to configure the internally-managed switches, such as the Intel® Omni-Path Director Class Switch 100 Series, using the Intel® Omni-Path Fabric Suite FastFabric toolset.

**Note:** For information about configuring switches made by other manufacturers, refer to the vendor's switch documentation.

### 7.1 Before You Begin

Before starting the chassis configuration, perform the following:

- Be sure you have completed all [Chassis Configuration Prerequisites](#) on page 52.
- Be sure you have the latest *Intel® Omni-Path Fabric Switches Release Notes* for reference.
- Gather your information for the configuration files, as needed.

Note that you can edit the files before you configure the chassis. Configuration files are located under the `/etc/opa` directory. Sample files are installed into `/usr/share/opa/samples` with the suffix `-sample`.

The following files are used to configure the chassis:

- `opafastfabric.conf`: Lists the default settings for most of the FastFabric command line options.
- `ports`: Lists the local HFI ports to use to access the fabrics.

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.

- `chassis`: Lists the TCP/IP names of the internally-managed switches in the cluster.

Intel recommends you use chassis names (the TCP/IP Ethernet management port names assigned). Enter one chassis name or IP address per line. **Note:** Do not list externally-managed switches in this file.

For example:

```
Chassis1
Chassis2
```

For more details about configuration files, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

For more details about the file format of the configuration files, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

**Note:** If you are configuring multiple subnets, refer to [Multi-Subnet Fabrics](#) on page 106 for additional instructions.





## 7.2 Configure the Chassis Using the FastFabric TUI Menu

You can configure the chassis using the FastFabric OPA Chassis Setup/Admin Menu.

### Assumptions

- You are logged into the management node.
- You have the latest *Intel® Omni-Path Fabric Switches Release Notes* for reference.

### Pre-Work

For the following menu items, prepare to answer questions or provide input as follows:

- **Updating the Chassis FM Security Files**
  - Choose 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.
  - Select security files or directory containing pem files
  - Choose parallel versus serial update
  - Set chassis password (default is to have password in fastfabric.conf or to use password-less SSH)
- **Checking the OPA Fabric Status**
  - Would you like to perform fabric error analysis? [y]:  
Allows you to start the analysis.
  - Clear error counters after generating report? [n]:  
Allows you to clear the error counters after generating the report.
  - Would you like to perform fabric link speed error analysis? [y]:  
Allows you to analyze fabric link speed errors.
  - Check for links configured to run slower than supported? [n]:  
Allows you to check for Links running slower than expected.
  - Check for links connected with mismatched speed potential? [n]:  
Allows you to check for links connected with mismatched speed.
  - Enter filename for results [/root/linkanalysis.res]:  
Allows you to enter a filename for the results or use the default file.
- **Control Chassis Fabric Manager**
  - Would you like to make sure the FM is not running? [n]:  
Allows you to ensure that the FM is not running.
  - Would you like to make sure the FM is running? [n]:  
Allows you to ensure that the FM is running.



- Would you like to run FM on slave MMs? [n]:  
Allows you to run FM on slave management modules.
- Would you like to do the operation in parallel? [y]:  
Allows you to perform operations in parallel (on multiple chassis). Doing the operation in parallel will finish the fastest.
- Would you like to change FM boot state to enable FM start at boot? [n]:  
Allows you to enable FM start on slave management modules at boot.
- Would you like to change FM boot state to disable FM start at boot? [n]:  
Allows you to disable FM start on slave management modules at boot.
- Would you like to be prompted for chassis' password? [n]:  
Allows you to be prompted for the chassis password.

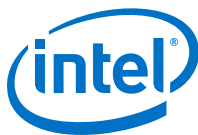
### Procedures

The following steps provide simplified instructions for first-time configuration of the chassis. For additional details, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

Step	Task/Prompt	Action
<b>Configuring the Chassis</b>		
1.	Access the <b>FastFabric OPA Chassis Setup/Admin menu</b> .	
	a) If you are not already logged into <b>Intel FastFabric OPA Tools</b> , at the command prompt...	Type <code>opafastfabric</code> and press <b>Enter</b> .
	b) Access the <b>FastFabric OPA Chassis Setup/Admin menu</b> .	Press <b>1</b> .
	c) Select menu items.	Select all operations by entering each menu item 0 – e. <b>NOTE:</b> You can skip any items that are not needed for your fabric. However, this procedure assumes that all items have been selected.
	d) Start the operations.	Press <b>P</b> . <b>NOTE:</b> Each selected item is preformed in the order of the menu list.
2.	<b>Edit the Configuration and Select/Edit Chassis File</b> (menu item 0)	
	a) Edit the <code>opafastfabric.conf</code> file.	Review the file with a focus on the following: <ul style="list-style-type: none"><li>• <code>FF_CHASSIS_LOGIN_METHOD</code></li><li>• <code>FF_CHASSIS_ADMIN_PASSWORD</code></li><li>• Select the location for the result files from FastFabric with the <code>FF_RESULT_DIR</code> parameter.</li></ul> If you made any changes, save and close the file. Press any key to continue.
	b) Edit the <code>ports</code> configuration file.	Review the file. If you made any changes, save and close the file. Press any key to continue.
<b>continued...</b>		



Step	Task/Prompt	Action
	c) Edit the chassis configuration file.	Create the file with a list of the chassis names (the TCP/IP Ethernet management port names assigned) or IP addresses. If you made any changes, save and close the file.
	d) Do you want to edit/review/change the files? [y]:	Type <b>n</b> and Press <b>Enter</b> .
	e) Continue to next step.	Press any key to continue.
3.	<b>Verify Chassis via Ethernet Ping</b> (menu item 1)	
	• If all chassis were found...	Press any key to continue to <b>Update the Chassis Firmware</b> .
	• If some chassis were not found...	Press <b>ESC</b> to exit the menu and review the following list for those chassis which were not found: <ul style="list-style-type: none"> <li>• Is chassis powered on and booted?</li> <li>• Is chassis connected to management network?</li> <li>• Are chassis IP address and network settings consistent with DNS or /etc/hosts ?</li> <li>• Is Management node connected to the management network?</li> <li>• Are Management node IP address and network settings correct?</li> <li>• Is management network itself up (including switches, routers, and others)?</li> <li>• Is correct set of chassis listed in the chassis file? You may need to repeat the previous step to review and edit the file.</li> </ul>
4.	(Optional) <b>Update Chassis Firmware</b> (menu item 2)	<b>NOTE:</b> Before continuing, refer to the <i>Intel® Omni-Path Fabric Switches Release Notes</i> for any prerequisites.
	a) For Directories, all .spkg files in the directory tree will be used. Enter Files/Directories to use (or none):	Specify the directory where the relevant firmware files have been stored and press <b>Enter</b> .
	b) Would you like to run the firmware now? [n]:	Type <b>y</b> and press <b>Enter</b> . <b>NOTE:</b> 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 <b>View opachassisadmin Result Files</b> option to review the result files from the update. Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more details.
	c) Continue to next step.	Press any key to continue.
5.	<b>Set Up Chassis Basic Configuration</b> (menu item 3)	
	a) For each prompt...	Accept the defaults.
	b) Continue to next step.	Press any key to continue.
6.	<b>Set Up Password-less ssh/scp</b> (menu item 4)	
	a) Would you like to override the default Chassis password? [n]:	Press <b>Enter</b> to accept the default Chassis password.
	b) Continue to next step.	Press any key to continue.
7.	<b>Reboot the Chassis</b> (menu item 5)	
	a) Performing Chassis Admin: Reboot Chassis	Press <b>Enter</b> to accept the default.
<b>continued...</b>		



Step	Task/Prompt	Action
	Would you like to be prompted for chassis' password? [n]:	
	b) Continue to next step.	Press any key to continue.
8.	<b>Get Basic Chassis Configuration</b> (menu item 6)	
	a) Continue to next step.	Press any key to continue.
9.	<b>Configure Chassis Fabric Manager</b> (menu item 7)	<b>NOTE:</b> For configuring multi-subnets: At least one subnet manager is required per subnet. Refer to the <i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i> for more information on how to configure a host SM node to manage more than one subnet.
	a) Performing Chassis Admin: Configure Chassis Fabric Manager (FM) Enter FM Config file to use (or none or generate):	Type <b>generate</b> and press <b>Enter</b> .
	b) For each prompt...	Accept the defaults.
	c) Would you like to restart the FM? [n]:	Type <b>y</b> and press <b>Enter</b> .
	d) Would you like to run the FM on slave MMs? [n]:	Choose: <ul style="list-style-type: none"> <li>Type <b>y</b>. If 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.</li> <li>Type <b>n</b>. If our 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.</li> </ul>
	e) Would you like to do the operation in parallel? [y]:	Press <b>Enter</b> .
	f) Would you like to enable FM start at boot? [n]:	Type <b>y</b> .
	g) Would you like to enable FM start on slave MMs at boot? [n]:	Choose: <ul style="list-style-type: none"> <li>Type <b>y</b>. If 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.</li> <li>Type <b>n</b>. If 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.</li> </ul>
	h) Would you like to be prompted for chassis' password? [n]:	Press <b>Enter</b> .
	i) Are you sure you want to proceed? [n]:	Type <b>y</b> .
	j) Continue to next step.	Press any key to continue.
10.	<b>Updating the Chassis FM Security Files</b> (menu item 8)	
	a) Enter Files/Directories to use (or none):	<b>NOTE:</b> You must enter a file or the task ends.
<b>continued...</b>		



Step	Task/Prompt	Action
	b) For subsequent prompts...	Provide the required information and press <b>Enter</b> .
	c) Continue to next step.	Press any key to continue.
11.	<b>Get the Chassis FM Security File</b> (menu item 9)	
	a) Continue to next step.	Press any key to continue.
<b>Verifying the Chassis Configuration</b>		
12.	<b>Check OPA Fabric Status</b> (menu item a)	
	a) For each prompt...	Provide the required information and press <b>Enter</b> .
	b) Continue to next step.	Press any key to continue.
13.	<b>Control Chassis Fabric Manager</b> (menu item b)	
	a) For each prompt...	Provide the required information and press <b>Enter</b> .
	b) Are you sure you want to proceed? [n]:	Select <b>y</b> .
	c) Continue to next step.	Press any key to continue.
14.	<b>Generate All Chassis Problem Report Information</b> (menu item c)	
	a) Would you like to be prompted for chassis' password? [n]:	Press <b>Enter</b> .
	b) Continue to next step.	Press any key to continue.
15.	<b>Run a Command on All Chassis</b> (menu item d)	
	a) If there are any other operations that need to be performed on all chassis...	Use the <b>Run a Command on All Chassis</b> 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. Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> .
	b) Continue to next step.	Press any key to continue.
16.	<b>View the opachassisadmin Result Files</b> (menu item e)	
	a) About to: vi /root/punchlist.csv /root/test.res /root/test.log	Press any key to review files or press <b>ESC</b> to abort.
	b) Continue to next step.	Press any key to continue.
	<b>End Task</b>	

### Next Steps

- To configure the externally-managed switches, go to [Configure the Externally-Managed Switches](#).

## 7.3 Configure the Chassis Using the FastFabric CLI Commands

You can configure the chassis using the CLI commands.

### Assumptions

- You are logged into the management node.
- You have the latest *Intel® Omni-Path Fabric Switches Release Notes* for reference.



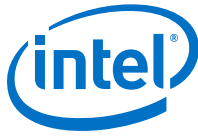
## Procedures

The following steps provide simplified instructions for first-time configuration of the chassis. For additional details, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

Step	Task/Prompt	Action
1.	(Optional) Edit or review the configuration files.	
	a) Edit the <code>/etc/opa/opafastfabric.conf</code> file.	Review the file with a focus on the following: <ul style="list-style-type: none"><li>FF_CHASSIS_LOGIN_METHOD</li><li>FF_CHASSIS_ADMIN_PASSWORD.</li><li>Select the location for the result files from FastFabric with the FF_RESULT_DIR parameter.</li></ul> If you made any changes, save and close the file.
	b) Edit the <code>/etc/opa/ports</code> file.	Review the file. If you made any changes, save and close the file.
	c) Edit the <code>/etc/opa/chassis</code> file.	Create the file with a list of the chassis names (the TCP/IP Ethernet management port names assigned) or IP addresses. <i>Note:</i> Do not list externally-managed switches in this file. If you made any changes, save and close the file.
2.	Verify the chassis is accessible over the management network.	Type <code>opapingall -C -p -f /etc/opa/chassis</code> .
	<ul style="list-style-type: none"><li>If all chassis were found...</li></ul>	Continue to the next step.
	<ul style="list-style-type: none"><li>If some chassis were not found...</li></ul>	Review the following list for those chassis that were not found: <ul style="list-style-type: none"><li>Is chassis powered on and booted?</li><li>Is chassis connected to management network?</li><li>Are chassis IP address and network settings consistent with DNS or <code>/etc/hosts</code> ?</li><li>Is Management node connected to the management network?</li><li>Are Management node IP address and network settings correct?</li><li>Is management network itself up (including switches, routers, and others)?</li><li>Is correct set of chassis listed in the chassis file? You may need to repeat the previous step to review and edit the file.</li></ul>
3.	Update the chassis firmware using either the parallel or serial method.	<b>NOTE:</b> Before continuing, refer to the <i>Intel® Omni-Path Fabric Switches Release Notes</i> for any prerequisites.
	a) For the parallel method...	Type <code>opachassisadmin -S -F chassisfile -P package -a run upgrade</code> .
	b) For the serial method...	Type <code>FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -P package -a run upgrade</code> .
	<b>NOTES:</b> During the firmware update, the following files are produced: <ul style="list-style-type: none"><li><code>test.res</code> – Appended with summary results of run</li><li><code>test.log</code> – Appended with detailed results of run</li><li><code>save_tmp/</code> – Contains a directory per failed operation with detailed logs</li><li><code>test_tmp*/</code> – Intermediate result files while operation is running</li></ul> If any chassis fails to be updated, refer to the <b>View opachassisadmin result files</b> section to review the result files from the update.	
continued..		



Step	Task/Prompt	Action
	Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more details.	
4.	Set up the chassis with the basic configuration settings using the command.	Type <b>opachassisadmin -S -F chassisfile configure</b> .
	a) Follow the system prompts to select and configure the items required.	This wizard gives you the option of setting up the following items on the chassis: <ul style="list-style-type: none"> <li>• Syslog server</li> <li>• NTP server</li> <li>• Timezone and DST information</li> <li>• Chassis link width</li> <li>• Configure Node Desc to match Ethernet* chassis name</li> <li>• Configure the Link CRC Mode</li> </ul>
5.	Set up secure password-less SSH, such that the Management Node can securely log into all the chassis as <b>admin</b> through the management network, without requiring a password.	Type <b>opasetupssh -p -S -C -F chassisfile</b> .
6.	Reboot all the selected chassis and ensure they reboot fully, as verified through ping over the management network.	Type <b>opachassisadmin -S -F chassisfile reboot</b> .
7.	Retrieve basic information from chassis such as syslog, NTP configuration, time zone, node description, and other information.	Type <b>opachassisadmin -F chassisfile getconfig</b> .
8.	Configure the Chassis Fabric Manager for any Intel® Omni-Path Chassis 100 Series chassis using either parallel or serial method.	<b>NOTE:</b> For configuring multi-subnets: At least one subnet manager is required per subnet. Refer to the <i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i> for more information on how to configure a host SM node to manage more than one subnet.
	a) For the parallel method...	Type <b>opachassisadmin -S -F chassisfile -P package -a run fmconfig</b> .
	b) For the serial method...	Type <b>FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -P package -a run fmconfig</b> .
9.	Update Chassis Fabric Manager Security Files to permit the chassis security files to be verified and updated using either parallel or serial method, as needed.	<b>Note:</b> 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 <i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i> for instructions on the administration tasks required to support these files.
	a) For the parallel method...	Type <b>opachassisadmin -S -F chassisfile -s securityFiles -a push fmsecurityfiles</b> .
	b) For the serial method...	Type <b>FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -s securityFiles -a push fmsecurityfiles</b> .
10.	Retrieve the Chassis Fabric Manager Security Files from the chassis.	Type <b>opachassisadmin -F chassisfile fmgetsecurityfiles</b> .
<b>Verifying the Chassis Configuration</b>		
11.	Check the OPA fabric status.	Type <b>opalinkanalysis -U -x snapshot_suffix reports verifyall &gt; results_file 2&gt;&amp;1</b> . <b>NOTE:</b> This step performs the following operations: <ul style="list-style-type: none"> <li>• Perform a fabric error analysis.</li> <li>• Clear the error counters after generating a report.</li> </ul>
<b>continued...</b>		



Step	Task/Prompt	Action
		<ul style="list-style-type: none"> <li>Perform a fabric link speed error analysis.</li> <li>Check for links configured to run slower than supported.</li> <li>Check for links connected with mismatched speed potential.</li> <li>Enter a filename for the results or save the results to the default location which is: /root/ffres/linkanalysis.res</li> </ul>
12.	To control the Chassis Fabric Manager, you can perform any or all of the following optional steps:	
	a) Restart all the Chassis Fabric Managers and run the Fabric Manager on the slave Management Module using either parallel or serial methods.	<b>NOTE:</b> There is a disruption as FMs are restarted. Using the serial method may reduce the disruption.
	<ul style="list-style-type: none"> <li>For parallel method...</li> </ul>	Type <code>opachassisadmin -S -F chassisfile -a restartall fmcontrol.</code>
	<ul style="list-style-type: none"> <li>For serial method...</li> </ul>	Type <code>FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -a restartall fmcontrol.</code>
	b) Restart the Master Chassis Fabric Managers and stop the Fabric Manager on the slave Management Modules using either parallel or serial methods.	<b>NOTE:</b> There is a disruption as FMs are restarted. Using the serial method may reduce the disruption.
	<ul style="list-style-type: none"> <li>For parallel method...</li> </ul>	Type <code>opachassisadmin -S -F chassisfile -a restart fmcontrol.</code>
	<ul style="list-style-type: none"> <li>For serial method...</li> </ul>	Type <code>FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -a restart fmcontrol.</code>
	c) Stop the Master Chassis Fabric Managers from running using either parallel or serial methods.	
	<ul style="list-style-type: none"> <li>For parallel method...</li> </ul>	Type <code>opachassisadmin -S -F chassisfile -a stop fmcontrol.</code>
	<ul style="list-style-type: none"> <li>For serial method...</li> </ul>	Type <code>FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -a stop fmcontrol.</code>
	d) Ensure all of the Chassis Fabric Managers are running, including the slaves using either parallel or serial methods.	
	<ul style="list-style-type: none"> <li>For parallel method...</li> </ul>	Type <code>opachassisadmin -S -F chassisfile -a runall fmcontrol.</code>
	<ul style="list-style-type: none"> <li>For serial method...</li> </ul>	Type <code>FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -a runall fmcontrol.</code>
	e) Ensure the Master Chassis Fabric Manager is running and the slaves are stopped using either parallel or serial methods.	
	<ul style="list-style-type: none"> <li>For parallel method...</li> </ul>	Type <code>opachassisadmin -S -F chassisfile -a run fmcontrol.</code>
	<ul style="list-style-type: none"> <li>For serial method...</li> </ul>	Type <code>FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -a run fmcontrol.</code>
	f) Change the Fabric Manager boot state to enable the Master Chassis Fabric Manager to start at boot.	Type <code>opachassisadmin -S -F chassisfile -I enable.</code>
	g) Change the Fabric Manager boot state to enable the all of the Chassis Fabric Managers to start at boot.	Type <code>opachassisadmin -S -F chassisfile -I enableall.</code>
<i>continued...</i>		





Step	Task/Prompt	Action
13.	Generate all chassis problem report information.	Type <code>opacaptureall -p -D 4 -f hostfile</code> .
14.	Run a command on all chassis.	Type <code>opacmdall -C -S -p -T timelimit -F chassisfile STDIN</code> .
15.	View <code>opachassisadmin</code> result files	Type <code>editor result_dir/result_file</code> . <b>NOTE:</b> In the line above, " <i>editor</i> " indicates the command line editor; for example, vi. The following default files are created: <ul style="list-style-type: none"> <li>• <code>punchlist.csv</code></li> <li>• <code>test.res</code></li> <li>• <code>test.log</code></li> </ul>
	<b>End Task</b>	

### Next Steps

- To configure the externally-managed switches, go to [Configure the Externally-Managed Switches](#).



