



Intel® Omni-Path Fabric Software

Installation Guide

Rev. 17.0

April 2020



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 product plans and roadmaps are subject to change without notice.

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.

Intel technologies may require enabled hardware, software or service activation.

No product or component can be absolutely secure.

Your costs and results may vary.

Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

Copyright © 2015–2020, Intel Corporation. All rights reserved.



Revision History

Date	Revision	Description
April 2020	17.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Updated Download and Install NVIDIA* Software (Optional) to include how to verify that GPUDirect is installed. Updated Using the INSTALL Command Line Options to remove note stating that openmpi_gcc_cuda_hfi (CUDA-enabled) is installed by default. Updated Install Using Linux* Distribution Software Packages Provided by Intel to include information about default installation options. Added new section, Reinstall Software After a Kernel Update.
January 2020	16.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Updated links to Intel® Omni-Path documentation set in Intel® Omni-Path Documentation Library. Updated instructions for Download the Intel® Omni-Path Software. Updated instructions for Download the Firmware. Updated Install Using CLI Commands to remove connected mode from IPV4 Configuration instructions. Updated Install Using Linux* Distribution Software Packages Provided by Intel to discuss virtual packages. Updated Repositories Included in the IFS Package to include Interoperate with the INSTALL Script. Updated Repository Deployment into the Environment with updated opacreaterrepo example. Updated Install a TFTP Server to include SLES 12 SPx and SLES 15 and onward. Updated instructions for Download the Fabric Manager GUI Software. Changed Elilo to Grub2 in Configure the Boot Loader.
October 2019	15.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Globally removed SHMEM; it is no longer supported. Globally removed references to RHEL* 7.5; it is no longer supported. Updated Preface to include new Best Practices. Combined sections, Setting Up Primarily Independent Subnets and Setting Up Overlapping Subnets, into Setting Up Subnets and included new optional step for multi-rails. Updated information for Setting Up Dual Rails for a Single Subnet and Configuring Dual Rails for Dual Subnets.
June 2019	14.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Added new section About User Queries Settings. Updated Repositories Included in the IFS Package to change OPA_COMMON to OPA_PKGS and OPA_CUDA to OPA_PKGS_CUDA. Updated Repository Deployment into the Environment to change OPA_COMMON to OPA_PKGS and OPA_CUDA to OPA_PKGS_CUDA and updated Services Node example.
March 2019	13.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Updated section Install Using Linux* Distribution Software Packages Provided by Intel to provide simplified installation instructions using new opacreaterrepo helper script.
February 2019	12.0	<p>Document has been updated with the following:</p>
continued...		



Date	Revision	Description
		<ul style="list-style-type: none"> Updated Using the INSTALL Command Line Options to remove <code>-r</code> and <code>--prefix</code> options.
December 2018	11.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Global change: For all applicable <code>opafastfabric</code> TUI menus, added <code>Q</code> (quit) as an additional method to <code>Exit</code> or <code>Return to Previous Menu</code> Removed OS Distribution RPMs from OS RPMs Installation Prerequisites and moved into <i>Intel® Omni-Path Fabric Software Release Notes</i> Updated Install the Intel Omni-Path Software to include new Linux* Distribution Software packages for installation. Updated Using the INSTALL Command Line Options <code>./INSTALL -G</code> for GPUDirect-supported Oses and additional non-root user set up for <code>--user_queries</code>. Added new section Install Using Linux* Distribution Software Packages Provided by Intel with RHEL* and SLES* instructions. Renamed section Install Using the Linux* Distribution Software Packages Provided by the OS Vendor. Updated Setting Up PXE Boot for Supported Operating Systems to be more generic by removing specific RHEL and SLES versions for subsections. Updated JRE 1.8 information in Install the Fabric Manager GUI Software on Windows* and Install the Fabric Manager GUI Software on Linux*. Updated Before You Begin and Install the Host Software on the Remaining Hosts Using the FastFabric TUI Menus in Install Host Software on Remaining Servers to include the RHEL 7.x need for <code>pigz</code> compression tool. Added Intel OPA Software Components to Packages Mapping to appendix.
October 2018	10.1	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Updated Fabric Setup Prerequisites to remove obsolete information for <code>rdma-ndd</code> in <code>infiniband-diags</code> package. Updated OS RMPs Installation Prerequisites to include the following: <ul style="list-style-type: none"> <code>rdma-ndd</code> in SLES* 12.3 and as a note in RHEL* that it is part of <code>rdma-core</code>. <code>kernel-syms</code> in SLES* 12.2 and SLES* 12.3.
September 2018	10.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Global change: Removed Verbs only builds for Open MPI and MVAPICH2 (gcc) and builds built with Intel compiler for Open MPI and MVAPICH2. Removed RHEL* 7.3. Added RHEL* 7.5 and SLES* 15. Updated Intel Omni-Path Architecture Overview. Updated OS RPMs Installation Prerequisites to include <code>mpi-selector</code> in SLES rpms. Removed unused rpms (related to <code>libelf</code>). Added new RHEL* 7.5 and SLES* 15 RPM tables. Updated Download and Install NVIDIA* Software (Optional). Updated Using the INSTALL Command Line Options Updated Install the Fabric Manager GUI Software on Windows* and Install the Fabric Manager GUI Software on Linux* to include Pre-requisites to installation. Updated System Requirements for a Windows* Environment and System Requirements for a Linux* Environment - Support for JRE 1.8 Updated Before You Begin and Upgrade the OPA-IFS Software section to include stopping/starting the FM. Added subsections for Starting the Fabric Manager and Stopping the Fabric Manager.
May 2018	9.1	<p>Document has been updated for OS RPMs Installation Prerequisites: Added <code>papi</code> to RHEL* RPM Matrix. Added RHEL*- and SLES*-specific notes for "additional, required development packages" to respective sections.</p>
continued...		