## 8.0 Configure the Externally-Managed Switches

---

This section provides information and procedures to configure the externally-managed switches, such as Intel® Omni-Path Switch 100 Series using the Intel® Omni-Path Fabric Suite FastFabric toolset.

### 8.1 Before You Begin

Before starting the switch configuration, perform the following:

- Be sure you have completed all [Switch Configuration Prerequisites](#) on page 53.
- Be sure you have the latest *Intel® Omni-Path Fabric Switches Release Notes* for reference.
- Gather your information for the configuration files, as needed.

Note that you can edit the files before you configure the switch. Configuration files are located under the `/etc/opa` directory. Sample files are installed into `/usr/share/opa/samples` with the suffix `-sample`.

The following files are used to configure the switch:

- `opafastfabric.conf`: Lists the default settings for most of the FastFabric command line options.
- `ports`: Lists the local HFI ports to use to access the fabric.
- `switches`: Lists the GUID, Node Description, and optional Distance of each externally-managed switch in the cluster.

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. *Note:* Do not list internally-managed chassis in this file.

For example:

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

For more details about configuration files, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

For more details about the file format of the configuration files, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

*Note:* If you are configuring multiple subnets, refer to [Multi-Subnet Fabrics](#) on page 106 for additional instructions.



## 8.2 Configure the Externally-Managed Switches Using FastFabric TUI Menu

You configure the externally-managed switches using the FastFabric OPA Switch Setup/Admin menu.

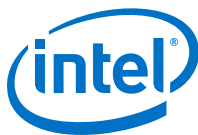
### Assumptions

- You are logged into the management node.
- You have the latest *Intel® Omni-Path Fabric Switches Release Notes* for reference.

### Procedures

The following steps provide simplified instructions for first-time configuration of the switch. For additional details, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

Step	Task/Prompt	Action
<b>Configuring the Externally-Managed Switches</b>		
1.	Access the <b>FastFabric OPA Switch Setup/Admin menu</b> .	
	a) If you are not already logged into <b>Intel FastFabric OPA Tools</b> , at the command prompt...	Type <code>opafastfabric</code> and press <b>Enter</b> .
	b) Access the <b>FastFabric OPA Switch Setup/Admin menu</b> .	Press <b>2</b> .
	c) Select menu items.	Select all operations by entering each menu item 0 – a. <b>NOTE:</b> You can skip any items that are not needed for your fabric. However, this procedure assumes that all items have been selected.
	d) Start the operations.	Press <b>P</b> . <b>NOTE:</b> Each selected item is preformed in the order of the menu list.
2.	<b>Edit the Configuration and Select/Edit Switch File</b> (menu item 0)	
	a) Edit the <code>opafastfabric.conf</code> file.	Review the file. If you made any changes, save and close the file. Press any key to continue.
	b) Edit the <code>ports</code> configuration file.	Review the file. If you made any changes, save and close the file. Press any key to continue.
	c) Edit the <code>switches</code> configuration file.	Create the file with a list of the switch node GUID and required switch names. If you made any changes, save and close the file.
	d) Do you want to edit/review/change the files? [y]:	Type <b>n</b> and Press <b>Enter</b> .
	e) Continue to next step.	Press any key to continue.
3.	<b>Generate or Update Switch File</b> (menu item 1)	
	a) For each prompt...	Provide the required information and press <b>Enter</b> .
	b) Continue to next step.	Press any key to continue.
<b>continued...</b>		



Step	Task/Prompt	Action
4.	<b>Test for Switch Presence</b> (menu item 2)	
	• If all switches were found...	Press any key to continue to <b>Verify Switch Firmware</b> .
	• If some switches were not found...	Press <b>ESC</b> to exit the menu and review the following list for those switches that were not found: <ul style="list-style-type: none"> <li>• Is switch powered on and booted?</li> <li>• Is switch connected to Intel® Omni-Path Fabric?</li> <li>• Is Subnet Manager running?</li> <li>• Is Management Node's Port active?</li> <li>• Is Management Node connected to the correct Intel® Omni-Path Fabric?</li> <li>• 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 <i>Intel® Omni-Path Fabric Switches Hardware Installation Guide</i>.</li> <li>• 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.</li> </ul>
5.	<b>Verify Switch Firmware</b> (menu item 3)	
	a) Continue to next step.	Press any key to continue.
6.	<b>Update Switch Firmware</b> (menu item 4)	<i>Note: Refer to the Intel® Omni-Path Fabric 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.</i>
	a) Enter Files/Directories to use (or none):	Enter the directory where the relevant firmware files are located.
	b) Would you like to reboot the switch after the update? [n]:	Type <b>y</b> .
	c) For subsequent prompts...	Press <b>Enter</b> to accept the defaults.
	d) Continue to next step.	Press any key to continue.
7.	<b>Set Up Switch Basic Configuration</b> (menu item 5)	
	a) For each prompt...	Press <b>Enter</b> to accept the defaults.
	b) Continue to next step.	Press any key to continue.
8.	<b>Reboot Switch</b> (menu item 6)	
	a) Continue to next step.	Press any key to continue.
9.	<b>Report Switch Firmware &amp; Hardware Info</b> (menu item 7)	
	a) Continue to next step.	Press any key to continue.
10.	<b>Get Basic Switch configuration</b> (menu item 8)	
	a) Continue to next step.	Press any key to continue.
11.	<b>Report Switch VPD Information</b> (menu item 9)	
	a) Continue to next step.	Press any key to continue.
<b>Verifying the Switch Configuration</b>		
12.	<b>View opaswitchadmin Result Files</b> (menu item a)	
<b>continued...</b>		



Step	Task/Prompt	Action
	a) About to: vi /root/punchlist.csv /root/test.res /root/test.log	Press any key to review files or press <b>ESC</b> to abort.
	b) Would you like to remove test.res test.log test_tmp* and save_tmp in /root ? [n]:	Press <b>Enter</b> to save or type <b>y</b> to remove the files.
	<b>End Task</b>	

### Next Steps

- To install the host software on the remaining servers, go to [Install Host Software on Remaining Servers](#).

## 8.3 Configure the Externally-Managed Switches Using FastFabric CLI Commands

You configure the externally-managed switches using the FastFabric OPA Switch Setup/Admin menu.

### Assumptions

- You are logged into the management node.
- You have the latest *Intel® Omni-Path Fabric Switches Release Notes* for reference.

### Procedures

The following steps provide simplified instructions for first-time configuration of the switch. For additional details, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

Step	Task/Prompt	Action
1.	(Optional) Edit or review the configuration files.	
	a) Edit the /etc/opa/opafastfabric.conf file.	Review the file. If you made any changes, save and close the file.
	b) Edit the /etc/opa/ports file.	Review the file. If you made any changes, save and close the file.
	c) Edit the /etc/opa/switches file.	Create the file with a list of the switch node GUID and required switch names. If you made any changes, save and close the file.
2.	(Optional) Generate or update the switches file. <b>NOTE:</b> You can also update switch names in the switches file by comparing the actual fabric to topology xml data.	Type <b>opagenswitches -s -o switches</b> .
3.	Test to see if each externally-managed switch is present.	Type <b>opaswitchadmin -I switchfile ping</b> .
	a) If all switches were found...	Continue to the next step.
	b) If some switches were not found...	Review the following list for those switches that were not found: <ul style="list-style-type: none"> <li>Is switch powered on and booted?</li> <li>Is switch connected to Intel® Omni-Path Fabric?</li> <li>Is Subnet Manager running?</li> <li>Is Management Node's Port active?</li> </ul>

*continued...*



Step	Task/Prompt	Action
		<ul style="list-style-type: none"><li>Is Management Node connected to the correct Intel® Omni-Path Fabric?</li><li>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 <i>Intel® Omni-Path Fabric Switches Hardware Installation Guide</i>.</li><li>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.</li></ul>
4.	Verify the switch firmware to check that each externally managed switch is operational and that its firmware is valid and accessible.	Type <code>opaswitchadmin -L switchfile fwverify</code> .
5.	Update the switch firmware version and the switch node name set.	<b>NOTE:</b> Be aware that non-parallel operation for a fabric with many externally-managed switches can take a significant amount of time.
	<ul style="list-style-type: none"><li>For the parallel method...</li></ul>	Type <code>opaswitchadmin -O -L switchfile -P packages -a run upgrade</code> .
	<ul style="list-style-type: none"><li>For the serial method...</li></ul>	Type <code>FF_MAX_PARALLEL=0 opaswitchadmin -O -L switchfile -P packages -a run upgrade</code> .
6.	Reboot the externally managed switches.	Type <code>opaswitchadmin -L switchfile reboot</code> .
7.	Set up the switch basic configuration for all externally managed switches.	Type <code>opaswitchadmin -L switchfile configure</code> . <b>NOTE:</b> This command runs a wizard to set up the externally-managed switch configuration.
8.	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.	Type <code>opaswitchadmin -L switchfile info</code> .
9.	Run the command to get the externally managed switch configuration report for all of the ports.	Type <code>opaswitchadmin -L switchfile getconfig</code> .
10.	Run the report that returns externally managed switch hardware vital product data (VPD) for all of the nodes listed in the <code>/etc/opa/switches</code> file.	Type <code>opaswitchadmin -L switchfile hwvpd</code> .
<b>Verifying the Switch Configuration</b>		
11.	View the <code>opaswitchadmin</code> results files.	Type <code>editor result_dir/result_file</code> . <b>NOTE:</b> In the line above, "editor" indicates the command line editor; for example, vi. The following default files are created: <ul style="list-style-type: none"><li><code>punchlist.csv</code></li><li><code>test.res</code></li><li><code>test.log</code></li></ul>
	<b>End Task</b>	

### Next Steps

- To install the host software on the remaining servers, go to [Install Host Software on Remaining Servers](#).



## 9.0 Install Host Software on Remaining Servers

---

This section provides information and procedures to install, configure, and verify the Intel® Omni-Path Fabric Host Software on the remaining hosts.

**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 Intel® OPA software.

### 9.1 Before You Begin

Before starting the host installation and configuration, perform the following:

- Gather your information for the configuration files, as needed.

Note that you can edit the files before you configure the hosts. Configuration files are located under the `/etc/opa` directory. Sample files are installed into `/usr/share/opa/samples` with the suffix `-sample`.

The following files are used to configure the hosts:

- `opafastfabric.conf`: Lists the default settings for most of the FastFabric command line options.

**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).

- Intel recommends that a FastFabric topology file is created as `/etc/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 User Guide* for more information about topology verification files.
- `ports`: Lists the local HFI ports to use to access the fabric.
  - 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.
- `hosts`: 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 (that is, 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.

- **allhosts:** Lists the Management Node's hosts name (the TCP/IP management network name, for example `mgmthost`) and includes the `hosts` file.

For example:

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

- **hostverify.sh:** Script to verify the configuration and performance of an individual node. This should be run using `opaverifyhosts`. It can also be run on an individual node directly.

For more details about configuration files, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

For more details about the file format of the configuration files, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

**Note:** If you are configuring multiple subnets, refer to [Multi-Subnet Fabrics](#) on page 106 for additional instructions.

- For the following Host Setup menu item, prepare to answer questions or provide input as follows:

- **Building Test Apps and Copying to Hosts**

Choose MPI Directory Selection:

- 0) /usr/mpi/gcc/mvapich2-x.x
- 1) /usr/mpi/gcc/mvapich2-x.x-hfi
- 2) /usr/mpi/gcc/openmpi-x.x.x
- 3) /usr/mpi/gcc/openmpi-x.x.x-hfi
- 4) /usr/mpi/intel/mvapich2-x.x-hfi
- 5) /usr/mpi/intel/openmpi-x.x.x-hfi
- 6) /usr/mpi/pgi/mvapich2-x.x-hfi
- 7) /usr/mpi/pgi/openmpi-x.x.x-hfi
- 8) Enter Other Directory

Choose MPI Directory Selection for SHMEM Job Launch

- 0) /usr/mpi/gcc/openmpi-x.x.x
- 1) /usr/mpi/gcc/openmpi-x.x.x-hfi
- 2) /usr/mpi/intel/openmpi-x.x.x-hfi





- 3) /usr/mpi/pgi/openmpi-x.x.x-hfi
- 4) Enter Other Directory
- 5) Skip MPI Directory Selection for SHMEM Job Launch

*Note:* "x.x" or "x.x.x" above refers to the latest supported versions defined in the *Intel® Omni-Path Fabric Software Release Notes*.

- For the following Host Verify menu item, prepare to answer questions or provide input as follows:
  - **Perform Single Host Verification**
    - Prior to using this operation, 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/share/opa/samples/hostverify.sh` to the directory pointed to by `FF_HOSTVERIFY_DIR`.  
Refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide* for information on `opaverifyhosts`.
    - Determine the tests to be run and time limit (in minutes) for the tests.

## 9.2 Install the Host Software on the Remaining Hosts Using the FastFabric TUI Menu

You can configure the host software using the FastFabric OPA Host Setup menu.

### Assumption

- You are logged into the management node.

### Procedures

The following steps provide simplified instructions for first-time configuration of the hosts. For additional details, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

Step	Task/Prompt	Action
<b>Configuring the Hosts</b>		
1.	Access the <b>FastFabric OPA Host Setup menu</b> .	
	a) If you are not already logged into <b>Intel FastFabric OPA Tools</b> , at the command prompt...	Type <code>opafastfabric</code> and press <b>Enter</b> .
	b) Access the <b>FastFabric OPA Host Setup menu</b> .	Press <b>3</b> .
	c) Select menu items.	<ul style="list-style-type: none"> <li>• Select items <b>0 – 2</b> and <b>4 – 8</b>.</li> <li>• Select item <b>3</b> if you are using <code>/etc/hosts</code> for name resolution (as opposed to using DNS).</li> </ul>
	d) Start the operations.	Press <b>P</b> . <b>NOTE:</b> Each selected item is preformed in the order of the menu list.
2.	<b>Edit Configuration and Select/Edit Host File</b> (menu item 0)	
	a) Edit the <code>opafastfabric.conf</code> file.	Review the file with a focus on the following: <ul style="list-style-type: none"> <li>• <code>FF_IPOIB_SUFFIX</code></li> </ul>
<i>continued...</i>		



Step	Task/Prompt	Action
		<ul style="list-style-type: none"> <li>FF_IPOIB_NETMASK</li> <li>FF_IPOIB_CONFIG</li> <li>FF_PRODUCT</li> <li>FF_PACKAGES</li> <li>FF_INSTALL_OPTIONS</li> <li>FF_UPGRADE_OPTIONS</li> </ul> <p>If you made any changes, save and close the file. Press any key to continue.</p>
	b) Edit the <code>hosts</code> configuration file.	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. If you made any changes, save and close the file.
	c) Do you want to edit/review/change the files? [y]:	Type <b>n</b> and Press <b>Enter</b> .
	e) Continue to next step.	Press any key to continue.
3.	<b>Verify Hosts Pingable</b> (menu item 1)	
	<ul style="list-style-type: none"> <li>If all hosts were found...</li> </ul>	Press any key to continue to <b>Set Up Password-Less SSH/SCP</b> .
	<ul style="list-style-type: none"> <li>If some hosts were not found...</li> </ul>	<p>Press <b>ESC</b> to exit the menu and review the following list for those hosts that were not found:</p> <ul style="list-style-type: none"> <li>Host powered on and booted?</li> <li>Host connected to management network?</li> <li>Host management network IP address and network settings consistent with DNS or <code>/etc/hosts</code>?</li> <li>Management node connected to the management network?</li> <li>Management node IP address and network settings correct?</li> <li>Management network itself up (including switches, routers, and others)?</li> <li>Correct set of hosts listed in the hosts file? You may need to repeat the previous step to review and edit the file.</li> </ul>
4.	<b>Set Up Password-Less SSH/SCP</b> (menu item 2)	
	a) Password for <code>root</code> on all hosts:	Type the password for <code>root</code> on all hosts and press <b>Enter</b> .
	b) Continue to next step.	Press any key to continue.
5.	<b>Copy <code>/etc/hosts</code> to All Hosts</b> (menu item 3)	
	a) Continue to next step.	Press any key to continue.
6.	<b>Show <code>uname -a</code> for All Hosts</b> (menu item 4)	<i>Note:</i> Refer to the <i>Intel® Omni-Path Fabric Switches Release Notes</i> to ensure that any prerequisites for the upgrade to the new firmware level have been met prior to performing the upgrade through FastFabric.
	a) Continue to next step.	Press any key to continue.
7.	<b>Install/Upgrade OPA Software</b> (menu item 5)	<i>Note:</i> 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 <a href="#">Upgrade the Software on the Remaining Servers</a> on page 126.
continued...		