Date	Revision	Description
April 2018	9.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Removed all instances of GASNET and OpenSHMEM (hfi, gcc) throughout document Removed references to <code>opaxlattice_topology_cust</code> throughout document Moved Software Installation Checklists to the Appendices Updated Document Organization Updated OS RPMs Installation Prerequisites Updated MPI Compiler Prerequisites Updated Download the Intel Omni-Path Software Updated Using the INSTALL Command Line Options Updated SLES* Distribution Software Added new Chapter Install and Upgrade Standalone Firmware Moved UEFI Secure Boot Support to Part 1, Installing the Software Updated and moved Setting Up PXE Boot for Supported Operating Systems to Part 1, Installing the Software Moved Install Intel® Omni-Path Fabric Suite Fabric Manager GUI to Part 1, Installing the Software Updated Download the Fabric Manager Software Updated Prepare Fabric Manager for Fabric Manager GUI Added new section QSFP-DD Cables Updated table in Setting Up Overlapping Subnets
February 2018	8.2	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Made corrections in section OS RPMs Installation Prerequisites: was <code>libjson-c</code>, should be <code>libjson-c2</code>.
November 2017	8.1	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Updated OS RPMs Installation Prerequisites Updated Setting Up PXE Boot for Supported Operating Systems
October 2017	8.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Updated and added information to the Fabric Design Prerequisites section Added "Sandia-OpenSHMEM (hfi, gcc)" to the OPA-Basic Package section Updated OS RPMs Installation Prerequisites to add RHEL* 7.4, SLES* 12.3 and an example of a install command Added <code>sandiashmem</code> to the "Supported Component (comp) Names" list in Using the INSTALL Command Line Options section Updated Upgrade Prerequisites to include link to OS RPMs Installation Prerequisites Removed HFI UEFI PXE Installation and Configuration Added new Appendix, Setting Up PXE Boot for Supported Operating Systems Moved Upgrade the HFI UEFI Firmware under Setting Up PXE Boot for Supported Operating Systems and updated the section 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 Intel® Omni-Path Documentation Library for details.
August 2017	7.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none"> Globally, updated the following filepaths: <ul style="list-style-type: none"> from <code>/etc/sysconfig/opa</code> to <code>/etc/opa</code> Added high level installation flowchart to section Installation Overview and updated the information

continued...



Date	Revision	Description
		<ul style="list-style-type: none">Updated section Fabric Manager GUI PackagesUpdated section OS RPMs Installation PrerequisitesAdded section Download and Install NVIDIA* Software (Optional)Updated section Install the Intel® Omni-Path SoftwareUpdated section Install Using the Linux* Distribution SoftwareUpdated section Install Intel® Omni-Path Fabric Suite Fabric Manager
April 2017	6.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none">Document has been restructured and rewritten for usability.Globally, updated the following filepaths:<ul style="list-style-type: none">from /usr/lib/opa-fm/samples/ to /usr/share/opa-fm/samples/from /usr/lib/opa-fm/etc/opafm* to /usr/lib/opa-fm/bin/opafm*from /etc/sysconfig/opafm.xml to /etc/opa-fm/opafm.xmlfrom /etc/sysconfig/* to /etc/*from /usr/lib/opa/ to /usr/share/opa/from /usr/lib/opa/src/* to /usr/src/opa/*Added Intel® Omni-Path Documentation Library to Preface.Added new section Intel® Omni-Path Architecture Overview.Added new section Install Using the Linux* Distribution Software.Added new section Performance Tuning Prerequisites.Deprecated opaxlattopology_cust in Fabric Setup Prerequisites.Added new section Update the Fabric Manager opafm Configuration File.Added new section Thermal Monitoring Module Verification.Updated Installation Checklist to include configuration and verification pointers.
Dec 2016	5.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none">Removed all information pertaining to RHEL* 7.1Updated Prepare Fabric Manager for Fabric Manager GUIAdded new section Multi-Rail Usage to Chapter Additional Installation and Setup TasksMoved Appendix "Multi-Subnet Fabrics" under new Chapter Additional Installation and Setup TasksUpdated HFI UEFI PXE Installation and Configuration sectionRemoved Appendix hfidiags User Guide. Moved to <i>Intel® Omni-Path Fabric Host Software User Guide</i>.
Aug 2016	4.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none">Added information on UEFI firmware PrerequisitesUpdated Install TUI menusUpdated ./INSTALL command information to include --user-spaceUpdated Fabric Manager GUI Installation informationUpdated "HFI UEFI PXE Installation and Configuration" sectionUpdated "hfidiags User Guide" section
May 2016	3.0	<p>Document has been updated with the following:</p> <ul style="list-style-type: none">Updated OS RPMs Installation PrerequisitesAdded MPI Compiler PrerequisitesAdded information about minimum disk space requirement for the Fabric Manager GUI installationReplaced Client/Server Configuration to Boot Over Fabric section with "HFI UEFI PXE Installation and Configuration"
February 2016	2.0	<p>Document has been updated with the following:</p>
continued...		



Date	Revision	Description
		<ul style="list-style-type: none">Updated OS RPMs Installation PrerequisitesAdded Client/Server Configuration to Boot Over Fabric appendix
November 2015	1.0	Initial Release



Contents

Revision History.....	3
Preface.....	14
Intended Audience.....	14
Intel® Omni-Path Documentation Library.....	14
How to Search the Intel® Omni-Path Documentation Set.....	16
Cluster Configurator for Intel® Omni-Path Fabric.....	17
Documentation Conventions.....	17
Best Practices.....	18
License Agreements.....	18
Technical Support.....	18
1.0 Introduction.....	19
1.1 Document Organization.....	19
2.0 Overview.....	21
2.1 Intel® Omni-Path Architecture Overview.....	21
2.1.1 Host Fabric Interface.....	23
2.1.2 Intel® OPA Switches.....	23
2.1.3 Intel® OPA Management.....	24
2.2 Intel® Omni-Path Software Overview.....	24
2.3 Installation Overview.....	26
2.4 Installation Packages.....	27
2.4.1 OPA-Basic Package.....	27
2.4.2 OPA-IFS Package.....	27
2.4.3 Fabric Manager GUI Packages.....	28
Part 1: Installing the Software.....	29
3.0 Installation Getting Started.....	30
3.1 Pre-Installation Requirements.....	30
3.1.1 Fabric Design Prerequisites.....	30
3.1.2 Fabric Setup Prerequisites.....	31
3.1.3 OS RPMs Installation Prerequisites.....	34
3.1.4 MPI Compiler Prerequisites.....	34
3.1.5 Performance Tuning Prerequisites.....	34
3.2 Download the Intel® Omni-Path Software.....	34
3.3 Unpack the Tar File.....	35
3.4 Download and Install NVIDIA* Software (Optional).....	35
4.0 Install the Intel® Omni-Path Software.....	38
4.1 Before You Begin.....	38
4.1.1 About User Queries Settings.....	39
4.1.2 Using the INSTALL Command Line Options.....	39
4.2 Install Using the TUI Menus.....	43
4.3 Install Using CLI Commands.....	45
4.4 Install Using Linux* Distribution Software Packages Provided by Intel.....	46
4.4.1 Repositories Included in the IFS Package.....	47