Step	Task/Prompt	Action
	a) Do you want to use ./IntelOPA-[Basic IFS].DISTRO.VERSION.tgz? [y]:	Press <b>Enter</b> to accept the default.
	b) Would you like to do a fresh [i]ninstall, an [u]pgrade or [s]kip this step? [u]:	Type <b>i</b> and press <b>Enter</b> .
	c) Are you sure you want to proceed? [n]:	Type <b>y</b> and press <b>Enter</b> .
	d) Complete the installation.	Press any key to continue.
	• If all hosts install...	Press any key to continue.
	• If any hosts fails to install...	Use the <b>View opahostadmin Results Files</b> menu item to review the result files from the update.
8.	<b>Configure IPoIB IP Address</b> (menu item 6)	
	a) Continue to next step.	Press any key to continue.
9.	<b>Build Test Apps and Copy to Hosts</b> (menu item 7)	
	a) For each prompt...	Provide your selections and press <b>Enter</b> .
	b) Continue to next step.	Press any key to continue.
10.	<b>Reboot Hosts</b> (menu item 8)	
	a) Continue to next step.	Press any key to continue.
11.	Ensure the hosts fully reboot, as verified through ping over the management network.	Perform Step 3.
<b>Optional Tasks</b> Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more information.		
11.	<b>Refresh SSH Known Hosts</b> (menu item 9)	This menu item refreshes 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.
12.	<b>Rebuild MPI Library and Tools</b> (menu item a)	This menu item prompts you for selection of which MPI to rebuild, and provides choices as to which available compiler to use.
13.	<b>Run a Command on All Hosts</b> (menu item b)	Intel recommends that you run the <code>date</code> command on all hosts to verify that the date and time are consistent. If needed, use the <b>Copy a File to All Hosts</b> menu item to copy the appropriate files to all hosts to enable and configure NTP.
14.	<b>Copy a File to All Hosts</b> (menu item c)	A file on the local host may be specified to be copied to all selected hosts.
<b>Verifying the Host Configuration</b>		
15.	<b>View opahostadmin Result Files</b> (menu item d)	
	a) About to: <code>vi /root/test.res /root/test.log</code>	Press any key to review files.
	b) Would you like to remove test.res test.log test_tmp* and save_tmp in /root ? [n]:	Press <b>Enter</b> to save or type <b>y</b> to remove the files.
	<b>End Task</b>	

### Next Steps

- To verify the host software is installed and running on the remaining servers, go to [Verify the Host Software on the Remaining Servers Using the FastFabric TUI Menu](#).



## 9.3 Verify the Host Software on the Remaining Servers Using the FastFabric TUI Menu

You can verify the host software using the FastFabric OPA Host Verification/Admin menu.

**Note:** 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.

### Assumption

- You are logged into the management node.

### Procedures

The following steps provide simplified instructions for first-time verification of the hosts. For additional details, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

Step	Task/Prompt	Action
1.	Access the <b>FastFabric OPA Host Verification/Admin Menu</b> .	
	a) If you are not already logged into <b>Intel FastFabric OPA Tools</b> , at the command prompt...	Type <code>opafastfabric</code> and press <b>Enter</b> .
	b) Access the <b>OPA Host Verification/Admin Menu</b> .	Press <b>4</b> .
	c) Select menu items.	Select items <b>0 – 8</b> .
	d) Start the operations.	Press <b>P</b> . <b>NOTE:</b> Each selected item is preformed in the order of the menu list.
2.	<b>Edit Configuration and Select/Edit Host File</b> (menu item 0)	
	a) Edit the <code>opafastfabric.conf</code> file.	Review the file with a focus on the following: <ul style="list-style-type: none"><li>FF_TOPOLOGY_FILE</li><li>FF_IPOIB_SUFFIX</li><li>FF_DEVIATION_ARGS</li><li>ff_host_basename_to_ipoib</li><li>ff_host_basename</li></ul> Review the following parameters which are used for overall fabric health checks: <ul style="list-style-type: none"><li>FF_ANALYSIS_DIR</li><li>FF_ALL_ANALYSIS</li><li>FF_FABRIC_HEALTH</li><li>FF_CHASSIS_CMDS</li><li>FF_CHASSIS_HEALTH</li><li>FF_ESM_CMDS</li></ul> FF_ALL_ANALYSIS should be updated to reflect the type of SM (esm or hostsm). If you made any changes, save and close the file. Press any key to continue.
	b) Edit the <code>ports</code> configuration file.	Review the file. If you made any changes, save and close the file.
continued...		



Step	Task/Prompt	Action
		Press any key to continue.
	b) Create or edit the <code>allhosts</code> configuration file.	Create the file with the Management Node's hosts name (the TCP/IP management network name, for example <code>mgmthost</code> ) and include the <code>hosts</code> file.  <i>Note:</i> If you have a cluster with mixed servers or HFI configurations, create a <code>/etc/opa/*hosts</code> file for each type of server configuration  If you made any changes, save and close the file.
	c) Do you want to edit/review/change the files? [y]:	Type <b>n</b> and Press <b>Enter</b> .
	e) Continue to next step.	Press any key to continue.
3.	<b>Summary of Fabric Components</b> (menu item 1)	
	a) After the operation completes...	Review the results against the expected configuration of the cluster.  <b>NOTE:</b> If components are missing, or degraded or omitted links are found, they should be corrected. Subsequent steps aid in locating any such links.
	b) Continue to next step.	Press any key to continue.
4.	<b>Verify Hosts Pingable, SSHable and Active</b> (menu item 2)	
	a) For each prompt...	Provide the required information and press <b>Enter</b> .
	b) After completion of the tests, you are prompted: Would you like to now use <code>/etc/opa/good</code> as Host File? [y]:	Press <b>Enter</b> to use the file or <b>n</b> to discard the file.
5.	<b>Perform Single Host Verification</b> (menu item 3)	
	a) Would you like to edit <code>/root/hostverify.sh</code> and copy to hosts? [y]:	Review the settings near the top and the list of TESTS selected.  <i>Note:</i> If you have a cluster with mixed servers or HFI configurations, ensure you add the proper settings for the server configuration (HFI PCIe bus, server memory size, expected single node HPL performance for server, etc) to the <code>/root/hostverify.sh</code> file  If you made any changes, save and close the file.
	b) For each prompt...	Provide the required information and press <b>Enter</b> .
	c) Start the tests.	Press any key to continue.
	d) Review the results file.	Press any key to view the file. Close the file.
	e) Repeat the <b>Perform Single Host Verification</b> for each of the host files.	<i>Note:</i> If you have a cluster with mixed servers or HFI configurations, Step 5 needs to be repeated for each <code>*host</code> file
	f) Continue to next step.	Press any key to continue.
6.	<b>Verify OPA Fabric Status and Topology</b> (menu item 4)	
	a) For each prompt...	Provide the required information and press <b>Enter</b> .
	a) Continue to next step.	Press any key to continue.
7.	<b>Verify Hosts See Each Other</b> (menu item 5)	
<b>continued...</b>		



Step	Task/Prompt	Action
	a) Continue to next step.	Press any key to continue.
8.	<b>Verify Hosts Ping via IPoIB</b> (menu item 6)	
	a) Continue to next step.	Press any key to continue.
9.	<b>Refresh SSH Known Hosts</b> (menu item 7)	
	a) Continue to next step.	Press any key to continue.
10.	<b>Check MPI Performance</b> (menu item 8)	<b>NOTE:</b> 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 intentionally uses techniques to reduce test runtime.
	a) For each prompt...	Press <b>Enter</b> to select the defaults.
	• If all hosts pass...	Continue to the next step.
	• If any hosts fail...	<ul style="list-style-type: none"> <li>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.</li> <li>Also verify that the HFI and any riser cards are properly seated.</li> </ul> Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more information.
	b) Continue to next step.	Press any key to continue.
<b>Optional Tasks</b>		
Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more information.		
11.	<b>Check Overall Fabric Health</b> (menu item 9)	This command 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 <a href="#">Configure and Initialize Health Check Tools Using FastFabric CLI Commands</a> on page 90 for more information.
12.	<b>Start or Stop Bit Error Rate Cable Test</b> (menu item a)	This command 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.
13.	<b>Generate All Hosts Problem Report Info</b> (menu item b)	This command collects configuration and status information from all hosts and generates a single *.tgz file that can be sent to an Intel support representative.
14.	<b>Run a Command on All Hosts</b> (menu item c)	This command runs the <code>cmdall</code> 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.
<b>Review the Host Verification</b>		
15.	<b>View opahostadmin Result Files</b> (menu item d)	A file on the local host may be specified to be copied to all selected hosts.
<i>continued...</i>		



Step	Task/Prompt	Action
	a) About to: vi /root/test.res /root/test.log	Press any key to review files.
	b) Would you like to remove test.res test.log test_tmp* and save_tmp in /root ? [n]:	Press <b>Enter</b> to save or type <b>y</b> to remove the files.
	<b>End Task</b>	

### Next Steps

- To set up and configure the management software for additional management nodes, go to [Set Up Additional Management Nodes](#).

## 9.4 Install the Host Software on the Remaining Hosts Using the FastFabric CLI Commands

You can configure the host software using the FastFabric OPA Host Setup menu.

### Assumption

- You are logged into the management node.

### Procedures

The following steps provide simplified instructions for first-time configuration of the hosts. For additional details, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

Step	Task/Prompt	Action
1.	(Optional) Edit or review the configuration files.	
	a) Edit the <code>/etc/opa/opafastfabric.conf</code> file.	Review the file with a focus on the following: <ul style="list-style-type: none"> <li>FF_IPOIB_SUFFIX</li> <li>FF_IPOIB_NETMASK</li> <li>FF_IPOIB_CONFIG</li> <li>FF_PRODUCT</li> <li>FF_PACKAGES</li> <li>FF_INSTALL_OPTIONS</li> <li>FF_UPGRADE_OPTIONS</li> </ul> If you made any changes, save and close the file.
	b) Create or edit the <code>/etc/opa/hosts</code> configuration file.	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. If you made any changes, save and close the file.
2.	Verify the hosts are pingable over the management network.	Type <code>opapingall -p</code> .
	<ul style="list-style-type: none"> <li>If all hosts were found...</li> </ul>	Continue to the next step.
	<ul style="list-style-type: none"> <li>If some hosts were not found...</li> </ul>	Review the following list for those hosts that were not found: <ul style="list-style-type: none"> <li>Host powered on and booted?</li> <li>Host connected to management network?</li> <li>Host management network IP address and network settings consistent with DNS or <code>/etc/hosts</code>?</li> <li>Management node connected to the management network?</li> </ul>

*continued...*



Step	Task/Prompt	Action
		<ul style="list-style-type: none"> <li>Management node IP address and network settings correct?</li> <li>Management network itself up (including switches, routers, and others)?</li> <li>Correct set of hosts listed in the hosts file? You may need to repeat the previous step to review and edit the file.</li> </ul>
3.	Set up secure password-less SSH, such that the Management Node can securely log into all the hosts as <code>root</code> through the management network, without requiring a password.	Type <code>opasetupssh -S -p -i "" -f hostfile</code> .
4.	(Optional) Copy the <code>/etc/hosts</code> file on this host to all the other selected hosts. <b>NOTE:</b> If DNS is being used, skip this step.	Type <code>opascpall -p -f hostfile /etc/hosts /etc/hosts</code> .
5.	(Optional) Copy the <code>/etc/resolv.conf</code> file on this host to all the other selected hosts.	Type <code>opascpall -p -f hostfile /etc/resolv.conf /etc/resolv.conf</code> .
6.	Show <code>uname -a</code> (OS version) on all the hosts.	Type <code>opacmdall -T 60 -f hostfile 'uname -a'</code> .
7.	Install the Intel® Omni-Path Fabric Host Software on all the hosts. <i>Note:</i> 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 <a href="#">Upgrade the Software on the Remaining Servers</a> on page 126.	Type <code>opahostadmin -f hostfile -d dir load</code> . By default, it looks in the current directory for the <code>IntelOPA-Basic.DISTRO.VERSION.tgz</code> file.
	• If all hosts install...	Continue to the next step.
	• If any hosts fails to install...	Use the <b>View opahostadmin</b> result files option to review the result files from the update.
8.	Configure IPoIB IP Address. This creates the <code>ifcfg-ib0</code> files on each host.	Type <code>opahostadmin -f hostfile configipoib</code> . <b>NOTE:</b> 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, <code>/etc/hosts</code> on the given host is checked.
9.	Build the MPI and/or SHMEM sample applications on the Management Node and copy the resulting object files to all the hosts. <i>Note:</i> This is in preparation for execution of MPI and/or SHMEM performance tests and benchmarks in a later step. <i>Note:</i> This option is only available when using the Intel® Omni-Path Fabric Host Software packaging of OFA (OFA Delta).	Type <code>MPICH_PREFIX=path_to_mpi</code> <code>cd /usr/src/opa/mpi_apps; make clobber quick</code> <code>opascpall -t -p -f hostfile source_dir dest_dir</code> .
10.	Reboot all the selected hosts.	Type <code>opahostadmin -f hostfile reboot</code> .
11.	Ensure the hosts fully reboot, as verified through ping over the management network.	Perform Step 2.
<b>Verifying the Host Configuration</b>		
15.	View opahostadmin result files.	Type <code>editor result_dir/result_file</code> . <b>NOTE:</b> In the line above, " <code>editor</code> " indicates the command line editor; for example, <code>vi</code> . The following default files are created: <ul style="list-style-type: none"> <li><code>test.res</code></li> </ul>
continued...		





Step	Task/Prompt	Action
		<ul style="list-style-type: none"> <li>test.log</li> </ul>
	<b>End Task</b>	

### Next Steps

- To verify the host software is installed and running on the remaining servers, go to [Verify the Host Software on the Remaining Servers Using CLI Commands](#).

## 9.5 Verify the Host Software on the Remaining Servers Using CLI Commands

You can verify the host software using the FastFabric OPA Host Verification/Admin menu.

**Note:** 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.

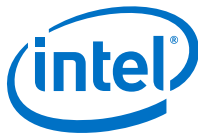
### Assumption

- You are logged into the management node.

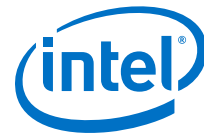
### Procedures

The following steps provide simplified instructions for first-time verification of the hosts. For additional details, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

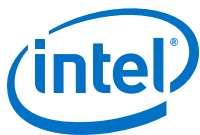
Step	Task/Prompt	Action
1.	(Optional) Edit or review the configuration files.	
	a) Edit the <code>/etc/opa/opafastfabric.conf</code> file.	Review the file with a focus on the following: <ul style="list-style-type: none"> <li>FF_TOPOLOGY_FILE</li> <li>FF_IPOIB_SUFFIX</li> <li>FF_DEVIATION_ARGS</li> <li>ff_host_basename_to_ipoib</li> <li>ff_host_basename</li> </ul> Review the following parameters which are used for overall fabric health checks: <ul style="list-style-type: none"> <li>FF_ANALYSIS_DIR</li> <li>FF_ALL_ANALYSIS</li> <li>FF_FABRIC_HEALTH</li> <li>FF_CHASSIS_CMDS</li> <li>FF_CHASSIS_HEALTH</li> <li>FF_ESM_CMDS</li> </ul> FF_ALL_ANALYSIS should be updated to reflect the type of SM (esm or hostsm). If you made any changes, save and close the file. Press any key to continue.
	b) Edit the <code>/etc/opa/ports</code> configuration file.	Review the file. If you made any changes, save and close the file. Press any key to continue.
<b>continued...</b>		



Step	Task/Prompt	Action
	b) Create or edit the <code>/etc/opa/allhosts</code> configuration file.	Create the file with the Management Node's hosts name (the TCP/IP management network name, for example <code>mgmthost</code> ) and include the <code>hosts</code> file.  <i>Note:</i> If you have a cluster with mixed servers or HFI configurations, create a <code>/etc/opa/*hosts</code> file for each type of server configuration  If you made any changes, save and close the file.
2.	Provide a brief summary of the counts of components in the fabric, including how many switch chips, hosts, and links are in the fabric.	Type <code>opafabricinfo</code> .
	a) After the operation completes...	Review the results against the expected configuration of the cluster.  <b>NOTE:</b> If components are missing, or degraded or omitted links are found, they should be corrected. Subsequent steps aid in locating any such links.
3.	(Optional) Verify each host is pingable.	Type <code>opapingall -p -f hostfile</code> .
	• If all hosts were found...	Continue to the next step.
	• If some hosts were not found...	Review the following list for those hosts that were not found: <ul style="list-style-type: none"> <li>• Host powered on and booted?</li> <li>• Host connected to management network?</li> <li>• Host management network IP address and network settings consistent with DNS or <code>/etc/hosts</code>?</li> <li>• Management node connected to the management network?</li> <li>• Management node IP address and network settings correct?</li> <li>• Management network itself up (including switches, routers, and others)?</li> <li>• Correct set of hosts listed in the hosts file? You may need to repeat the previous step to review and edit the file.</li> </ul>
4.	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.	Type <code>opafindgood -R -A -Q -f hostfile</code> .
	a) Review the <code>opasorthosts</code> files.	The following files are created in <code>opasorthosts_order</code> with all duplicates removed in the <code>OPA_CONFIG_DIR/</code> directory: <ul style="list-style-type: none"> <li>• good</li> <li>• alive</li> <li>• running</li> <li>• active</li> <li>• bad</li> <li>• quarantined</li> </ul>
5.	Perform a single host test on all hosts.	Type <code>opaverifyhosts -k -c -u hostverify.res -T timelimit -f hostfile test</code> .  <i>Note:</i> If you have a cluster with mixed servers or HFI configurations, ensure you add the proper settings for the server configuration (HFI PCIe bus, server memory size, expected single node HPL performance for server, etc) to the <code>/root/hostverify.sh</code> file.  Repeat the this step for each of the <code>*host</code> files.
continued...		



Step	Task/Prompt	Action
6.	Verify OPA Fabric status and topology.	Type <code>opalinkanalysis -U -x snapshot_suffix all verifyall &gt; \$FF_RESULT_DIR/linkanalysis.res 2&gt;&amp;1</code> . <b>NOTE:</b> The results can be seen in the <code>\$FF_RESULT_DIR/linkanalysis.res</code> file. A punch list of issues is appended to the <code>\$FF_RESULT_DIR/punchlist.csv</code> file.
	a) (Optional) To clear error counters after generating the report...	Add <code>clearerrors</code> and optionally <code>clearhwerrors</code> options to the <code>opalinkanalysis</code> run. <i>Note:</i> Clearing of hardware counters ( <code>-A</code> option) is optional and may affect the PM and other tools. See "PM Running Counters to Support <code>opareport</code> " section in the <i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i> for more information.
7.	Verify that each host can see all the others through queries to the Subnet Administrator.	Type <code>opahostadmin -f hostfile sacache</code> .
8.	Verify that IPoIB is properly configured and running on all the hosts.	Type <code>opahostadmin -f hostfile ipoibping</code> .
	• If successful...	Continue to next step.
	• If not successful...	Verify that the management host has IPoIB configured.
9.	Refresh the SSH <code>known_hosts</code> file on the Management Node to include the IPoIB hostnames of all the hosts.	Type <code>opasetupssh -p -U -f hostfile</code> .
10.	Perform a quick check of PCIe and MPI performance through end-to-end latency and bandwidth tests. <b>NOTE:</b> 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 intentionally uses techniques to reduce test runtime.	Type <code>opahostadmin -f hostfile mpiperfdeviation</code> .
	• If all hosts pass...	Continue to the next step.
	• If any hosts fail...	<ul style="list-style-type: none"> <li>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.</li> <li>Also verify that the HFI and any riser cards are properly seated.</li> </ul> Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more information.
<b>Optional Tasks</b>		
11.	Baseline the present fabric configuration for use in future fabric health checks. <b>NOTE:</b> This should be performed after configuring any additional Management Nodes.	Type <code>opaallanalysis -b</code> .
12.	Perform host and/or ISL cable testing. <b>NOTE:</b> The test allows for starting and stopping an extended Bit Error Rate test. The system prompts to clear hardware counters. <b>NOTE:</b> 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 <code>opatop</code> , <code>opareport</code> , and/or the Fabric Manager GUI. Once the desired test time has elapsed, stop the test using the command:	Type <code>opacabletest -C -f hostfile start</code> .
	a) To stop the test...	Type <code>opacabletest -f hostfile stop</code> .
<b>continued...</b>		



Step	Task/Prompt	Action
<b>Review the Host Verification</b>		
13.	View opahostadmin result files.	Type <i>editor result_dir/result_file</i> . <b>NOTE:</b> In the line above, "editor" indicates the command line editor; for example, vi. The following default files are created: <ul style="list-style-type: none"><li>• test.res</li><li>• test.log</li></ul>
	<b>End Task</b>	

### Next Steps

- To set up and configure the management software for additional management nodes, go to [Set Up Additional Management Nodes](#).



## 10.0 Set Up Additional Management Nodes

---

If the fabric has more than one Management Node, you need 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 procedures 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 106.

### 10.1 Before You Begin

Before starting the set up of the additional management nodes, perform the following:

- If required, upgrade the software to add additional IFS components using the procedure documented in [Upgrade from OPA-Basic to OPA-IFS Software Package](#) on page 128.
  - 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.

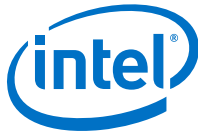
**Note:** Do not uninstall or replace existing configuration files that were previously created, especially IPoIB-related configuration files.

### 10.2 Configure Additional Management Nodes Using FastFabric TUI Menus

You can configure additional management nodes using the FastFabric TUI menus.

#### Assumption

- You are logged into the target Management Node.



## Procedures

Repeat the following steps on *each* additional Management Node.

Step	Task/Prompt	Action
1.	Copy the FastFabric configuration files from the initial Management Node to each new management node.	At a minimum, the following files should be copied:  <code>/etc/opa/opafastfabric.conf</code> <code>/etc/opa/ports</code> <code>/etc/opa/topology*.xml</code> <code>/etc/opa/hosts</code> <code>/etc/opa/allhosts</code> <code>/etc/opa/switches</code> <code>/etc/opa/chassis</code>
2.	Log into the FastFabric TUI. At the prompt...	Type <b>opafastfabric</b> .
3.	From the main menu, access the <b>FastFabric OPA Host Setup menu</b> .	Press <b>3</b> .
	a) <b>Edit Configuration and Select/Edit Host File</b>	Select <b>0</b> and press <b>P</b> to start the operation.
	b) Edit the <code>hosts</code> file.	Edit the <code>hosts</code> file such that the file on each Management Node omits itself.
	c) Complete the task.	Press <b>X</b> to return to the main menu.
4.	Access the <b>OPA Host Verification/Admin Menu</b> .	Press <b>4</b> .
	a) <b>Edit Configuration and Select/Edit Host File</b>	Select <b>0</b> and press <b>P</b> to start the operation.
	b) Edit the <code>allhosts</code> file.	Edit the <code>allhosts</code> file such that the file on each Management Node specifies itself.
	c) Complete the task.	Press <b>X</b> to return to the main menu.
5.	If Fabric Manager is to be run on the new management nodes, copy the Fabric Manager configuration file ( <code>/etc/opa-fm/opafm.xml</code> ) from the initial Management Node to each new management node.	
	a) Edit the <code>opafm.xml</code> file on each Management Node as needed.	Refer to the <i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i> for more information on how to configure the Fabric Manager.
6.	Access the <b>FastFabric OPA Host Setup menu</b> .	Press <b>3</b> .
	a) <b>(Linux) Set Up Password-Less SSH/SCP</b>	Select <b>2</b> and press <b>P</b> to start the operation.
	b) Password for root on all hosts:	Type the password for root on all hosts and press <b>Enter</b> .
	c) Upon completing the task, you are returned to the <b>FastFabric OPA Host Setup menu</b> .	
	d) <b>Refresh SSH Known Hosts</b>	Select <b>9</b> and press <b>P</b> to start the operation.
	e) Complete the operation.	Press any key to end the operation.
	<b>End Task</b>	

## Next Steps

- To configure and run the health check tool, go to [Perform Initial Health Check](#).



## 10.3 Configure Additional Management Nodes Using FastFabric CLI Commands

You can configure additional management nodes using the FastFabric CLI commands.

### Assumption

- You are logged into the target Management Node.

### Procedures

Repeat the following steps on *each* additional Management Node.

Step	Task/Prompt	Action
1.	Copy the FastFabric configuration files from the initial Management Node to each new management node.	At a minimum, the following files should be copied:  <pre> /etc/opa/opafastfabric.conf /etc/opa/ports /etc/opa/topology*.xml /etc/opa/hosts /etc/opa/allhosts /etc/opa/switches /etc/opa/chassis </pre>
2.	Edit the <code>/etc/opa/hosts</code> configuration file.	Edit the <code>hosts</code> file such that the file on each Management Node omits itself.
3.	Edit the <code>/etc/opa/allhosts</code> configuration file.	Edit the <code>allhosts</code> file such that the file on each Management Node specifies itself.
4.	If Fabric Manager is to be run on the new management nodes, copy the Fabric Manager configuration file ( <code>/etc/opa-fm/opafm.xml</code> ) from the initial Management Node to each new management node.	
5.	Edit the <code>opafm.xml</code> file on each Management Node as needed.	Refer to the <i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i> for more information on how to configure the Fabric Manager.
6.	Set up secure password-less SSH, such that the Management Node can securely log into all the hosts as <code>root</code> through the management network, without requiring a password.	Type <code>opasetupssh -S -p -i "" -f hostfile</code> .
7.	Refresh the SSH <code>known_hosts</code> file on the Management Node to include the IPoIB hostnames of all the hosts.	Type <code>opasetupssh -p -U -f hostfile</code> .
	<b>End Task</b>	

### Next Steps

- To configure and run the health check tool, go to [Perform Initial Health Check](#).



## 11.0 Perform Initial Health Check

---

The Health Check tool performs the initial fabric verification and creates a baseline of the hardware and software configuration. Once a good baseline has been established, you use the tools to compare the present fabric against the baseline and check its health. Baselines are rerun when changes occur such as fabric upgrades, hardware replacements or changes and software configuration changes.

### 11.1 Before You Begin

Before starting the health check configuration, perform the following:

- Be sure you have the latest *Intel® Omni-Path Fabric Switches Release Notes* for reference.
- Gather your information for the configuration files, as needed.

Note that you can edit the files before you configure the health check. Configuration files are located under the `/etc/opa` directory. Sample files are installed into `/usr/share/opa/samples` with the suffix `-sample`.

The following files are used to configure the health check:

- `opafastfabric.conf`: Lists the default settings for most of the FastFabric command line options.

Review and update the following parameters as needed:

- `FF_ANALYSIS_DIR`

This parameter should be updated to reflect the type of SM (esm or hsm).

- `FF_ALL_ANALYSIS`

**NOTE:** If you are running in a back-to-back configuration, this parameter should not contain `chassis`.

- `FF_FABRIC_HEALTH`
- `FF_CHASSIS_CMDS`
- `FF_CHASSIS_HEALTH`
- `FF_ESM_CMDS`

- `esm_chassis`: Lists the chassis names, using the assigned TCP/IP Ethernet management port names or IP addresses, that are running SMs if using Embedded SM(s) in the Intel® Omni-Path Fabric Chassis.

Intel recommends you use chassis names (the TCP/IP Ethernet management port names assigned). Enter one chassis name or IP address per line.

For example:

```
Chassis1
Chassis2
```





For more details about configuration files, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

For more details about the file format of the configuration files, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

## 11.2 Configure and Initialize Health Check Tools Using FastFabric TUI Menu

The health check tools may be run on one or more Management Nodes within the cluster. You set up and use the Intel® Omni-Path Fabric Suite FastFabric health check tools using the FastFabric TUI Menu.

For more information about health check tools, see the detailed discussion in the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

### Assumption

- You are logged into the target management node.

### Procedures

Repeat the following steps on *each* Management Node that will use the health check tools.

Step	Task/Prompt	Action
1.	Create the <code>/etc/opa/esm_chassis</code> file.	Create the file with a list of the chassis names using the assigned TCP/IP Ethernet management port names.
2.	Access the <b>FastFabric OPA Host Verification/Admin Menu</b> .	
	a) If you are not already logged into <b>Intel FastFabric OPA Tools</b> , at the command prompt...	Type <code>opafastfabric</code> and press <b>Enter</b> .
	b) Access the <b>OPA Host Verification/Admin Menu</b> .	Press <b>4</b> .
2.	<b>Edit Configuration and Select/Edit Host File</b> (menu item 0)	Select item <b>0</b> and press <b>P</b> .
	a) Edit <code>opafastfabric.conf</code> file.	Review the following parameters: <ul style="list-style-type: none"> <li>FF_ANALYSIS_DIR This parameter should be updated to reflect the type of SM (esm or hsm).</li> <li>FF_ALL_ANALYSIS <b>NOTE:</b> If you are running in a back-to-back configuration, this parameter should not contain chassis.</li> <li>FF_FABRIC_HEALTH</li> <li>FF_CHASSIS_CMDS</li> <li>FF_CHASSIS_HEALTH</li> <li>FF_ESM_CMDS</li> </ul>
3.	<b>Check Overall Fabric Health</b> (menu item 9)	Select item <b>9</b> and press <b>P</b> .
	a) Performing Host Admin: Check Overall Fabric Health Baseline present configuration? [n]:	Press <b>Enter</b> .
4.	Check the results.	
	a) If no errors were encountered...	Continue to next step.
<i>continued...</i>		



Step	Task/Prompt	Action
	b) If any errors are encountered...	Perform the following: 1. Resolve the errors. 2. Rerun 3. and 4. until a clean run occurs.
5.	Create a cluster configuration baseline.	
	a) <b>Check Overall Fabric Health</b> (menu item 9)	Select item <b>9</b> and press <b>P</b> .
	b) Performing Host Admin: Check Overall Fabric Health Baseline present configuration? [n]:	Press <b>y</b> and press <b>Enter</b> .
6.	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 <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more information about opaallanalysis and its automated use.
	<b>End Task</b>	

### Next Steps

- Before you run top500 HPL (High Performance Linpack) runs or customer acceptance tests, Intel recommends that you follow all steps outlined in the *Intel® Omni-Path Fabric Setup Guide*.
- To run top500 HPL2, go to [Run a Sample High Performance Linpack 2 Configuration](#).

## 11.3 Configure and Initialize Health Check Tools Using FastFabric CLI Commands

The health check tools may be run on one or more Management Nodes within the cluster. You set up and use the Intel® Omni-Path Fabric Suite FastFabric health check tools using CLI commands.

For more information, see the detailed discussion in the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

### Assumption

- You are logged into the target management node.

### Procedures

Repeat the following steps on *each* Management Node that will use the health check tools.

Step	Task/Prompt	Action
1.	Edit /etc/opa/opafastfabric.conf file.	Review the following parameters: <ul style="list-style-type: none"><li>• FF_ANALYSIS_DIR This parameter should be updated to reflect the type of SM (esm or hsm).</li><li>• FF_ALL_ANALYSIS <b>NOTE:</b> If you are running in a back-to-back configuration, this parameter should not contain chassis.</li><li>• FF_FABRIC_HEALTH</li></ul>
continued...		



Step	Task/Prompt	Action
		<ul style="list-style-type: none"> <li>FF_CHASSIS_CMDS</li> <li>FF_CHASSIS_HEALTH</li> <li>FF_ESM_CMDS</li> </ul>
2.	Create the <code>/etc/opa/esm_chassis</code> file.	Create the file with a list of the chassis names using the assigned TCP/IP Ethernet management port names.
3.	Perform a health check.	Type <b>opaallanalysis -e</b> .
4.	Check the results.	
	a) If no errors were encountered...	Continue to next step.
	b) If any errors are encountered...	Perform the following: <ol style="list-style-type: none"> <li>1. Resolve the errors.</li> <li>2. Rerun <a href="#">3.</a> and <a href="#">4.</a> until a clean run occurs.</li> </ol>
5.	Create a cluster configuration baseline. <i>Note:</i> This may also be done using the FastFabric TUI menu by selecting <b>Check Overall Fabric Health</b> and answering <b>y</b> to the question: <code>Baseline present configuration? [n]:</code>	Type <b>opaallanalysis -b</b> .
6.	If required, schedule regular runs of <code>opaallanalysis</code> through cron or other mechanisms.	Refer to the Linux* OS documentation for more information on cron. Also refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more information about <code>opaallanalysis</code> and its automated use.
	<b>End Task</b>	

### Next Steps

- Before you run top500 HPL (High Performance Linpack) runs or customer acceptance tests, Intel recommends that you follow all steps outlined in the *Intel® Omni-Path Fabric Setup Guide*.
- To run top500 HPL2, go to [Run a Sample High Performance Linpack 2 Configuration](#).



## 12.0 Perform High Performance Linpack 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.

You run the initial HPL2 benchmark to gather preliminary baseline numbers. The sample configurations provided in the `/usr/src/opa/mpi_apps/hpl-config` folder should perform within 10 – 20% of optimal HPL2 results for the cluster. For further performance tuning, refer to the *Intel® Omni-Path Fabric Performance Tuning User Guide*.

The basic HPL2 process is:

1. **Perform the initial run.**

This should be 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.

2. **Perform full scale runs based on your cluster size using sample or customized configurations.**

Assorted sample `HPL.dat` files are provided in the `/usr/src/opa/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 two-processor node, 2 GB of node memory is assumed. For each cluster size, four 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`.

Customized files can be generated using `hpl_dat_gen`.

3. **Review the results.**

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/src/opa/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 User Guide* for more information on `scpall`.

## 12.1 Before You Begin

Before starting the High Performance Linpack benchmark, perform the following:

- Ensure that you have followed all steps outlined in the *Intel® Omni-Path Fabric Setup Guide*.
- Create the file `/usr/src/opa/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.
- Determine the HPL2 runs in advance, as needed.
  - Select an initial, small configuration to verify HPL has been properly compiled.
  - Select one or more full scale configurations based on the cluster size and problem size.

*Notes:* You will need to the following commands to perform the runs:

- Select configuration command usage:  
`./config_hpl2 config_name[problem_size]`
- Run command usage:  
`run_hpl number_processes`

For example, `./config_hpl2 256s` selects the configuration and, subsequently, `./run_hpl2 256` runs the configuration.

## 12.2 Run a Sample High Performance Linpack 2 Configuration

You can run sample HPL2 configurations to gather baseline numbers based on pre-configured `HPL.dat` files found in the `/usr/src/opa/mpi_apps/hpl-config` folder.

### Assumptions

- You are logged into the host.
- You have determined the configurations you will run.

### Procedure

*Note:* It is best to start with a small configuration to verify HPL has been properly compiled, then move on to full scale HPL2 runs.

**(Host)** To run a sample HPL2 configuration, perform the following:



Step	Task/Prompt	Action
1.	Change directories to the <code>mpi_apps</code> folder.	Type <code>cd /usr/src/opa/mpi_apps</code>
2.	If you do not know which configuration file you will use, view the sample configuration files.	Type <code>./config_hpl2</code> .
3.	Select an initial run with a very small problem size to determine if the run is successful.	Type <code>./config_hpl2 &lt;number of processes&gt;&lt;problem size&gt;</code> . For example, <code>./config_hpl2 32s</code>
4.	Run the configuration.	Type <code>./run_hpl2 &lt;number of processes&gt;&lt;problem size&gt;</code> For example, <code>./run_hpl2 32</code>
5.	Review the results. <b>NOTE:</b> Performance of this run is expected to be low.	
	• If the run is successful...	Continue to the next step.
	• If it is not successful...	Troubleshoot.
6.	View the sample configuration files to find a configuration that simulates your cluster.	Type <code>./config_hpl2</code> .
7.	Select a full scale run to that simulates your cluster.	Type <code>./config_hpl2 &lt;number of processes&gt;&lt;problem size&gt;</code> . For example, <code>./config_hpl2 9216m</code>
8.	Run the configuration.	Type <code>./run_hpl2 &lt;number of processes&gt;&lt;problem size&gt;</code> For example, <code>./run_hpl2 9216</code>
9.	Review the results.	
	<b>End Task</b>	

### Next Steps

- To install and start up the Fabric Manager GUI software, refer to [Install Intel® Omni-Path Fabric Suite Fabric Manager GUI](#).

## 12.3 Generate and Run a Custom High Performance Linpack Configuration

To generate a custom HPL2 configuration, you use `hpl_dat_gen` to probe 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.

### Assumptions

- You are logged into the host.
- You have determined the configurations you will run.