4.4.2 Repository Deployment into the Environment.....	47
4.4.3 Using IFS Repository on Linux* OS.....	48
4.5 Install Using the Linux* Distribution Software Packages Provided by the OS Vendor.....	50
4.5.1 Using RHEL* Distribution Software.....	50
4.5.2 Using SLES* Distribution Software.....	51
4.6 Configure IPoIB IPV6.....	53
4.7 Reinstall Software After a Kernel Update.....	54
4.7.1 Reinstall Using TUI Menus.....	54
4.7.2 Reinstall Using CLI Commands.....	55
4.7.3 Reinstall Using Linux Distribution Software Packages Provided by Intel.....	55
5.0 Install and Upgrade Standalone Firmware.....	58
5.1 Before You Begin.....	58
5.2 Download the Firmware.....	58
5.3 Installing and Upgrading rpms.....	59
5.3.1 Upgrade the HFI UEFI Firmware.....	60
5.4 UEFI Secure Boot Support.....	62
5.4.1 UEFI Firmware Support.....	62
5.4.2 Linux* Support.....	62
5.5 Setting Up PXE Boot for Supported Operating Systems.....	63
5.5.1 PXE Server.....	64
5.5.1.1 Configure PXE Server with RHEL*.....	64
5.5.1.2 Configure PXE Server with SLES*.....	69
5.5.2 Client Node.....	75
5.5.2.1 Before You Begin.....	76
5.5.2.2 Install RHEL* over Intel® OP HFI.....	76
5.5.2.3 Install SLES* over Intel® OP HFI.....	78
6.0 Install Intel® Omni-Path Fabric Suite Fabric Manager GUI.....	82
6.1 Before You Begin.....	82
6.1.1 System Requirements for a Windows* Environment.....	82
6.1.2 System Requirements for a Linux* Environment.....	82
6.1.3 Download the Fabric Manager GUI Software.....	83
6.2 Install the Fabric Manager GUI Software on Windows*.....	83
6.3 Install the Fabric Manager GUI Software on Linux*.....	86
6.4 Prepare Fabric Manager for Fabric Manager GUI.....	87
6.5 Start the Fabric Manager GUI on Windows*.....	89
6.6 Start the Fabric Manager GUI on Linux*.....	89
6.7 Uninstall Fabric Manager GUI.....	90
6.7.1 Uninstall Fabric Manager GUI Software on Windows*.....	90
6.7.1.1 Clear the Fabric Manager GUI Cache on Windows*.....	91
6.7.2 Uninstall Fabric Manager GUI Software on Linux*.....	91
6.7.2.1 Clear the Fabric Manager GUI Cache on Linux*.....	92
Part 2: Configuring the Software.....	93
7.0 Configuration Getting Started.....	94
7.1 Pre-Configuration Requirements.....	94
7.1.1 Chassis Configuration Prerequisites.....	94
7.1.2 Switch Configuration Prerequisites.....	95
7.2 How to Use the FastFabric TUI.....	95



8.0 Configure the Chassis.....	98
8.1 Before You Begin.....	98
8.2 Configure the Chassis Using the FastFabric TUI Menu.....	99
8.3 Configure the Chassis Using the FastFabric CLI Commands.....	104
9.0 Configure the Externally-Managed Switches.....	108
9.1 Before You Begin.....	108
9.2 Configure the Externally-Managed Switches Using FastFabric TUI Menu.....	109
9.3 Configure the Externally-Managed Switches Using FastFabric CLI Commands.....	111
10.0 Install Host Software on Remaining Servers.....	113
10.1 Before You Begin.....	113
10.2 Install the Host Software on the Remaining Hosts Using the FastFabric TUI Menu.....	115
10.3 Verify the Host Software on the Remaining Servers Using the FastFabric TUI Menu.....	118
10.4 Install the Host Software on the Remaining Hosts Using the FastFabric CLI Commands.....	121
10.5 Verify the Host Software on the Remaining Servers Using CLI Commands.....	123
11.0 Set Up Additional Management Nodes.....	127
11.1 Before You Begin.....	127
11.2 Configure Additional Management Nodes Using FastFabric TUI Menus.....	127
11.3 Configure Additional Management Nodes Using FastFabric CLI Commands.....	129
12.0 Perform Initial Health Check.....	130
12.1 Before You Begin.....	130
12.2 Configure and Initialize Health Check Tools Using FastFabric TUI Menu.....	131
12.3 Configure and Initialize Health Check Tools Using FastFabric CLI Commands.....	132
13.0 Perform High Performance Linpack Benchmark	134
13.1 Before You Begin.....	135
13.2 Run a Sample High Performance Linpack 2 Configuration.....	135
13.3 Generate and Run a Custom High Performance Linpack Configuration.....	136
14.0 Additional Installation and Setup Tasks.....	138
14.1 QSFP-DD Cables.....	138
14.2 Multi-Subnet Fabrics.....	138
14.2.1 Setting Up Subnets.....	139
14.3 Multi-Rail Usage.....	141
14.3.1 Multi-Rail Overview.....	141
14.3.2 Setting Up Dual Rails for a Single Subnet.....	143
14.3.3 Configuring Dual Rails for Dual Subnets.....	145
14.3.4 IP Routing Guidance.....	147
15.0 Installation Verification and Additional Settings.....	149
15.1 LED Link and Data Indicators.....	149
15.2 Thermal Monitoring Module Verification.....	149
15.3 Adapter and Other Settings.....	150
15.4 ARP Neighbor Table Setup for Large Clusters.....	150
15.5 SM Loop Test.....	151