### Procedures

**(Host)** To generate and run a custom HPL2 configuration, perform the following:



Step	Task/Prompt	Action
1.	Change directories to the <code>mpi_apps</code> folder.	Type <code>cd /usr/src/opa/mpi_apps</code>
2.	Start the dat generator tool.	Type <code>./hpl_dat_gen.</code>
3.	Provide the following information:	
	a) # of compute nodes [1]?	Type a number and press <b>Enter</b> .
	b) # of cores per node [72]?	Type a number and press <b>Enter</b> .
	c) # of RAM per node (in MB) [64156]?	Type a number and press <b>Enter</b> .
	d) Memory pressure (range between 0.1 and 0.9) [0.3]?	Type a decimal number and press <b>Enter</b> .
	e) From the results, note the number of processes. This will be added to the run command: <code>./run_hpl2 &lt;number of processes&gt;</code> .  <pre> 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 (&gt;= 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 (&gt;=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 (&gt; 0) </pre>	Continue to the next step.
3.	Run the configuration.	Type <code>./run_hpl2 &lt;number of processes&gt;</code> . For example: <code>./run_hpl2 9216</code>
4.	Review the results.	
	<b>End Task</b>	

### Next Steps

- To install and start up the Fabric Manager GUI software, refer to [Install Intel® Omni-Path Fabric Suite Fabric Manager GUI](#).



## 13.0 Install Intel® Omni-Path Fabric Suite Fabric Manager GUI

---

Intel® Omni-Path Fabric Suite Fabric Manager GUI can be run on a Linux\* or Windows\* desktop/laptop system with TCP/IP connectivity to the Intel® Omni-Path Fabric Suite Fabric Manager. Network connectivity to the Fabric Executive (FE) component of the Fabric Manager allows continuous, remote, “out of band” monitoring. If configured with alternate FE connection information, the Fabric Manager GUI automatically fails over to an alternative FE if the connection to the original is lost.

This section provides information and procedures to install and set up the Fabric Manager GUI on your desktop or laptop.

### 13.1 Before You Begin

The following sub-sections provide information and instructions that are necessary before you start the Intel® Omni-Path Fabric Suite Fabric Manager GUI installation.

#### 13.1.1 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\* 10.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)

#### 13.1.2 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)

### 13.1.3 Download the Fabric Manager Software

1. Using a web browser, type `downloadcenter.intel.com` in the address field and press **Enter**.
2. In the "Search downloads" field, type `Fabric Manager GUI` and press **Enter**.
3. From the search drop-down, select the Intel® Omni-Path Fabric Suite Fabric Manager GUI software.
4. In the Available Downloads list, select the file(s) you need for the OS you have installed on your fabric.

*Note:* You can download either the executable file to install or the source files to build your Intel OPA Fabric Manager GUI jar file. This document only addresses installing through the executable.

5. Review the Intel Software License Agreement.
6. Click "I accept the terms in the license agreement."
7. Save the download to your hard drive.

## 13.2 Install the Fabric Manager GUI Software on Windows\*

You install the Fabric Manager GUI software on your Windows\* desktop using the installation wizard.

### Assumption

- You have downloaded the installation executable for Windows\* to your desktop.

### Procedure

To install the software, perform the following:

Step	Task/Prompt	Action
1.	Start the wizard.	Using File Explorer, navigate to the target folder and double-click the <code>IntelOPA-FMGUI.windows-x.x.x.x.exe</code> file on the desktop
<i>continued...</i>		



Step	Task/Prompt	Action
		where <code>x.x.x.x</code> is the version number of the Fabric Manager GUI application being installed.
	a) At the prompt: Do you want to allow the following program to make changes to this computer?	Click <b>Yes</b> .
	b) Installer Language...	Select the language to be used for the installation screens and then click <b>OK</b> .
	c) Fabric Manager GUI Setup dialog...	
	• If this is a first-time installation...	Click <b>Next</b> .
	• If Fabric Manager GUI is already installed on your system and you are re-installing...	Click <b>Yes</b> to overwrite the existing software and then click <b>Next</b> .
	d) After reviewing the Licence Agreement...	Select the checkbox "I accept the terms of the License Agreement" and click <b>Next</b> .
	e) Choose Install Location...	Intel recommends that you use the default file location. Click <b>Install</b> .
	f) Completing Fabric Manager GUI Setup...	Click <b>Finish</b> .
2.	(Optional) Change the database location.	
	a) Change directory.	Navigate to <code>&lt;app_data_path&gt;\Intel\FabricManagerGUI</code> where <code>&lt;app_data_path&gt;</code> is a user's application data folder. Under Windows 7/8, it's <code>C:\Users\&lt;user_name&gt;\AppData\Roaming</code>
	b) Create file <code>settings.xml</code> .	<ol style="list-style-type: none"> <li>1. Right-click in the folder and select <b>New &gt; Text Document</b>.</li> <li>2. Name the file <code>settings.xml</code>.</li> <li>3. Click <b>Yes</b> to change the extension.</li> </ol>
	c) Open the file in a text editor and add the contents...	<pre>&lt;?xml version="1.0" encoding="UTF-8" standalone="no"?&gt; &lt;!DOCTYPE properties SYSTEM "http://java.sun.com/dtd/properties.dtd"&gt; &lt;properties&gt;   &lt;comment&gt;FM GUI Application configuration   settings&lt;/comment&gt;   &lt;entry key="db.connection.url"&gt;jdbc:hsqldb:file:c:   \temp\db\mydbname;    hsqldb.result_max_memory_rows=1000;hsqldb.default_table   _type=cached;          hsqldb.log_size=50;&lt;/entry&gt; &lt;/properties&gt;</pre>
	Change <code>c:\temp\db\mydbname</code> to the location for the database file. <b>NOTE:</b> Only change the database file path. Omitting the other options will cause database performance issues in HSQLDB.	Save and close the file.
	<b>End Task</b>	

### Next Steps

- To prepare the Fabric Manager for Fabric Manager GUI, refer to [Prepare Fabric Manager for Fabric Manager GUI](#).



### 13.3 Install the Fabric Manager GUI Software on Linux\*

You install the Fabric Manager GUI software on your Linux\* desktop using the rpm file.

#### Assumption

- You have downloaded the installation rpm for Linux\* to your desktop.

#### Procedure

To install the software, perform the following:

Step	Task/Prompt	Action
1.	Install the rpm.	
	a) Log into the system where the Fabric Manager GUI will be installed.	Or become root by typing <b>su</b> and entering the super-user password.
	b) Open a Terminal window in X Windows.	
	c) Change directories to the directory where you downloaded the rpm file.	Type <b>cd <i>directory_path_name</i></b> .
	d) Install the package.	Type <b>rpm -iv IntelOPA-FMGUI-linux-x.x.x.x.noarch.rpm</b> where <i>x.x.x.x</i> is the version number of the Fabric Manager GUI application being installed.
	e) Wait for the successful completion of the command.	Continue to next step.
2.	(Optional) Change the database location.	
	a) Change directory.	Type <b>cd ~/.Intel\FabricManagerGUI</b> .
	b) Create file <b>settings.xml</b> .	Type <b>touch settings.xml</b>
	c) Using a text editor, add the following contents...	<pre>&lt;?xml version="1.0" encoding="UTF-8" standalone="no"?&gt; &lt;!DOCTYPE properties SYSTEM "http://java.sun.com/dtd/properties.dtd"&gt; &lt;properties&gt;   &lt;comment&gt;FM GUI Application configuration   settings&lt;/comment&gt;   &lt;entry     key="db.connection.url"&gt;jdbc:hsqldb:file:/var/temp/db/     mydbname;      hsqldb.result_max_memory_rows=1000;hsqldb.default_table     _type=cached;     hsqldb.log_size=50;&lt;/entry&gt; &lt;/properties&gt;</pre>
	d) Change <b>/var/temp/db/mydbname</b> to the location for the database file. <b>NOTE:</b> Only change the database file path. Omitting the other options will cause database performance issues in HSQLDB.	Save and close the file.
	<b>End Task</b>	

#### Next Steps

- To prepare the Fabric Manager for Fabric Manager GUI, refer to [Prepare Fabric Manager for Fabric Manager GUI](#).



## 13.4 Prepare Fabric Manager for Fabric Manager GUI

Fabric Manager GUI requires a running fabric with FE available. This task show you how to prepare the Fabric Manager to work with the Fabric Manager GUI.

### Assumptions

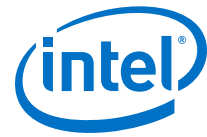
- You have completed the installation of the Fabric Manager GUI on your system.
- Before running Fabric Manager GUI, ensure that Fabric Manager has the proper configuration and is running normally.

### Procedure

To install the software, perform the following:

Step	Task/Prompt	Action
1.	On the management node, edit the <code>/etc/opa-fm/opa_fm.xml</code> file.	<b>NOTE:</b> The following steps contain the necessary edits to <code>opa_fm.xml</code> .
2.	Change the <code>&lt;SubnetPrefix&gt;</code> for <code>fm0</code> to <code>0xfe80000000001000</code>	<pre> &lt;!-- Shared Instance config, applies to all components: SM, PM and FE --&gt; &lt;Shared&gt;   &lt;!-- Fm.Shared.Start controls overall startup of the Instance. --&gt;   &lt;!-- If 0, none of the components in the Instance are started. --&gt;   &lt;!-- If 1, instance is enabled and Fm.Sm.Start, Fm.Pm.Start, etc --&gt;   &lt;!-- control startup of each manager. The default for each manager --&gt;   &lt;!-- is defined by Common.Sm.Start, Common.Pm.Start, etc --&gt;   &lt;!-- ESM does not support Start via XML configuration. Use CLI commands --&gt;   &lt;Start&gt;1&lt;/Start&gt;   &lt;!-- &lt;StartupRetries&gt;5&lt;/StartupRetries&gt; --&gt;   &lt;!-- &lt;StartupStableWait&gt;10&lt;/StartupStableWait&gt; --&gt;    &lt;!-- Name, Hfi, Port, and PortGUID are ignored for ESM since they --&gt;   &lt;!-- are automatically set --&gt;   &lt;Name&gt;fm0&lt;/Name&gt; &lt;!-- also for logging with _sm, _fe, _pm appended --&gt;   &lt;Hfi&gt;1&lt;/Hfi&gt; &lt;!-- local HFI to use for FM instance, 1=1st HFI --&gt;   &lt;Port&gt;1&lt;/Port&gt; &lt;!-- local HFI port to use for FM instance, 1=1st Port --&gt;   &lt;PortGUID&gt;0x0000000000000000&lt;/PortGUID&gt; &lt;!-- local port to use for FM --&gt;   &lt;SubnetPrefix&gt;0xfe80000000001000&lt;/SubnetPrefix&gt; &lt;!-- should be unique --&gt; </pre> <p>To set up multiple rails in a single or multi-subnet, refer to <a href="#">Multi-Rail Usage</a> on page 110 for additional instructions.</p>
3.	Ensure that the FE is enabled.	<p>In the <code>&lt;FE&gt;</code> section, "Start" is 1.</p> <pre> &lt;Fe&gt;   &lt;Start&gt;1&lt;/Start&gt; &lt;!-- default FE startup for all instances --&gt; </pre>
4.	Ensure that FE has proper security setup.	
	<ul style="list-style-type: none"> <li>If you want to run FE without SSL, ensure that the <code>SslSecurityEnable</code> is 0.</li> </ul> <p>Also, when you run Fabric Manager GUI, set up FE host as unsecured (clear the "Secure" checkbox).</p>	<pre> &lt;!-- OpenSSL FE network security parameters --&gt; &lt;SslSecurityEnable&gt;0&lt;/SslSecurityEnable&gt; </pre>

*continued...*



Step	Task/Prompt	Action
	<ul style="list-style-type: none"> <li>If you want to run FE with SSL, ensure that you set <code>SslSecurityEnable</code> to 1.</li> </ul> <p>To set up keys and <code>opafm.xml</code> properly, refer to additional steps in sections and appendices found in the <i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i>:</p> <ul style="list-style-type: none"> <li>"Out-of-Band Security"</li> <li>"Core-Level Public Key Infrastructure (PKI) Best Practices Guidelines"</li> <li>"Advanced-Level Public Key Infrastructure Best Practices Guidelines"</li> <li>"SSL Key Creation for Fabric Manager GUI"</li> </ul> <p>And, when you run Fabric Manager GUI, set up FE host as "Secure" with the key files (select the "Secure" checkbox).</p>	<pre>&lt;!-- OpenSSL FE network security parameters --&gt; &lt;SslSecurityEnable&gt;1&lt;/SslSecurityEnable&gt;</pre>
5.	Ensure that PM is enabled.	<p>In the <code>&lt;Pm&gt;</code> section, "Start" is 1.</p> <pre>&lt;Pm&gt;   &lt;!-- ESM does not support Start via XML   configuration. Use CLI commands --&gt;   &lt;Start&gt;1&lt;/Start&gt; &lt;!-- default PM startup for all   instances --&gt;</pre>
6.	If you made any changes...	Save and close the file.
7.	Restart the Fabric Manager so that your changes take effect.	Type <code>systemctl restart opafm</code> .
	<b>End Task</b>	

### Next Steps

- To start the Fabric Manager GUI on Windows\*, go to [Start the Fabric Manager GUI on Windows\\*](#).
- To start the Fabric Manager GUI on Linux\*, go to [Start the Fabric Manager GUI on Linux\\*](#).

## 13.5 Start the Fabric Manager GUI on Windows\*

You start the Intel® Omni-Path Fabric Suite Fabric Manager GUI on Windows\* through the Start menu.

**Note:** Intel recommends that you accept the default startup options (settings). Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager GUI Online Help* for procedures to set user preferences.

### Using Windows Vista and Windows 7

To start the application:

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

The Fabric Manager GUI application starts.



*Notes:* It is recommended that you create a shortcut to put an icon on your desktop.

- For Windows Vista, right-click the application name in step 2 and select **Send To > Desktop (Create Shortcut)**.
- For Windows 7, right-click the application name in step 2 and select "Pin to Start Menu".

### Using Windows 8.x

To start the application:

1. From the **Start** menu, select the Search (magnifying glass) icon.
2. In the search box, begin typing **Fabric Manager GUI**.  
The application will appear for selection.
3. Select the application.

The Fabric Manager GUI application starts.

*Note:* It is recommended that you pin the application to your Start menu. Right-click the application name in step 3 and select "Pin to Start".

## 13.6 Start the Fabric Manager GUI on Linux\*

You can start the Fabric Manager GUI on Linux\* from either a terminal window on X Windows or the K Desktop Environment (KDE).

*Note:* Intel recommends that you accept the default startup options (settings). Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager GUI Online Help* for procedures to set user preferences.

### Using X Windows

To start the application from a terminal window on X Windows:

1. Open a terminal window in X Windows.
2. Type **fmgui** and press **Enter**.

The Fabric Manager GUI starts.

### Using K Desktop Environment (KDE)

To start the 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 starts.

## 13.7 Uninstall Fabric Manager GUI

This section provides instructions and information for uninstalling the Fabric Manager GUI from your desktop. These instructions are provided for reference only.



**Note:** 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.

### 13.7.1 Uninstall Fabric Manager GUI Software on Windows\*

#### Using Windows\* Vista and Windows\* 7

To uninstall the software, perform the following:

1. From the **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.

#### Using Windows\* 8.x

To uninstall the software, perform the following:

1. Right-click the **Start** menu and select **Programs and Features**.
2. In the program list, select **Intel Fabric Manager GUI**.
3. Click the **Uninstall** button at the top of the list.
4. Follow the instructions on the uninstall window.

The Fabric Manager GUI is uninstalled.

#### 13.7.1.1 Clear the Fabric Manager GUI Cache on Windows\*

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.

#### Using Windows\* Vista and Windows\* 7

To clear the cache, perform the following:

1. Select **Start** menu.
2. Select **All Programs**.
3. Select **Intel**.
4. Select **Omni-Path**.
5. Select **Clear FM GUI Cache**.  
A command window is displayed.
6. Type **y** and press **Enter**.
7. Press any key to complete the process.



### Using Windows\* 8.x

To clear the cache, perform the following:

1. Select the Start button (window icon).
2. Click the Search (magnifying glass) icon.
3. Begin typing **Clear FM GUI Cache**.

The application will display in the results pane.

4. Select **Clear FM GUI Cache**.

A command window is displayed.

5. Type **y** and press **Enter**.
6. Press any key to complete the process.

## 13.7.2 Uninstall Fabric Manager GUI Software on Linux\*

To uninstall the software, perform the following:

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 `# 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.

### 13.7.2.1 Clear the Fabric Manager GUI Cache on Linux\*

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.

#### Using X Windows

To clear the cache, perform the following:

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.

#### Using K Desktop Environment (KDE)

To clear the cache, perform the following:

1. In KDE or GNOME, select **Applications**.





2. Select **Fabric**; if not available, try the **System** category.
3. Select **Clear FM GUI Cache**.

The Fabric Manager GUI cache is cleared.



## 14.0 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.

### 14.1 Multi-Subnet Fabrics

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

- **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 separately. This single-subnet scenario is created using the default FastFabric instructions previously described in this document.

- **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 [Setting Up Primarily Independent Subnets](#).

- **The subnets are overlapping.**

If multiple components are common to more than one subnet, such as FastFabric nodes, servers, and others, refer to the following instructions for [Setting Up Overlapping Subnets](#).

#### 14.1.1 Setting Up Primarily Independent Subnets

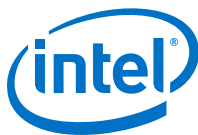
This section provides additional instructions performed for specific steps during the installation and configuration process. Use the table below as both instructions and a checklist for setting up your primarily independent subnets.

##### Procedures

Perform the installation and configuration as described in this guide. Append the following actions during the process.