Part 3: Upgrading the Software.....	152
16.0 Upgrade Getting Started.....	153
16.1 Upgrade Prerequisites.....	153
16.2 Download the Intel® Omni-Path Software.....	153
16.3 Unpack the Tar File.....	154
17.0 Upgrade the Intel® Omni-Path Software.....	155
17.1 Before You Begin.....	155
17.2 Upgrade Scenarios.....	155
17.3 Upgrade the OPA-Basic Software.....	155
17.4 Upgrade the OPA-IFS Software.....	156
17.4.1 Stopping the Fabric Manager.....	157
17.4.2 Upgrade the Software on the Management Nodes.....	157
17.4.3 Update the Fabric Manager opafm Configuration File.....	158
17.4.4 Starting the Fabric Manager.....	159
17.4.5 Upgrade the Software on the Remaining Servers.....	159
17.5 Upgrade from OPA-Basic to OPA-IFS Software Package.....	161
Appendix A Software Installation Checklists.....	163
A.1 OPA-Basic Checklists.....	163
A.1.1 OPA-Basic Pre-Installation Checklist.....	163
A.1.2 Install the OPA-Basic Software Checklist.....	163
A.1.3 Upgrade the OPA-Basic Software Checklist.....	164
A.2 OPA-IFS Checklists.....	164
A.2.1 OPA-IFS Pre-Installation Checklist.....	164
A.2.2 Install and Configure the OPA-IFS Software Checklist.....	164
A.2.3 Upgrade the OPA-IFS Software Checklist.....	165
Appendix B Intel® OPA Software Components to Packages Mapping.....	166



Figures

1	Intel® OPA Fabric.....	22
2	Intel® OPA Building Blocks.....	23
3	Intel® OPA Fabric and Software Components.....	25



Tables

1	BIOS versus Boot Mode.....	75
2	Intel® OPA Software Components Mapping.....	166



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 the next generation, High-Performance Computing (HPC) network solution that is designed to cost-effectively meet the growth, density, and reliability requirements of large-scale HPC clusters.

Both the Intel® OP Fabric and standard InfiniBand* (IB) are able to send Internet Protocol (IP) traffic over the fabric, or *IPoFabric*. In this document, however, it may also be referred to as *IP over IB* or *IPoIB*. From a software point of view, IPoFabric behaves the same way as IPoIB, and in fact uses an `ib_ipoib` driver to send IP traffic over the `ib0/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 URL, under *Latest Release Library*:

<https://www.intel.com/content/www/us/en/design/products-and-solutions/networking-and-io/fabric-products/omni-path/downloads.html>

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

Task	Document Title	Description
Using the Intel® OPA documentation set	<i>Intel® Omni-Path Fabric Quick Start Guide</i>	A roadmap to Intel's comprehensive library of publications describing all aspects of the product family. This document outlines the most basic steps for getting your Intel® Omni-Path Architecture (Intel® OPA) cluster installed and operational.
Setting up an Intel® OPA cluster	<i>Intel® Omni-Path Fabric Setup 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 best known methods are defined to facilitate the installation process and troubleshooting.
continued...		



Task	Document Title	Description
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.
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 graphical user interface (GUI) of the Intel® Omni-Path Fabric Chassis Viewer GUI. This document provides task-oriented procedures for configuring and managing the Intel® Omni-Path Switch family. Help: GUI embedded help files
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>	Provides instructions for using the set of fabric management tools designed to simplify and optimize common fabric management tasks. The management tools consist of Text-based User Interface (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. This document is a user-friendly alternative to traditional command-line tools for day-to-day monitoring of fabric health. Help: Fabric Manager GUI embedded help files
Configuring and administering Intel® HFI and IPoIB driver Running MPI applications on Intel® OPA	<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 cluster administrators and Message-Passing Interface (MPI) application programmers.
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.
continued...		



Task	Document Title	Description
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 LNet router on Intel® OPA	<i>Intel® Omni-Path IP and LNet 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 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.
Using NVMe* over Fabrics on Intel® OPA	<i>Configuring Non-Volatile Memory Express* (NVMe*) over Fabrics on Intel® Omni-Path Architecture Application Note</i>	Describes how to implement a simple Intel® Omni-Path Architecture-based point-to-point configuration with one target and one host server.
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)	
	<i>Intel® Omni-Path Fabric Unified Extensible Firmware Interface (UEFI) Release Notes</i>	
	<i>Intel® Omni-Path Fabric Thermal Management Microchip (TMM) Release Notes</i>	
	<i>Intel® Omni-Path Fabric Firmware Tools Release Notes</i>	

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 18 for more information.
For more information, visit www.intel.com.
- Text in **bold** font indicates user interface elements such as menu items, buttons, check boxes, key names, key strokes, or column headings. For example:
Click the **Start** button, point to **Programs**, point to **Accessories**, and then click **Command Prompt**.
Press **CTRL+P** and then press the **UP ARROW** key.
- Text in `Courier` font indicates a file name, directory path, or command line text. For example:
Enter the following command: `sh ./install.bin`
- Text in *italics* indicates terms, emphasis, variables, or document titles. For example:
Refer to *Intel® Omni-Path Fabric Software Installation Guide* for details.
In this document, the term *chassis* refers to a managed switch.

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

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



Best Practices

- Intel recommends that users update to the latest versions of Intel® Omni-Path firmware and software to obtain the most recent functional and security updates.
- To improve security, the administrator should log out users and disable multi-user logins prior to performing provisioning and similar tasks.

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 138.

1.1 Document Organization

This manual is organized as follows:

- Introduction
- Overview

Part 1: Installing the Software

- Installation Getting Started
- Install Intel® Omni-Path Software
- Install and Upgrade Standalone Firmware
- Install Intel® Omni-Path Fabric Suite Fabric Manager GUI

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



- 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

- Software Installation Checklists
- Intel® OPA Software Components to Packages Mapping



2.0 Overview

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

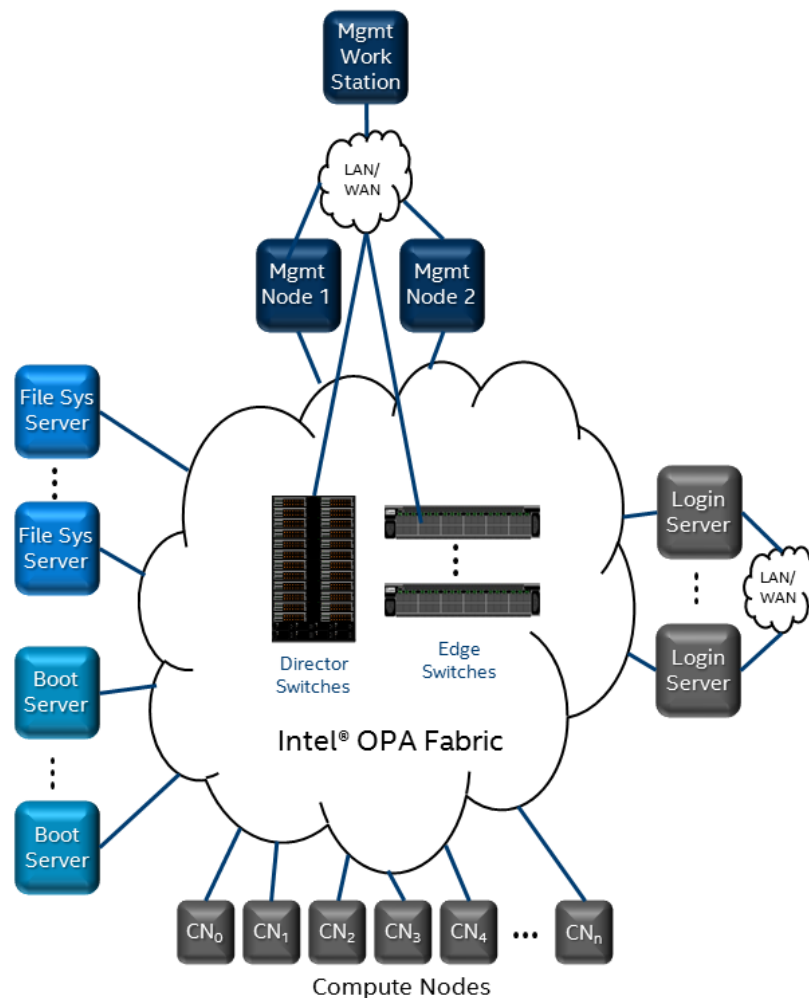
2.1 Intel® Omni-Path Architecture Overview

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. 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. Intel® OPA also focuses on HPC needs, 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 provides traffic isolation between jobs or between users.

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

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

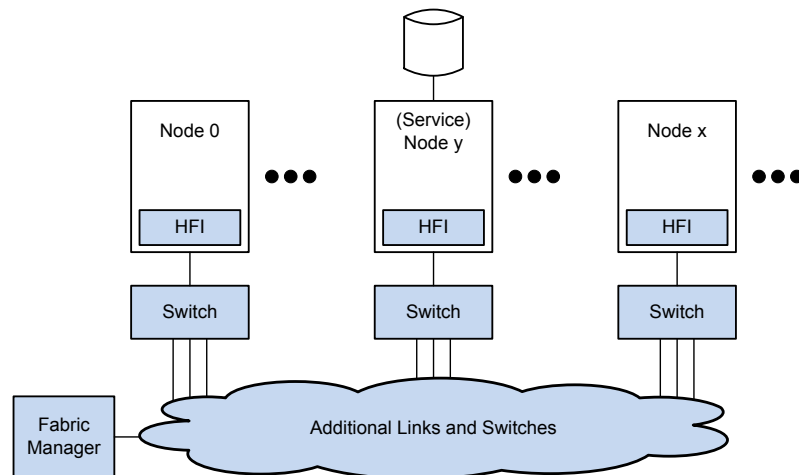


2. The Performance Scaled Messaging 2 (PSM2) API provides HPC-focused transports and an evolutionary software path from the Intel® True Scale Fabric.
3. OFA Verbs provides support for existing remote direct memory access (RDMA) applications and includes extensions to support Intel® OPA fabric management.
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 combines the Intel® Omni-Path Host Fabric Interfaces (HFIs), Intel® Omni-Path switches, and fabric management and development tools into an end-to-end solution. 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 supports redundant Fabric Managers that centrally 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.

The Primary Fabric Manager sends management packets over the fabric. These packets are sent *in-band* over the same wires as regular Intel® Omni-Path packets using dedicated buffers on a specific virtual lane (VL15). End-to-end reliability protocols detect lost packets.

2.2 Intel® Omni-Path Software Overview

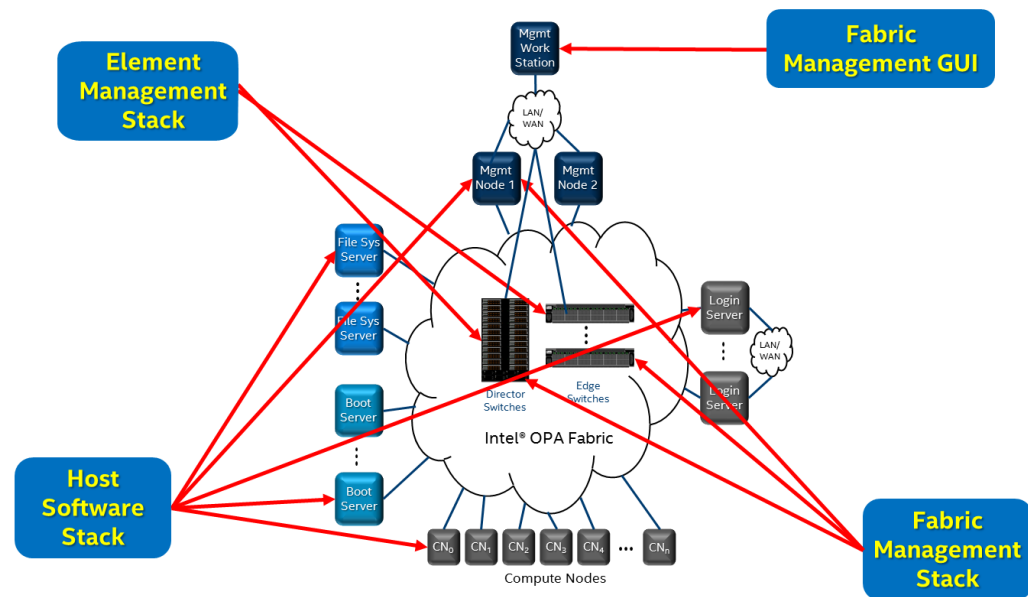
For software applications, Intel® OPA maintains consistency and compatibility with existing Intel® True Scale Fabric and InfiniBand* APIs through the open source OpenFabrics Alliance* (OFA) software stack on Linux* distribution releases.