Step	Task/Prompt	Action
1.	<b>Installation Prerequisites</b>	
	a) <a href="#">Fabric Design Prerequisites</a>	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.
	b) <a href="#">Fabric Setup Prerequisites</a>	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 <i>Intel® Omni-Path Fabric Host Software User Guide</i> for more information on configuring IPoIB.
2.	<b>Install the Intel® Omni-Path Software</b>	As stated.
3.	<b>Configure the Software</b>	
	a) <a href="#">Configure the Chassis</a>	<ul style="list-style-type: none"> <li>When creating the <code>chassis</code> 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.</li> <li>When editing the <code>ports</code> 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.</li> </ul>
	b) <a href="#">Configure the Externally-Managed Switches</a>	<ul style="list-style-type: none"> <li>When editing the <code>ports</code> 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.</li> <li>If required, the <code>switches</code> file may specify a <code>hfi:port</code> per switch. However, if <code>hfi:port</code> is not specified, all the <code>hfi:port</code> specified in the <code>ports</code> file are searched to locate the given switch's Node GUID.</li> </ul>
	c) <a href="#">Install the Host Software on the Remaining Hosts Using the FastFabric TUI Menu</a>	When creating the <code>hosts</code> 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.
	d) <a href="#">Verify the Host Software on the Remaining Servers Using the FastFabric TUI Menu</a>	<ul style="list-style-type: none"> <li><b>Edit Configuration and Select/Edit Host File</b> Create the <code>allhosts</code> 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 <code>ports</code> 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.</li> <li><b>Verify Hosts See Each Other</b> This can be run for each subnet by using the <code>allhosts</code> files specific to each subnet (i.e., those listing only hosts in a single subnet).</li> <li><b>Check MPI Performance</b> This can be run for each subnet by using the <code>allhosts</code> files specific to each subnet (i.e., those listing only the hosts in a single subnet).</li> </ul>
<b>continued...</b>		



Step	Task/Prompt	Action
	e) Configure Additional Management Nodes Using FastFabric TUI Menus	When copying FastFabric configuration files to the additional Fabric Management Nodes, be sure to also copy the additional <code>hosts</code> , <code>chassis</code> , and <code>allhosts</code> files that were created per subnet.  <i>Note:</i> 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 <code>opafabricanalysis</code> may indicate different port numbers or host files used for FastFabric, and MPI may need to list different hosts.
	f) Configure and Initialize Health Check Tools Using FastFabric CLI Commands	Make sure the <code>/etc/opa/ports</code> file lists each of the Fabric Management Node local HFIs and ports that are connected to a unique subnet. When running <code>opareports</code> , <code>opafabricinfo</code> , <code>opafabricanalysis</code> , or <code>opaallanalysis</code> , the default is to use the <code>ports</code> file. If required, the <code>-p</code> and <code>-t</code> options or the <code>PORTS/PORTS_FILE</code> 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 <code>esm_chassis</code> and <code>chassis</code> files used must list all relevant Intel chassis in all subnets.
	g) Run a Sample High Performance Linpack 2 Configuration	Run HPL2 for each subnet by creating <code>mpi_hosts</code> files specific to each subnet, that is, only listing hosts in a single subnet.
	h) Install Intel® Omni-Path Fabric Suite Fabric Manager GUI	At least one subnet manager is required per subnet. Refer to the <i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i> for more information on how to configure a host SM node to manage more than one subnet.
	<b>End Task</b>	

### 14.1.2 Setting Up Overlapping Subnets

This section provides additional instructions performed for specific steps during the installation and configuration process. Use the table below as both instructions and a checklist for setting up your overlapping subnets.

#### Procedures

Perform the installation and configuration as described in this guide. Append the following actions during the process.

Step	Task/Prompt	Action
1.	<b>Installation Prerequisites</b>	
	a) Fabric Design Prerequisites	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 required, between routers, IPoIB routers, and other devices.
<b>continued...</b>		



Step	Task/Prompt	Action
	b) <a href="#">Fabric Setup Prerequisites</a>	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 <i>Intel® Omni-Path Fabric Host Software User Guide</i> for more information on configuring IPoIB.
2.	<b>Install the Intel® Omni-Path Software</b>	As stated.
3.	<b>Configure the Software</b>	
	a) <a href="#">Configure the Chassis</a>	<ul style="list-style-type: none"> <li>When creating the <code>chassis</code> 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.</li> <li>When editing the <code>ports</code> 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.</li> </ul>
	b) <a href="#">Configure the Externally-Managed Switches</a>	<ul style="list-style-type: none"> <li>When editing the <code>ports</code> 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.</li> <li>If required, the <code>switches</code> file may specify a <code>hfi:port</code> per switch. However, if <code>hfi:port</code> is not specified, all the <code>hfi:port</code> entries specified in the <code>ports</code> file are searched to locate the given switch's node GUID.</li> </ul>
	c) <a href="#">Install the Host Software on the Remaining Hosts Using the FastFabric TUI Menu</a>	<ul style="list-style-type: none"> <li>When creating the <code>hosts</code> file, list 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.</li> <li>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.</li> <li>Refer to the <i>Intel® Omni-Path Fabric Host Software User Guide</i> for more information on configuring IPoIB.</li> </ul>
	d) <a href="#">Verify the Host Software on the Remaining Servers Using the FastFabric TUI Menu</a>	<ul style="list-style-type: none"> <li><b>Edit Configuration and Select/Edit Host File</b> Create the <code>allhosts</code> 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 <code>ports</code> 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.</li> <li><b>Verify Hosts See Each Other</b> This can be run for each subnet by using the <code>allhosts</code> files specific to each subnet (i.e., those listing only hosts in a single subnet).</li> <li><b>Check MPI Performance</b> This can be run for each subnet by using the <code>allhosts</code> files specific to each subnet (i.e., those listing only the hosts in a single subnet). This is currently not available for OFED.</li> </ul>
<b>continued...</b>		



Step	Task/Prompt	Action
	e) Configure Additional Management Nodes Using FastFabric TUI Menus	When copying FastFabric configuration files to the additional Fabric Management Nodes, be sure to also copy the additional <code>hosts</code> , <code>chassis</code> , and <code>allhosts</code> files that were created per subnet.  <i>Note:</i> 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 <code>opafabricanalysis</code> may indicate different port numbers or host files used for FastFabric, and MPI may need to list different hosts.
	f) Configure and Initialize Health Check Tools Using FastFabric CLI Commands	Make sure the <code>/etc/opa/ports</code> file lists each of the Fabric Management Node local HFIs and ports that are connected to a unique subnet. When running <code>opareports</code> , <code>opafabricinfo</code> , <code>opafabricanalysis</code> , or <code>opaallanalysis</code> , the default is to use the <code>ports</code> file. If required, the <code>-p</code> and <code>-t</code> options or the <code>PORTS/PORTS_FILE</code> 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 <code>esm_chassis</code> and <code>chassis</code> files used must list all relevant Intel chassis in all subnets.
	g) Run a Sample High Performance Linpack 2 Configuration	Run HPL2 for each subnet by creating <code>mpi_hosts</code> files specific to each subnet, that is, only listing hosts in a single subnet.
	h) Install Intel® Omni-Path Fabric Suite Fabric Manager GUI	At least one subnet manager is required per subnet. Refer to the <i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i> for more information on how to configure a host SM node to manage more than one subnet.
	End Task	

## 14.2 Multi-Rail Usage

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

### 14.2.1 Multi-Rail Overview

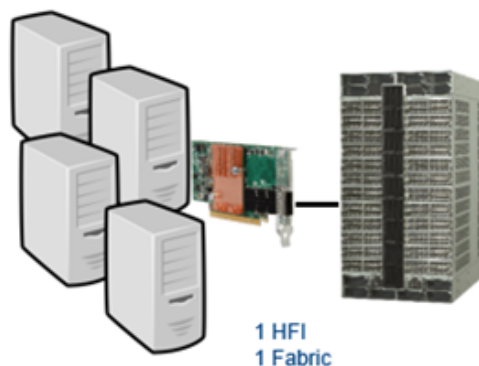
A multi-rail configuration provides load balancing and failover capabilities, adding a higher degree of fabric redundancy. If one HFI or an entire subnet fails, traffic can be moved to the remaining switches.

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.

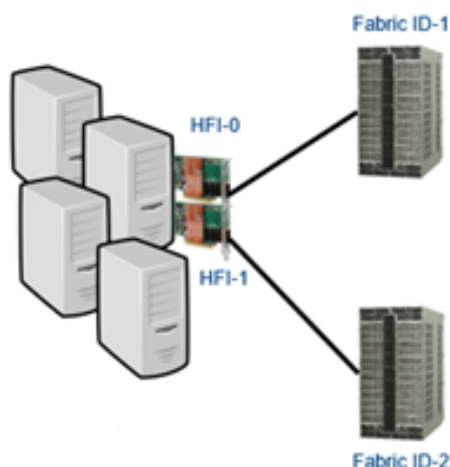
*Note:* Subnets can also be referred to as *planes* or *fabrics*. Rails are also referred to as *HFIs*.

Three basic scenarios include:

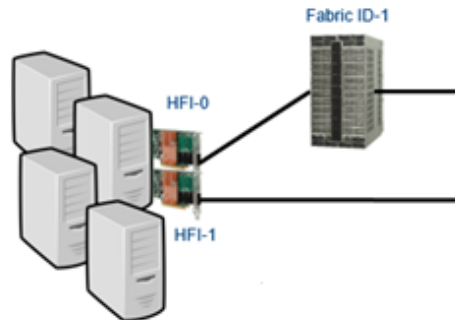
- Single-rail in a single subnet: This scenario, shown in the following figure, consists of one HFI in a server connected to one subnet. This is the default configuration during installation. This configuration provides the performance required by most applications in use today.



- Dual-rail in dual subnets: This scenario, shown in the following figure, consists of two HFIs in the same server connected to separate subnets. Depending on the platform, this configuration may provide 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.



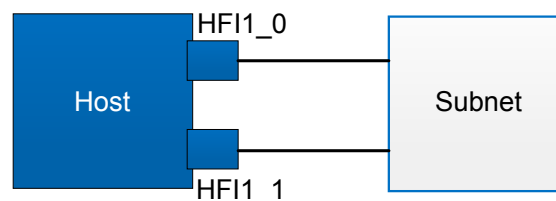
**Note:** Other multi-rail scenarios can be configured. A single Host FM server can manage multiple subnets up to the supported number of FM instances.

## 14.2.2 Setting Up Dual Rails for a Single Subnet

Support for single rail 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 the dual-rail configuration, you need to change the SubnetPrefix for `fm0` as described in the steps below.

### Assumptions

- HFIs have been installed in the host servers.
- HFIs have been cabled as shown below:



- The single subnet has already been configured during software installation.

### Procedures

Perform the following steps to set the SubnetPrefix:

1. On the Host, open the `/etc/opa-fm/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. Note that this is the default configuration for a single-rail in a single subnet.

```
<!-- 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 a unique value other than 0xfe80000000000000. A recommended value to use would be **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. Edit fm1 settings as shown below:
  - Start = 1 (enables FM to start an instance on the second HFI)
  - HFI = 2
  - Port = 1

- Verify that the SubnetPrefix is the same as fm0.

An example of the change 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>0xfe8000000001000</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>
```

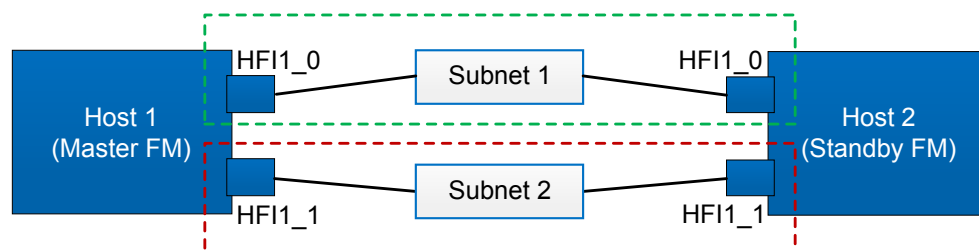
- Save the opafm.xml file.
- Restart the Fabric Manager using `systemctl restart opafm`.  
Host 1 starts up as the master Fabric Manager.
- Copy the opafm.xml file to other hosts using dual rail with single subnet.
- Restart the Fabric Manager on the other hosts using `systemctl restart opafm`.

### 14.2.3 Configuring Dual Rails for Dual Subnets

A common multi-rail, multi-subnet scenario is to set up dual rails for dual subnets where you configure two HFIs in each host to two different subnets to provide failover and load-balancing functions.

#### Assumptions

- HFIs have been installed in the host servers.
- HFIs have been cabled as shown below:



- Two subnets have already been configured during software installation.

#### Procedures

Perform the following steps to set up the master Fabric Manager on Host 1 and standby Fabric Manager on Host 2:

- On Host 1, open the `/etc/opa-fm/opafm.xml` file for editing.
- 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 a unique value other than 0xfe80000000000000. A recommended value to use would be **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>0xfe8000000001001</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 (enables FM to start an instance on the 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>0xfe8000000001001</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.



## 15.0 Installation Verification and Additional Settings

This section provides instructions for verifying that the software has been properly installed and configured, the Intel® Omni-Path Fabric drivers are loaded, and that the fabric is active and ready to use. Information on HFIs and performance tuning is also provided.

### 15.1 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*.

### 15.2 Thermal Monitoring Module Verification

The following procedure provides steps to check the version of your Thermal Monitoring Module (TMM) and to update it to the latest version.

Note the new field for TMM firmware version, an optional micro-controller for thermal monitoring on vendor-specific HFI adapters using the SMBus.

1. Navigate to the Intel® Omni-Path Host Fabric Interface (HFI) firmware updates directory:

```
cd /lib/firmware/updates
```

2. Check the file TMM firmware version in the `hfil_smbus.fw` file using:

```
opatmmtool -f hfil_smbus.fw fileversion
```

3. Check the current TMM firmware version using: `opatmmtool -fwversion`.
4. Check the TMM firmware version in the HFI revision snapshot:

```
# opahfirev
```

Output Example, TMM version is shown in bold type:

```
# opahfirev
#####
node145 - HFI 0000:81:00.0
HFI: hfil_0
Board: ChipABI 3.0, ChipRev 7.17, SW Compat 3
SN: 0x0063be82
Location:Discrete Socket:1 PCISlot:00 NUMANode:1 HFI0
Bus: Speed 8GT/s, Width x16
GUID: 0011:7501:0163:be82
SiRev: B1 (11)
TMM: 10.0.0.0.696
#####
```



5. If the `fwversion` in step 3 and 4 is less than the `fileversion` in step 2, update the TMM firmware version in the HFI:

```
opatmmtool -f hfi1_smbus.fw update
```

## 15.3 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, a lower than expected performance may occur. Use the following command to check PCIe Bus link speed and width.

```
lspci -vv -d :24f0 | grep LnkSta:
```