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

Embedded Management Stack

- 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 (elements).
- Provides system management capabilities, including signal integrity, thermal monitoring, and voltage monitoring, among others.
- Accessed via Ethernet* port using command line interface (CLI) or graphical user interface (GUI).

User documents:

- *Intel® Omni-Path Fabric Switches GUI User Guide*
- *Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide*

Host Software Stack

- Runs on all Intel® OPA-connected host nodes and supports compute, management, and I/O nodes.
- Provides a rich set of APIs including OFI, PSM2, sockets, and OFA verbs.
- Provides high performance, highly scalable MPI implementation via OFA, PSM2, and an extensive set of upper layer protocols.
- 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.

User documents:

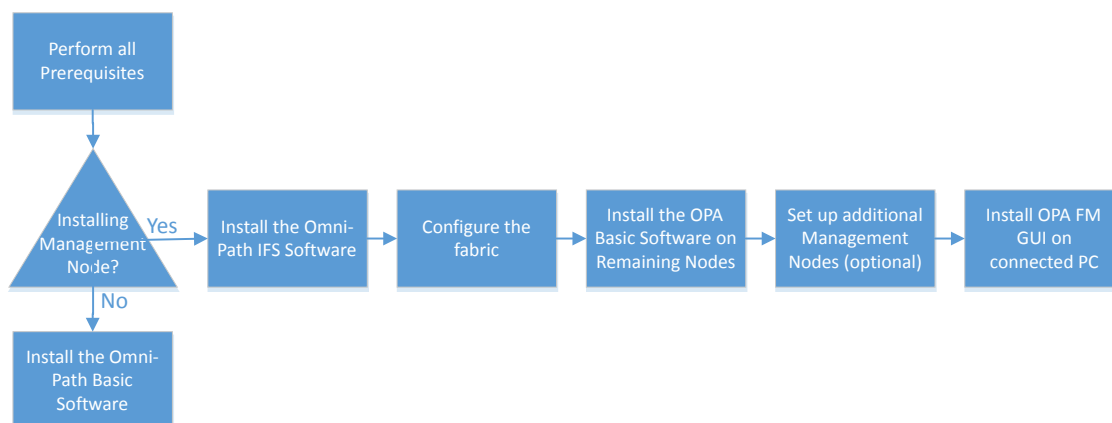
- *Intel® Omni-Path Fabric Host Software User Guide*

continued...

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

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)
- Open MPI (hfi, 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.



Part 1: Installing the Software



3.0 Installation Getting Started

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

3.1 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.

3.1.1 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 75 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.
- 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 `.xlsx` files installed into `/usr/share/opa/samples/` on the management node. After a plan is established, `opaxlattopology` or `opaxltopology_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** section 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.

3.1.2 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`, `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 Host Software on Remaining Servers 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](#).

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 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.

Related Links

[Install Host Software on Remaining Servers](#) on page 113

[Configure the Chassis](#) on page 98

3.1.3 OS RPMs Installation Prerequisites

If you are using the `INSTALL` script for installation, OS RPMs must be installed before you can install the Intel® Omni-Path software.

Refer to the *Intel® Omni-Path Fabric Software Release Notes* for the list of OS RPMs required for each supported OS in this release.

3.1.4 MPI Compiler Prerequisites

MVAPICH2 and Open MPI that are downloaded with the OPA-Basic Software package are selectable in the Intel OPA Install Menu. These two MPIs were built with Intel compilers. To use either of these MPIs for compiling and running your applications and benchmarks, you need to install the latest version of the Intel® Parallel Studio XE 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/intel-compilers-redistributable-libraries-by-version>

3.1.5 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.

3.2 Download the Intel® Omni-Path Software

If the OS you installed did not include the OPA RPMs, download the software package from the Intel Resource & Design Center using the following procedures.

1. Using a web browser, go to <https://www.intel.com/content/www/us/en/design/products-and-solutions/networking-and-io/fabric-products/omni-path/downloads.html>.
2. Click on Latest Release Library.



3. In the table, click the drop-down arrow for Software and Drivers to select the files you need for the OS you have installed on your fabric.

Or, use the search box above the table to enter your search criteria.

- For OPA-Basic software, select:

```
Intel® Omni-Path Basic Software - DISTRO.VERSION - OPA release
```

- For OPA-IFS software, select:

```
Intel® Omni-Path IFS Software - DISTRO.VERSION - OPA release
```

NOTE

DISTRO.VERSION refers to the distribution and CPU.

4. Locate your target software and click the **Download** button.
5. Review the Intel Software License Agreement.
6. Click **Accept**.

The zipped software is saved to your computer.

3.3 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
```

3.4 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.



IMPORTANT

You must perform step 2 below to avoid issues where the HFI GPUDirect*-enabled driver may not be able to load due to missing symbols. Failure to perform this step will require that you uninstall and then reinstall the driver using the proper instructions.

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/>

Verifying GPUDirect* is Installed

To ensure that the GPUDirect* capable software is installed, verify the following (note bold text in sample outputs):

- For PSM, run:

```
$ rpm -qa | grep psm | grep cuda
```



Sample output:

```
libpsm2-devel-11.2.89-1cuda.x86_64  
libpsm2-debuginfo-11.2.89-1cuda.x86_64  
libpsm2-compat-11.2.89-1cuda.x86_64  
libpsm2-11.2.89-1cuda.x86_64
```

- For MPI, run:

```
$ /usr/mpi/gcc/openmpi-3.1.4-cuda-hfi/bin/ompi_info -a | grep  
"mpi_built_with_cuda_support" | grep true
```

Sample output:

```
MCA opal base: informational "opal_built_with_cuda_support" (current value:  
"true", data source: default, level: 4 tuner/basic, type: bool, synonyms:  
mpi_built_with_cuda_support)  
MCA mpi base: parameter "mpi_built_with_cuda_support" (current value: "true",  
data source: default, level: 4 tuner/basic, type: bool, synonym of:  
opal_built_with_cuda_support)
```

- For Driver, run:

```
$ modinfo hfi1 | grep gpu-direct
```

Sample output:

```
version: 10.10.1.0 gpu-direct
```



4.0 Install the Intel® Omni-Path Software

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

You install the software using one of the following methods:

- TUI menus (recommended)
- CLI commands
- Linux* Distribution Software packages provided by Intel
- Linux* Distribution Software packages provided by the OS vendor

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 in the Appendices to track your installation of the software.

Related Links

[Install Using the TUI Menus](#) on page 43

[Install Using CLI Commands](#) on page 45

[Install Using Linux* Distribution Software Packages Provided by Intel](#) on page 46

[Install Using the Linux* Distribution Software Packages Provided by the OS Vendor](#) on page 50

4.1 Before You Begin

Before starting the installation, perform the following:

- Refer to the *Intel® Omni-Path Fabric Software Release Notes* for the list of compatible operating systems and required OS RPMs Installation Prerequisites.
- 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 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.

4.1.1 About User Queries Settings

During installation, the default user queries setting permits non-root users and tools to query the fabric. This setting enables read/write permissions on Userspace Management Datagrams (UMADs). The UMAD interface allows Subnet Administration (SA) queries, Performance Administration (PA) queries, Subnet Management Agent (SMA) queries, and Performance Management Agent (PMA) queries from user space applications.

NOTE

Allowing non-root users read/write access to issue queries can be a security risk as non-root users can potentially send MADs to change the fabric state. Also, given the nature of the UMAD interface, sending MADs and receiving the corresponding responses require read/write permission, both for queries and state changes. Setting the interface to read-only for non-root users disables the ability to issue queries.

In Intel® OPA, the *MgmtAllowed* configuration setting in the switches can be used to prevent a compute node from accessing or altering the fabric configuration. A compute node without *MgmtAllowed* cannot issue SMA and PMA transactions to other nodes, thus protecting the fabric from malicious or buggy application code. Such a node will also be restricted in its access to SA information and will not be permitted to access PA information. However, when *umad* is enabled for local use, a malicious application could potentially alter its own SMA or PMA settings, for example, taking the node's port down or observing or clearing PMA counters.

Many of the Intel® OPA basic tools make use of UMAD. Intel® Omni-Path Fabric Suite FastFabric tools are typically run on a management node and frequently issue SA, PA, SMA, and PMA operations; however they are often run as root.

NOTE

The user query setting you choose should depend on whether there are other tools or applications run as non-root that need access to UMAD and whether non-root users are trusted. For example, some non-root applications may use UMAD to issue SA queries for *name services* style queries.

4.1.2 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 [-v|-vv] -R osver -B osver [-a|-n|-U|-u|-s|-O|-N|-i comp |
-e comp] [-E comp] [-D comp] [--user-space] [--without-depcheck]
[--rebuild] [--force] [--answer keyword=value]
```



or

```
./INSTALL -C
```

or

```
./INSTALL -V
```

Options

No option selected	Displays the Intel OPA Software TUI.
-v	Provides verbose logging. Logs to the <code>/var/log/opa.log</code> file.
-vv	Provides very verbose debug logging. Logs to the <code>/var/log/opa.log</code> file.
-R <i>osver</i>	Force install for specific OS kernel version, rather than running kernel.
-B <i>osver</i>	Run build for all components targetting specific OS kernel version.
-a	Installs all Upper Layer Protocols (ULP) and drivers with the default options, including UEFI firmware.
<hr/> NOTE This rebuilds the RPMs except for CUDA* which is not rebuilt unless <code>-G</code> is specified. <hr/>	
-n	Installs all ULPs and drivers with the default options, but does not change the autostart options.
-U	Upgrades/reinstalls all presently installed ULPs and drivers with the default options, and does not change the autostart options.
-u	Uninstalls all ULPs and drivers with the default options.
-s	Enables autostart for all installed software.
-O	Keeps the current modified rpm configuration file.
-N	Uses a new default rpm configuration file.
-i <i>comp</i>	Installs the given component with the default options. This option can appear multiple times on a command line.



IMPORTANT

Using this command to upgrade or downgrade an individual component against the existing IFS will update all previously installed components to the version of the individual component being installed.

`-e comp`

Uninstalls the given component with the default options. This option can appear multiple times on a command line.

`-E comp`

Enables autostart of given component. This option can appear with `-D` or multiple times on a command line.

NOTE

To control which installed software is configured for autostart, combine this option with `-a`, `-n`, `-i`, `-e`, and `-U` options.

`-D comp`

Disables autostart of the given component. This option can appear with `-E` or multiple times on a command line.

NOTE

To control which installed software is configured for autostart, combine this option with `-a`, `-n`, `-i`, `-e`, and `-U` options.

`--user-space`

Skips kernel space components during installation.

`--without-depcheck`

Disables the check of OS dependencies.

`--rebuild`

Forces a rebuild of OFA_DELTA srpms.

`--force`

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. Refer to [Other Information](#).

NOTE

Intel® Omni-Path is compatible with GPUDirect* for RHEL* 7.6, SLES* 12 SP3, SLES* 12 SP4, and SLES* 15.

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

NOTE

Under some OS versions (for example, SLES* 12 SP4), non-root users may need to add `/usr/sbin` to their `$PATH` in order to execute query commands.