## 15.4 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 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/neighbor/default/gc_thresh1
```

You may repeat for `gc_thresh2` and `gc_thresh3`.

## 15.5 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*.



## **Part 3: Upgrading the Software**





## 16.0 Upgrade Getting Started

---

This section provides instructions and information for getting started with the Intel® Omni-Path Fabric Suite upgrade.

### 16.1 Upgrade Prerequisites

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/opa-fm/opafm.xml
  - /etc/opa/\*
  - /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.
- Add or remove OS RPMs per [OS RPMs Installation Prerequisites](#) on page 33.

### 16.2 Download the Intel® Omni-Path Software

If the OS you installed did not include the OPA RPMs, download the software package from an Intel web page or other Intel secured location using the following procedures.

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` and press **Enter**.
3. In the Description list select the "Intel® Omni-Path Fabric Software (Including Intel® Omni-Path Host Fabric Interface Driver)" for the version you are wanting to install .

*Note:* The latest version of each type of each download type is showing in the list. To show previous versions select "Show more" at the bottom of the list.

4. In the "Available Downloads" list, select the file(s) you need for the OS you have installed on your fabric.

- For OPA-Basic software, select:

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

- For OPA-IFS software, select:

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



*Note:* `DISTRO.VERSION` refers to the distribution and CPU.

5. Review the Intel Software License Agreement.
6. Click "I accept the terms in the license agreement."
7. Save the download to your hard drive.

## 16.3      **Unpack the Tar File**

You unpack the tar file using the following procedure.

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.
  - For OPA-Basic, use:

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

- For OPA-IFS, use:

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



## 17.0 Upgrade the Intel® Omni-Path Software

This section provides information and procedures to upgrade to the Intel® Omni-Path Software.

### 17.1 Before You Begin

Before starting the upgrade, perform the following:

- Refer to the Release Notes for the list of compatible operating systems.
- Be sure you have completed all [Upgrade Prerequisites](#) on page 121.
- You have downloaded and extracted the software package per [Upgrade Getting Started](#) on page 121.
- If your `opafm.xml` configuration file contains custom changes that you want to carry forward, make a backup copy of it.

### 17.2 Upgrade Scenarios

The following upgrades are available:

- To upgrade to a new version of the Intel® Omni-Path Software for compute nodes (Basic), go to [Upgrade the OPA-Basic Software](#).
- To upgrade to a new version of the Intel® Omni-Path Software for Management Nodes (IFS), go to [Upgrade the OPA-IFS Software](#).
- To upgrade a Management Node from OPA-Basic to OPA-IFS, go to [Upgrade from OPA-Basic to OPA-IFS Software Package](#).

### 17.3 Upgrade the OPA-Basic Software

You upgrade the Intel® Omni-Path Software using the `IntelOPA-Basic.DISTRO.VERSION.tgz` package file.

#### Assumption

- You are logged in.

#### Procedures

Perform the following steps to upgrade the Intel® OP Software:

Step	Task/Prompt	Action
1.	At the command prompt, change directory to the location of the new Basic software package.	Type the following and press <b>Enter</b> : <code>cd IntelOPA-Basic.DISTRO.VERSION</code> where <code>DISTRO.VERSION</code> is the distribution and CPU.
2.	At the command prompt, start the install script.	Type <code>./INSTALL</code> and press <b>Enter</b> .
<i>continued...</i>		



Step	Task/Prompt	Action
3.	Select 1) Install/Uninstall Software.	Type <b>1</b> .
4.	Review the items to be upgraded.	Accept the defaults. Type <b>N</b> to go to the next page. <b>NOTE:</b> If you need to change any item, enter the alphanumeric character associated with the item to toggle between Upgrade or Don't Install.
5.	Start the upgrade.	Type <b>P</b> to perform the actions.
6.	For each system prompt...	Accept the defaults by pressing <b>Enter</b> to continue.
7.	When the Intel OPA Autostart Menu displays, review the items.	Intel recommends leaving all of the Autostart selections set to the default values. <b>NOTE:</b> If you need to change any item, enter the alphanumeric character associated with the item to toggle between Enable or Disable.
8.	Run the OPA Autostart operations.	Type <b>P</b> .
9.	For each system prompt, "Hit any key to continue..."	Press any key. <b>NOTE:</b> When the installation completes, you are returned to the main menu.
10.	Exit the installation menu.	Type <b>X</b> .
11.	Reboot the server.	Type <code>reboot</code> and press <b>Enter</b> .
	<b>End Task</b>	

## 17.4 Upgrade the OPA-IFS Software

You can upgrade the OPA-IFS software to a new version using the Intel OPA Software and FastFabric TUI menus. You begin by upgrading each Management Node and then the remaining servers.

### 17.4.1 Upgrade the Software on the Management Nodes

To upgrade the Intel® Omni-Path Software, you first upgrade the Management Nodes using the `IntelOPA-IFS.DISTRO.VERSION.tgz` package file.

#### Assumption

- You are logged in to the target Management Node.

#### Procedures

Perform the following steps to upgrade the Intel® OP Software on each Management Node:

Step	Task/Prompt	Action
1.	At the command prompt, change directory to the location of the new IFS software package.	For IFS, type the following and press <b>Enter</b> : <code>cd IntelOPA-IFS.DISTRO.VERSION</code> where <code>DISTRO.VERSION</code> is the distribution and CPU.
2.	At the command prompt, start the install script.	Type <code>./INSTALL</code> and press <b>Enter</b> .
3.	Select 1) Install/Uninstall Software.	Type <b>1</b> .
4.	Review the items to be upgraded.	Accept the defaults.
continued...		



Step	Task/Prompt	Action
		Type <b>N</b> to go to the next page. <b>NOTE:</b> If you need to change any item, enter the alphanumeric character associated with the item to toggle between <i>Upgrade</i> or <i>Don't Install</i> .
5.	Start the upgrade.	Type <b>P</b> to perform the actions.
6.	For each system prompt...	Accept the defaults by pressing <b>Enter</b> to continue.
7.	When the Intel OPA Autostart Menu displays, review the items.	Intel recommends leaving all of the Autostart selections set to the default values. <b>NOTE:</b> If you need to change any item, enter the alphanumeric character associated with the item to toggle between <i>Enable</i> or <i>Disable</i> .
8.	Run the OPA Autostart operations.	Type <b>P</b> .
9.	For each system prompt, "Hit any key to continue..."	Press any key. <b>NOTE:</b> When the installation completes, you are returned to the main menu.
10.	Exit the installation menu.	Type <b>X</b> .
11.	Reboot the server.	Type <i>reboot</i> and press <b>Enter</b> .
	<b>End Task</b>	

### Next Steps

- If you have customizations to add to your `opafm.xml` configuration file, go to [Update the Fabric Manager opafm Configuration File](#).
- If you are ready to upgrade the fabric, go to [Upgrade the Software on the Remaining Servers](#).

## 17.4.2 Update the Fabric Manager opafm Configuration File

The `opafm` is designed so that any new parameters that might not appear in your old `opafm.xml` file will automatically be set to their defaults in the software. If you have customized your pre-upgrade `opafm.xml` file, it is recommended that you manually update your new `opafm.xml` file with the latest comments and defaults. This will aid you with future customizations and make comparisons with the `opafm.xml` in future upgrades easier.

### Assumption

- You have upgraded the Fabric Manager, accepting all the defaults.

### Procedures

Perform the following steps to transfer customizations into a new `opafm.xml` file:

Step	Task/Prompt	Action
1.	Compare your retained customized <code>/etc/opa-fm/opafm.xml</code> file to the new default <code>/usr/share/opa-fm/opafm.xml</code> file.	Use the following command to compare files: <code>opafmconfigdiff /etc/opa-fm/opafm.xml /usr/share/opa-fm/opafm.xml</code> <b>NOTE:</b> The resulting differences will highlight your previous customizations.

*continued...*



Step	Task/Prompt	Action
		Refer to the <i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i> for more information on <code>opafmconfigdiff</code> .
2.	Create a new <code>opafm.xml</code> .	Use <code>/usr/share/opa-fm/opafm.xml</code> file as a base for the new file by copying this read-only file to a temporary location to make the subsequent edits.
3.	Update the new <code>opafm.xml</code> with customizations necessary for your fabric.	Use the resulting differences file from Step 1 to locate the customizations and add them to the new <code>opafm.xml</code> .
4.	Update <code>/etc/opa-fm/opafm.xml</code> .	Use the new file from Step 3 to replace your <code>/etc/opa-fm/opafm.xml</code> file.
5.	Restart the Fabric Manager	Type <code>reboot</code> and press <b>Enter</b> .
	<b>End Task</b>	

### Next Steps

- If you are ready to upgrade the fabric, go to [Upgrade the Software on the Remaining Servers](#).

## 17.4.3 Upgrade the Software on the Remaining Servers

After upgrading the Management Nodes, you upgrade the fabric software on the remaining servers using the FastFabric OPA Host Setup menu.

### Assumptions

- You have upgraded the Management Nodes.
- You are logged in.

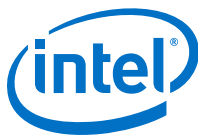
### Procedures

Perform the following steps to upgrade the servers:

Step	Task/Prompt	Action
<b>Configuring the Hosts</b>		
1.	Access the <b>FastFabric OPA Host Setup</b> menu.	
	a) If you are not already logged into <b>Intel FastFabric OPA Tools</b> , at the command prompt...	Type <code>opafastfabric</code> and press <b>Enter</b> .
	b) Access the <b>FastFabric OPA Host Setup</b> menu.	Press <b>3</b> .
	c) Select menu items.	<ul style="list-style-type: none"><li>• Select items <b>0</b> (optional) and <b>5</b>.</li></ul>
	d) Start the operations.	Press <b>P</b> . <b>NOTE:</b> Each selected item is preformed in the order of the menu list.
2.	(Optional) <b>Edit Configuration and Select/Edit Host File</b> (menu item 0)	
	a) Edit the <code>opafastfabric.conf</code> file.	Review the file and change as needed. If you made any changes, save and close the file. Press any key to continue.
	b) Edit the <code>hosts</code> configuration file.	Review the file and change as needed. If you made any changes, save and close the file.
<b>continued...</b>		



Step	Task/Prompt	Action
	c) Do you want to edit/review/change the files? [y]:	Type <b>n</b> and Press <b>Enter</b> .
	e) Continue to next step.	Press any key to continue.
3.	<b>Install/Upgrade OPA Software</b> (menu item 5)	
	a) Do you want to use ./IntelOPA-[Basic IFS].DISTRO.VERSION.tgz? [y]:	Press <b>Enter</b> to accept the default.
	b) Would you like to do a fresh [i]ninstall, an [u]pgrade or [s]kip this step? [u]:	Press <b>Enter</b> to accept the default.
	c) Are you sure you want to proceed? [n]:	Type <b>y</b> and press <b>Enter</b> .
	d) When the Intel OPA Autostart Menu displays, review the items.	Intel recommends leaving all of the Autostart selections set to the default values. <b>NOTE:</b> If you need to change any item, enter the alphanumeric character associated with the item to toggle between <i>Enable</i> or <i>Disable</i> .
	e) Run the OPA Autostart operations.	Type <b>P</b> .
	f) For each system prompt, "Hit any key to continue..."	Press any key. <b>NOTE:</b> When the installation completes, you are returned to the main menu.
	g) Complete the installation.	Press any key to continue.
	• If all hosts install...	Press any key to continue.
	• If any hosts fails to install...	Use the <b>View opahostadmin Result Files</b> menu item to review the result files from the update.
<b>Optional Tasks</b> Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more information.		
4.	<b>Reboot Hosts</b> (menu item 8)	<b>(Linux)</b> This menu item 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.
5.	<b>Run a Command on All Hosts</b> (menu item b)	<b>(Linux)</b> For any other setup operations that need to be performed on all hosts, this menu item executes the specified Linux* shell command against all selected hosts. It can also execute a sequence of commands separated by semicolons. <i>Note:</i> Check the relevant release notes for the new Intel® Omni-Path Fabric Host Software release being installed for any additional required steps.
6.	<b>Copy a File to All Hosts</b> (menu item c)	This menu item specifies a file on the local host to be copied to all selected hosts. <i>Note:</i> 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.
7.	<b>View opahostadmin Result Files</b> (menu item d)	
	a) About to: vi /root/test.res /root/test.log	Press any key to review files.
	b) Would you like to remove test.res test.log test_tmp* and save_tmp in /root ? [n]:	Press <b>Enter</b> to save or type <b>y</b> to remove the files.
	<b>End Task</b>	



### Next Steps

- To verify the host software has been upgraded and running on the remaining servers, go to [Verify the Host Software on the Remaining Servers Using the FastFabric TUI Menu](#).

## 17.5 Upgrade from OPA-Basic to OPA-IFS Software Package

You can upgrade from OPA-Basic to OPA-IFS in order to install the FastFabric and OPA FM software.

### Assumptions

- OPA-Basic is installed on the server.
- You are logged in.

### Procedures

Perform the following steps to upgrade to OPA-IFS:

Step	Task/Prompt	Action
1.	At the command prompt, change directory to the location of the new IFS software package.	For IFS, type the following and press <b>Enter</b> : <code>cd IntelOPA-IFS.DISTRO.VERSION</code> where <i>DISTRO.VERSION</i> is the distribution and CPU.
2.	At the command prompt, start the install script.	Type <code>./INSTALL</code> and press <b>Enter</b> .
3.	Select 1) Install/Uninstall Software.	Type <b>1</b> .
4.	Review the items to be upgraded. <b>NOTE:</b> Ensure that FastFabric and OPA_FM show Install and all other selections show Up To Date.	Type <b>N</b> to go to the next page. <b>NOTE:</b> If you need to change any item, enter the alphanumeric character associated with the item to toggle between Upgrade or Don't Install.
5.	Start the upgrade.	Type <b>P</b> to perform the actions.
6.	For each system prompt...	Accept the defaults by pressing <b>Enter</b> to continue.
7.	When the Intel OPA Autostart Menu displays, review the items.	Intel recommends leaving all of the Autostart selections set to the default values. <b>NOTE:</b> If you need to change any item, enter the alphanumeric character associated with the item to toggle between Enable or Disable.
8.	Run the OPA Autostart operations.	Type <b>P</b> .
9.	For each system prompt, "Hit any key to continue..."	Press any key. <b>NOTE:</b> When the installation completes, you are returned to the main menu.
10.	Exit the installation menu.	Type <b>X</b> .
11.	Reboot the server.	Type <code>reboot</code> and press <b>Enter</b> .
	<b>End Task</b>	

### Next Steps

- If you are setting up additional Management Nodes, go to [Set Up Additional Management Nodes](#) on page 85.





## Appendix A UEFI Secure Boot Support

---

Intel® Omni-Path software includes support for Unified Extensible Firmware Interface (UEFI) Secure Boot, both in the UEFI firmware environment and in Linux\*.

### A.1 UEFI Firmware Support

All UEFI drivers shipped as part of the Intel® Omni-Path Fabric 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.

### A.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. Refer to the [Client Node](#) on page 142 section in the [Setting Up PXE Boot for Supported Operating Systems](#) on page 130 appendix 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\*, refer to 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](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\*, refer to the SLES\* documentation at [https://www.suse.com/documentation/sled11/book\\_sle\\_admin/data/sec\\_uefi\\_secboot.html](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
- hfidiagnostics

## Appendix B Setting Up PXE Boot for Supported Operating Systems

Different operating systems require specific instructions for setting up PXE Boot on the Intel® Omni-Path Fabric.

The basic requirement for the preferred environment consists of a cluster with at least two nodes connected to a switch. One of the nodes is the Client node, where the OS will be installed over HFI. The other node will be a Preboot Execution Environment (PXE) Server to which the Client will connect to initialize and perform the OS installation.

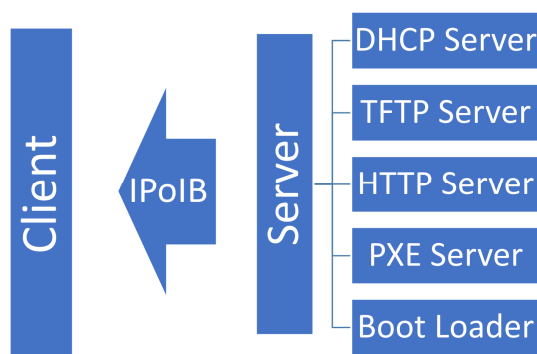
The communication over Intel® Omni-Path (IP over InfiniBand\*) proceeds as follows:

1. Client requests to DHCP Server and receives the IP address of TFTP Server and boot image (the name of bootloader).
2. Client contacts the TFTP Server and receives the boot image.
3. Client executes the boot image and downloads from PXE Server all other needed files like kernel and initial root file system (`initramfs`) image and loads them.

*Note:* The `initramfs` image does not contain all the modules and firmware required to use IPoIB with HFI. Therefore it **needs to be updated** before booting over HFI.

4. The installation of the target OS starts on the Client machine using the ISO image of the target OS downloaded from the HTTP Server.

The figure below shows a high-level overview of the communication between Client and Server.



The following procedures describe how to configure and boot a server over Intel® Omni-Path using the HFI UEFI firmware. Note that it does not cover all possible PXE configurations. The examples given are targeted for servers running the latest, supported Red Hat\* Enterprise Linux\* (RHEL\*) and SUSE\* Linux\* Enterprise Server (SLES\*) operating systems.



- For information on other PXE server configurations, please refer to your OS vendor documentation.
- For PXE boot information on earlier supported RHEL\* and SLES\* OS versions, contact your Intel Support Personnel.

## B.1 PXE Server

You use the Intel® Omni-Path Host Fabric Interface (HFI) Unified Extensible Firmware Interface (UEFI) firmware to configure a server to boot over Intel® Omni-Path.

This sections describes the procedures for different, supported operating systems.

### B.1.1 Configure PXE Server with RHEL\* 7.4

#### B.1.1.1 Before You Begin

The OPA-IFS package contains the HFI driver and the required IP over IB software stack. The package also installs the Fabric Manager. Note that the Fabric Manager must 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 Suite Fabric Manager software must be running.

Usually Fabric Manager is running on the **Switch**, but in a simple Back-to-Back configuration, Fabric Manager will be started on a head node (PXE server).

**Notes:** If you need to install and start the Fabric Manager, perform the following steps:

- a. Install Intel® Omni-Path Fabric Suite Fabric Manager using the command:

```
yum install opa-fm
```

- b. Start up Fabric Manager using the following command:

```
service opafm start
```

2. The HFI driver (hfi1) and IP over InfiniBand\* modules must be running. Use the commands:

```
modprobe hfi1  
modprobe ib_ipoib
```

#### B.1.1.2 Configure 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. Add `ib0` network interface configuration file: `/etc/sysconfig/network-scripts/ifcfg-ib0`



2. Add the following content related to DHCP configuration to the configuration file:

```
DEVICE=ib0
TYPE=Infiniband
BOOTPROTO=static
IPADDR=192.168.100.1
NETMASK=255.255.255.0
NETWORK=192.168.100.0
BROADCAST=192.168.100.255
ONBOOT=yes
USERCTL=no
```

3. Run the following command, where `ib0` is the name of the network interface configured in step 2:

```
ifup ib0
```

### B.1.1.3 Configure DHCP Server

The following procedure describe how to configure DHCP for dynamic IP or fixed IP addressing.

#### For Dynamic IP Addressing

To install a DHCP server for dynamic IP addressing, follow the steps below:

1. Install DHCP package using the command:

```
yum install dhcp
```

2. Add DHCP configuration file.

The following is an example of `/etc/dhcp/dhcpd.conf`

```
#
# DHCP Server Configuration file.
#
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 dhcpd start
```

4. Enable DHCP service to be run always with OS start:

```
systemctl enable dhcpd.service
```

### For Fixed IP Addressing

To install a DHCP server for fixed IP addressing, follow the steps below:

1. Use the GUID from the client HFI in this example:

```
001175010165abb2
```

2. Configure the `dhcpd.conf` file as the following, using 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. Start DHCP server using the command:

```
service dhcp start
```



4. The output from `/var/log/messages` for 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
```

#### B.1.1.4 Install a TFTP Server

**Prerequisite:** XINETD may have been installed with OS. If not, use the following command:

```
yum install xinetd
```

To install a TFTP server, perform the following steps:

1. Install TFTP package using the command:

```
yum install tftp-server
```

2. Add or modify 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 defaults to `yes` and should be changed to `no` before starting the XINETD server.

3. Start XINETD server using the command:

```
service xinetd start
```

4. Enable XINETD service to be run always with OS start:

```
systemctl enable xinetd.service
```



### B.1.1.5 Create Directory Structure for PXE Server

The PXE Server is contained within a directory structure indicated by the `server_args` argument of the TFTP configuration file.

1. Create the base directory structure using the command:

```
mkdir -p /var/lib/tftpboot/uefi
```

### B.1.1.6 Configure the HTTP Server

To configure the HTTP Server, perform the following steps:

1. Install the TFTP package using the command:

```
yum install httpd
```

2. Add the configuration file `/etc/httpd/conf.d/netinstall.conf` with the following content:

```
Alias /netinstall /var/www/netinstall
<Directory /var/www/netinstall>
Options Indexes FollowSymLinks
Order Deny,Allow
Allow from all
</Directory>
```

3. Create a new directory for iso image of the system that will be installed over the Intel® Omni-Path Host Fabric Interface. Use the following command:

```
mkdir -p /var/www/netinstall/ISO/
```

4. Start the HTTP server using the command:

```
service httpd start
```

5. Enable the HTTP service to be run always with OS start:

```
systemctl enable httpd.service
```

6. Whenever the content of `/var/www/netinstall` is modified use the command:

```
service httpd reload
```

### Post-requisite

To see if the HTTP server was started successfully, open a web browser and enter <IP address of the PXE server>/netinstall.

For example:, if IP address of the PXE server is 100.100.200.200, use `http://100.100.200.200/netinstall/`.



### B.1.1.7 Configure the Boot Loader

To configure a PXE server to use the **Elilo** boot loader, perform the following steps.

**Note:** For configuration details of other boot loaders, please refer to your OS vendor documentation.

1. Download the latest source files for Elilo: <https://sourceforge.net/projects/elilo/files/elilo/>
2. Unpack the archive file.

```
tar xzf elilo-3.16-all.tar.gz
```

3. Rename the efi file and copy it to the PXE Server directory.

```
mv elilo-3.16-x86_64.efi bootx64.efi
cp bootx64.efi /var/lib/tftpboot/uefi/
```

4. Unpack the next archive file and copy the text menu message file to the PXE Server directory.

```
tar xzf elilo-3.16-source.tar.gz
cp elilo-3.16-source/examples/textmenu_chouser/textmenu-message.msg /var/lib/tftpboot/uefi/
```

5. Create the Elilo configuration file: `/var/lib/tftpboot/uefi/elilo.conf`.

The following is an example for RHEL\* 7.x using `vmlinux` kernel file and `initramfs.img` as a minimal file system image:

```
chooser=textmenu
message=textmenu-message.msg
prompt
delay=20

default=rhel7
image=vmlinux
label=rhel7
description = "RHEL 7"
initrd = initramfs.img
read-only
```

## B.1.2 Configure PXE Server with SLES\* 12 SP3

### B.1.2.1 Before You Begin

The OPA-IFS package contains the HFI driver and the required IP over IB software stack. The package also installs the Fabric Manager. Note that the Fabric Manager must 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 Suite Fabric Manager software must be running.

Usually Fabric Manager is running on the **Switch**, but in a simple Back-to-Back configuration, Fabric Manager will be started on a head node (PXE server).





**Notes:** If you need to install and start the Fabric Manager, perform the following steps:

- a. Install Intel® Omni-Path Fabric Suite Fabric Manager using the command:

```
zypper install opa-fm
```

- b. Start up Fabric Manager using the following command:

```
service opafm start
```

2. The HFI driver (hfi1) and IP over InfiniBand\* modules must be running. Use the commands:

```
modprobe hfi1  
modprobe ib_ipoib
```

### B.1.2.2 Configure 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. Add `ib0` network interface configuration file: `/etc/sysconfig/network-scripts/ifcfg-ib0`
2. Add the following content related to DHCP configuration to the configuration file:

```
DEVICE=ib0  
TYPE=Infiniband  
BOOTPROTO=static  
IPADDR=192.168.100.1  
NETMASK=255.255.255.0  
NETWORK=192.168.100.0  
BROADCAST=192.168.100.255  
ONBOOT=yes  
USERCTL=no
```

3. Run the following command, where `ib0` is the name of the network interface configured in step 2:

```
ifup ib0
```

### B.1.2.3 Configure DHCP Server

The following procedure describe how to configure DHCP for dynamic or fixed IP addressing.



### For Dynamic IP Addressing

To install a DHCP server for dynamic IP addressing, follow the steps below:

1. Install DHCP package using the command:

```
zypper install dhcp-server
```

2. Add DHCP configuration file.

The following is an example: `/etc/dhcpd.conf`

```
#
# DHCP Server Configuration file.
#
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. Modify `/etc/sysconfig/dhcpd` by adding at the end of file line:

```
DHCPD_INTERFACE="ib0"
```

4. Start DHCP server using the command:

```
service dhcpd start
```

5. Enable DHCP service to be run always with OS start:

```
systemctl enable dhcpd.service
```

### For Fixed IP Addressing

To install a DHCP server for fixed IP addressing, follow the steps below:

1. Use the GUID from the client HFI in this example:

```
001175010165abb2
```



2. Configure the `dhcpd.conf` file as the following, using the standard `ff:00:00:00:00:02:00:00:01:57:`

```
[root@phs1fnivd14u25 dhcpd]# 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. Start DHCP server using the command:

```
service dhcp start
```

4. The output from `/var/log/messages` for 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
```



#### B.1.2.4 Install a TFTP Server

**Prerequisite:** XINETD may have been installed with OS. If not, use the following command:

```
zypper install xinetd
```

To install a TFTP server, perform the following steps:

1. Install TFTP package using the command:

```
zypper install tftp
```

2. Add or modify TFTP configuration file. An example is shown below: `/etc/xinetd.d/tftp`

```
service tftp
{
    socket_type           = dgram
    protocol              = udp
    wait                  = yes
    flags                  = IPv6 IPv4
    user                  = root
    server                 = /usr/sbin/in.tftpd
    server_args            = -u tftp -s /srv/tftpboot
    # per_source           = 11
    # cps                   = 100 2
    disable                = no
}
```

*Note:* The `disable` field defaults to `yes` and should be changed to `no` before starting the XINETD server.

3. Start XINETD server using the command:

```
service xinetd start
```

4. Enable XINETD service to be run always with OS start:

```
systemctl enable xinetd.service
```

#### B.1.2.5 Create Directory Structure for PXE Server

The PXE Server is contained within a directory structure indicated by the `server_args` argument of the TFTP configuration file.

1. Create the base directory structure using the command.

```
mkdir -p /srv/tftpboot/uefi
```

2. Change the SUSE default access permission of the TFTP Server directory.

```
chmod 755 /srv/tftpboot/
```



### B.1.2.6 Configure the HTTP Server

To configure the HTTP Server, perform the following steps:

1. Install the TFTP package using the command:

```
zypper install apache2
```

2. Add the configuration file `/etc/apache2/conf.d/netinstall.conf` with the following content:

```
Alias /netinstall /var/www/netinstall
<Directory /var/www/netinstall>
Options Indexes FollowSymLinks
Require all granted
</Directory>
```

*Note:* The configuration file above is valid for Apache\* version 2.4 or higher.

3. Create a new directory for iso image of the system that will be installed over Intel® Omni-Path Host Fabric Interface. Use the following command:

```
mkdir -p /srv/www/netinstall/ISO/
```

4. Start the HTTP server using the command:

```
service apache2 start
```

5. Enable the HTTP service to be run always with OS start:

```
systemctl enable apache2.service
```

6. Whenever the content of `/srv/www/netinstall` is modified use the command:

```
service apache2 reload
```

#### Post-requisite

To see if the HTTP server was started successfully, open a web browser and enter <IP address of the PXE server>/netinstall.

For example:, if IP address of the PXE server is 100.100.200.200, use `http://100.100.200.200/netinstall/`.

### B.1.2.7 Configure the Boot Loader

To configure a PXE server to use the **Elilo** boot loader, perform the following steps.

*Note:* For configuration details of other boot loaders, please refer to your OS vendor documentation.

1. Download the latest source files for Elilo: <https://sourceforge.net/projects/elilo/files/elilo/>



2. Unpack the archive file.

```
tar xzf elilo-3.16-all.tar.gz
```

3. Rename the `efi` file and copy it to the PXE Server directory.

```
mv elilo-3.16-x86_64.efi bootx64.efi
cp bootx64.efi /srv/tftpboot/uefi/
```

4. Unpack the next archive file and copy the text menu message file to the PXE Server directory.

```
tar xzf elilo-3.16-source.tar.gz
cp elilo-3.16-source/examples/textmenu_chooser/textmenu-message.msg /srv/
tftpboot/uefi/
```

5. Create the Elilo configuration file: `/srv/tftpboot/uefi/elilo.conf`.

The following is an example for SUSE\* Linux 12.x using `vmlinux` kernel file and `initramfs.img` as a minimal file system image:

```
chooser=textmenu
message=textmenu-message.msg
prompt
delay=20

default=suse12
image=vmlinux
label=suse12
description = "SUSE Linux 12"
initrd = initramfs.img
read-only
```

## B.2 Client Node

You install the operating system over Intel® OP HFI for the Client node.

The following table shows the BIOS versus Boot Mode information.

**Table 3. BIOS versus Boot Mode**

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	Customized by UEFI driver	Customized by UEFI driver
continued...							



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
					Boot from Ethernet		
UEFI	Legacy BIOS	Yes (BIOS that runs UEFI driver in this mode)	Achieved pre boot	Provided by UEFI driver	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

This sections describes the procedures for different, supported operating systems.

## B.2.1 Before You Begin

Before installing the operating system over Intel® OP HFI, perform the following:

- Perform the steps in [Upgrade the HFI UEFI Firmware](#) if needed.
- Be sure to build the `initrd.img` image file on a machine that uses the same OS distribution that it targets. For instance, you build RHEL\* 7.2 `initrd.img` on a machine that has Linux RHEL\* 7.2 running on it.

## B.2.2 Upgrade the HFI UEFI Firmware

The HFI UEFI firmware is available as part of the `hfi1-uefi.x86_64 RPM`, released with the Intel® Omni-Path Software package.

To upgrade the HFI UEFI firmware, perform the following steps:

1. Install the new HFI UEFI rpm.

```
# rpm -i hfi1-uefi-X.X-XX.x86_64.rpm
```

2. Change directory to `/usr/share/opa/bios_images/`.

```
# cd /usr/share/opa/bios_images/
```

3. Update the HFI firmware. The basic command is:

```
# hfi1_eprom [See Additional Options] -w -o HfiPcieGen3Loader_1.6.0.0.0.rom -b HfiPcieGen3_1.6.0.0.0.0.efi
```



Example output of update:

```
Using device: /sys/bus/pci/devices/0000:81:00.0/resource0
Erasing loader file... done
Writing loader file... done
Erasing driver file... done
Writing driver file... done
```

### Additional Options:

The following information provides guidance for various firmware update options.

- To update all of the HFIs, add the `-d all` option before the `-w` option.

```
# hfil_eprom -d all -w -o HfiPcieGen3Loader_1.6.0.0.rom -b
HfiPcieGen3_1.6.0.0.0.efi
```

- To update each HFI separately (using `-d` only), perform the following steps:

*Note:* This procedure needs to be repeated for every HFI card installed in the system (listed by `lspci` command).

- a. Collect the PCIe addresses (first column in the output below) for each HFI to be updated using `lspci |grep HFI`.

Example output is:

```
# lspci |grep HFI
05:00.0 Fabric controller: Intel Corporation Omni-Path HFI Silicon 100
Series [discrete] (rev 11)
81:00.0 Fabric controller: Intel Corporation Omni-Path HFI Silicon 100
Series [discrete] (rev 11)
```

- b. Add `-d <resource0_file_full_path>` as shown in the example below:

```
# hfil_eprom -d /sys/bus/pci/devices/<hfi_pcie_address>/resource0 -w -
o HfiPcieGen3Loader_1.6.0.0.rom -b HfiPcieGen3_1.6.0.0.0.efi
```

where `<hfi_pcie_address>` is the value read from the first column of the `lspci` command output. From the example output in [3.a](#), the `<resource0_file_full_path>` is: `/sys/bus/pci/devices/0000:81:00.0/resource0`.

4. Read loader file version.

```
# hfil_eprom -V -o
```

Example output:

```
Using device: /sys/bus/pci/devices/0000:02:00.0/resource0
loader file version: 1.6.0.0.0
```

5. Read driver file version.

```
# hfil_eprom -V -b
```





Example output:

```
Using device: /sys/bus/pci/devices/0000:02:00.0/resource0
driver file version: 1.6.0.0.0
```

6. Reboot the server.

```
# reboot
```

### B.2.3 Install RHEL\* 7.4 over Intel® OP HFI

To edit the `initrd.img`, perform the following steps:

1. Log in as root.
2. Decompress the `.iso` archive with RHEL\* 7.4 distribution.  
The `initrd.img` is located in the `/images/pxeboot` folder.
3. Create `/home/images/initrd` folder.

```
mkdir -p /home/images/initrd
```

4. Copy `initrd.img` to `home/images/initrd`.

```
cp <.iso decompressed archive path>/images/pxeboot/initrd.img /home/images/
initrd
```

5. Change directories to `/home/images/initrd`.

```
cd /home/images/initrd
```

6. Check archive type.

For example:

```
file initrd.img
initrd.img: XZ compressed data
```

7. Extract `initrd.img` according to archive type:

- For XZ archive

```
xzcat ./initrd.img | cpio -d -i -m
```

- For GZ archive

```
gunzip < initrd.img | cpio -i --make-directories
```

8. Move `initrd.img` out of the folder.

```
mv ./initrd.img ../initrd.img
```

**Note:** From this point on in the instructions, all commands assume that the present working directory (`pwd`) is `/home/images/initrd` and all commands are relative to this path.



9. Copy firmware files.

```
cp --recursive /lib/firmware/updates/ usr/lib/firmware/updates/
```

The list of required firmware files are:

```
/lib/firmware/updates/hfil_dc8051_d.fw
/lib/firmware/updates/hfil_dc8051.fw
/lib/firmware/updates/hfil_fabric_d.fw
/lib/firmware/updates/hfil_fabric.fw
/lib/firmware/updates/hfil_pcie_d.fw
/lib/firmware/updates/hfil_pcie.fw
/lib/firmware/updates/hfil_platform.dat
/lib/firmware/updates/hfil_sbus_d.fw
/lib/firmware/updates/hfil_smbus.fw
/lib/firmware/updates/hfi_dc8051.bin
```

10. Add `ib_ipoib.conf` file.

```
mkdir etc/modules-load.d/
touch etc/modules-load.d/ib_ipoib.conf
echo "# Load ib_ipoib.ko at boot" >> etc/modules-load.d/ib_ipoib.conf
echo "ib_ipoib" >> etc/modules-load.d/ib_ipoib.conf
```

11. Generate `initrd-new.img` file and give permissions.

```
cd /home/images/initrd
find . -print |cpio -o -H newc | xz --format=lzma > ../initrd-new.img
chmod 777 ../initrd-new.img
```

**Note:** Remember to use appropriate Linux kernel file (`vmlinuz`) along with `initrd.img`.

12. Sample `elilo.conf`.

```
prompt
chooser=textmenu
message=textmenu-message.msg
delay=20

default=rhel74
image=vmlinuz
    label=rhel74
    description = "RHEL"
    initrd=initrd-new.img
    read-only
    append="ip=ib0:dhcp rd.driver.post=ib_ipoib rd.net.timeout.carrier=50
inst.repo=http://192.168.100.1/netinstall/RHEL74"
```

## B.2.4 Install SLES\* 12 SP3 over Intel® OP HFI

To edit the `initrd.img`, perform the following steps:

1. Log in as root.
2. Mount the SLES\* 12 SP3 `.iso` file.

```
sudo mkdir /mnt/iso
sudo mount -o loop SLE-12-SP2-Server-DVD-x86_64-GM-DVD1.iso /mnt/iso
```



3. Create a working directory /home/new\_initrd.

```
mkdir /home/new_initrd
```

4. Copy linux file to /home/new\_initrd.

```
cp /mnt/iso/boot/x86_64/loader/linux /home/new_initrd
```

**Note:** linux (Linux kernel – corresponds to “vmlinuz” in RHEL) will not be modified. Please remember to use this file as the kernel file provided for PXE Boot (in elilo.conf).

5. Copy initrd.img to home/new\_initrd.

```
cp /mnt/iso/boot/x86_64/loader/initrd /home/new_initrd/initrd.img
```

**Note:** initrd (initial root filesystem image) will be modified in this procedure.

6. Change directories to /home/new\_initrd.

```
cd /home/new_initrd
```

7. Create a new initrd folder and move initrd.img into it.

```
mkdir initrd  
mv initrd.img ./initrd
```

8. Extract initrd file.

```
cd initrd  
xzcat ./initrd.img | cpio -d -i -m
```

9. Move initrd.img out of the folder.

```
mv initrd.img ../  
cd ..
```

10. Go to initrd/parts directory.

```
cd initrd/parts
```

11. Copy 00\_lib file into new\_initrd directory (outside of initrd).

```
cp 00_lib ../../  
cd ../..
```

12. Extract 00\_lib file.

```
unsquashfs 00_lib
```

**Note:** 00\_lib will be extracted to a newly created ./squashfs-root directory.



After this step, the present working directory (/home/new\_initrd) should look like this:

```
# ls
00 lib
initrd
initrd.img
linux
squashfs-root
```

13. Copy five firmware drivers to squashfs-root/lib/firmware directory.

```
cp --recursive /lib/firmware/updates/ ./squashfs-root/lib/firmware/
```

14. Copy the following additional drivers from system that are not present in the default initrd.img.

```
cp --recursive /lib/modules/$(uname -r)/updates/ifs-kernel-updates/
squashfsroot/ lib/modules/$(uname -r)/initrd/
```

15. In order to keep the existing module dependencies, temporarily copy some modules from <initrd-root>/lib/modules/\$(uname -r)/initrd into squashfsroot/lib/modules/\$(uname -r)/initrd.

```
cp initrd/lib/modules/$(uname -r)/initrd/lz4_decompress.ko squashfsroot/
lib/modules/$(uname -r)/initrd/
cp initrd/lib/modules/$(uname -r)/initrd/squashfs.ko squashfsroot/
lib/modules/$(uname -r)/initrd/
cp initrd/lib/modules/$(uname -r)/initrd/loop.ko squashfs-root/lib/modules/$(uname
-r)/initrd/
```

16. Create a backup copy of an existing <initrd-root>/lib/modules directory.

```
cp -R initrd/lib/modules modules_backup
```

17. Temporarily remove the modules directory from initrd.

```
rm -rf initrd/lib/modules
```

18. Copy modules directory from squashfs to initrd.

```
cp -R squashfs-root/lib/modules initrd/lib
```

19. Change root to initrd root directory and rebuild module dependencies.

```
cd initrd
chroot $(pwd)
depmod -a -v
exit
cd ..
```



20. Go to `<initrd-root>/lib/modules/$(uname -r)/initrd` and remove all files added in 15.

```
rm initrd/lib/modules/$(uname -r)/initrd/lz4_decompress.ko
rm initrd/lib/modules/$(uname -r)/initrd/loop.ko
rm initrd/lib/modules/$(uname -r)/initrd/squashfs.ko
```

21. Remove modules directory from squashfs.

```
rm -rf squashfs-root/lib/modules
```

22. Copy a modified modules directory back to the squashfs.

```
cp -R initrd/lib/modules squashfs-root/lib
```

23. Remove modules directory from initrd.

```
rm -rf initrd/lib/modules
```

24. Restore (from the backup done in 16) the previous content of modules directory.

```
cp -R modules_backup initrd/lib/modules
```

25. Pack squashfs-root contents again to 00\_lib.

```
mksquashfs squashfs-root 00_lib_new
```

26. Overwrite the existing 00\_lib in the initrd directory.

```
cp 00_lib_new initrd/parts/00_lib
```

27. Compress again initrd contents (from initrd-root).

```
cd initrd
find . -print | cpio -ov -H newc | xz --format=lzma > ../initrd-sles.img
```

28. Change privileges of a new initrd.

```
chmod 777 ../initrd-sles.img
```

29. Update `elilo.conf` to include the line `append="insmod=ib_ipoib netdevice=ib0 netsetup=dhcp"`.

This forces an automatic `ib_ipoib` module start, uses `ib0` as a default network device, and configures IP address via DHCP.

```
prompt
chooser=textmenu
message=textmenu-message.msg
delay=20

default=sles123
image=loader/linux
  label=sles123
  description = "SLES"
  initrd=loader/initrd-new.img
```



```
read-only
append="insmod=ib_ipoib netdevice=ib0 netsetup=dhcp install=http://
192.168.100.1/netinstall/SLES_123"
```