- `--no_user_queries` Specifies that 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,
rdma_ndd, delta_ipoib, delta_srp, delta_srpt, opafm, opamgt_sdk,
mvapich2_gcc_hfi,
openmpi_gcc_hfi, openmpi_gcc_cuda_hfi, mpisrc, hfi1_uefi, delta_debug
```

Supported Component (comp) Name Aliases:

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

Components for -G (CUDA-based) Installations:

For RHEL*:

```
ifs-kernel-updates-devel, kmod-ifs-kernel-updates, ifs-kernel-updates-debuginfo,
libpsm2, libpsm2-debuginfo, libpsm2-compat, libpsm2-devel
```

For SLES*:

```
ifs-kernel-updates-devel, ifs-kernel-updates-kmp-default, libpsm2,
libpsm2-compat, libpsm2-devel
```

4.2 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.

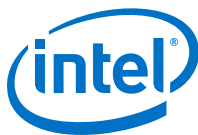
Procedures

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

CAUTION

Do not interrupt an operation mid-process. Some operations may take a few minutes to complete.

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"> • For Basic, type the following and press Enter: <code>cd IntelOPA-Basic.DISTRO.VERSION</code> • For IFS, type the following and press Enter: <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 Enter .
continued...		



Step	Task/Prompt	Action
		<i>Notes:</i> <ul style="list-style-type: none">To install IFS with different root directory, use <code>chroot</code>.When the kernel version in the <code>chroot</code> environment is different from the host's kernel version, use <code>./INSTALL -R</code> to force the IFS installation with the target OS kernel version.
3.	Select 1) Install/Uninstall Software.	Type 1 .
4.	Review the items to be installed.	Accept the defaults (No action required). Type N to go to the next page. NOTE: If you need to change any item, enter the alphanumeric character associated with the item to toggle between <code>Install</code> or <code>Don't Install</code> .
5.	Start the installation.	Type P to perform the actions. <i>Note:</i> This may take a few minutes.
6.	Preparing OFA <i>VERSION</i> release for Install... Rebuild OFA SRPMs (a=all, p=prompt per SRPM, n=only as needed?) [n]:	Press Enter to accept the default. NOTE: The system will display prompts that require your response throughout the installation.
7.	For each system prompt...	Accept the defaults by pressing Enter to continue. <i>Note:</i> Some of the default processes may take a few minutes to complete.
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. NOTE: If you need to change any item, enter the alphanumeric character associated with the item to toggle between <code>Enable</code> or <code>Disable</code> .
9.	Run the OPA Autostart operations.	Type P .
10.	For each system prompt, "Hit any key to continue..."	Press any key. NOTE: When the installation completes, you are returned to the main menu.
Set up IPoIB IPV4 Configuration		
11.	Select 2) Reconfigure OFA IP over IB.	Type 2 .
12.	Configure OFA IP over IB IPV4 addresses now? [n]:	Type y .
13.	How many OFA IP over IB interfaces would you like to configure? [1]:	Type the number of interfaces and press Enter .
14.	For each interface, provide the following information: <ul style="list-style-type: none">OFA IP over IB interface nameIPV4 address in dot notation (or dhcp)IPV4 netmask in dot notation	Provide the required information and press Enter . NOTE: After setup is complete, you are returned to the main menu.
15.	Exit out of the TUI to the command prompt.	Type X .
continued...		



Step	Task/Prompt	Action
16.	Reboot the server.	Type <code>reboot</code> and press Enter . Caution: Please do not interrupt the reboot process. Depending on your operating system, the reboot may take a few minutes.
17.	Verify the installation was successful.	Type <code>opaconfig -V</code> and press Enter .
	End Task	

Related Links

[Configure IPoIB IPV6](#) on page 53

[Upgrade the HFI UEFI Firmware](#) on page 60

[Configuring the Software](#) on page 93

4.3 Install Using CLI Commands

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

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.

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"> For Basic, type the following and press Enter: <code>cd IntelOPA-Basic.DISTRO.VERSION</code> For IFS, type the following and press Enter: <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 -n</code> and press Enter . Note: <ul style="list-style-type: none"> To install IFS with different root directory, use <code>chroot</code>. When the kernel version in the <code>chroot</code> environment is different from the host's kernel version, use <code>./INSTALL -R</code> to force the IFS installation with the target OS kernel version.
3.	At the command prompt, reboot the server.	Type <code>reboot</code> and press Enter .
Set up IPoIB IPV4 Configuration		<i>continued...</i>



Step	Task/Prompt	Action
4.	Manually edit or create the <code>/etc/sysconfig/network-scripts/ifcfg-ibX</code> file.	<p>Note: Use the OS distribution-supplied instructions for setting up network interfaces.</p> <p>For example (RHEL*):</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</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 Enter .
	End Task	

Related Links

[Using the INSTALL Command Line Options](#) on page 39

[Configure IPoIB IPV6](#) on page 53

[Upgrade the HFI UEFI Firmware](#) on page 60

[Configuring the Software](#) on page 93

4.4 Install Using Linux* Distribution Software Packages Provided by Intel

The Intel® Omni-Path Software IFS package contains the OS-specific repository for installing the Intel® OPA software.

This section provides the instructions for installing using the IFS package repository.

Since 10.10.1, Intel introduced virtual packages to facilitate IFS installation.

- A virtual package prefixed with `opameta_` is a meta-package for an IFS component in the `INSTALL` script. Installing a meta-package will install the corresponding component.
- A virtual package prefixed with `opanode_` is an alias package for a typical IFS installation on an HPC node.

Default Installation Options

This installation method will install OPA packages with default options. To install with different options, set the following system environment variables in advance of the installation.

- Variable Name: `OPA_UDEV_RULES`
Values:
 - 1 - Non root user can query the fabric (default)
 - 0 - Only root user can query the fabric
- Variable Name: `OPA_ARPTABLE_TUNING`
Values:
 - 1 - Enable adjust kernel ARP table for large fabric (default)



0 - Disable adjust kernel ARP table for large fabric

Alternatively, you can use `opasystemconfig` to change the options after installation. Refer to *Intel® Omni-Path Fabric Host Software User Guide* for the details.

4.4.1 Repositories Included in the IFS Package

The IFS package contains two repositories:

- OPA_PKGS – Contains all software needed to be installed on the compute node, management node, or service node, such as storage node.
- OPA_PKGS_CUDA – Contains all software needed to be installed on the node that includes NVIDIA* cards.

NOTE

These two repositories cannot coexist on a node. Ensure that only one exists or is enabled.

Interoperate with the INSTALL Script

Intel recommends that you **do not** mix yum/zypper repository-based install with script-based install. Doing so may cause unexpected behaviors. However, you can switch from one install mechanism to another.

- Switching from script-based install to yum/zypper-based install:
Customer can switch to the yum/zypper-based install at any time. No special actions are required.
To switch to a yum/zypper-based install: If Intel® OPA is already partially or fully installed with the script, the yum/zypper-based install will identify the installed packages and skip them during installation.
- Switching from yum/zypper-based install to script-based install:
The meta and alias packages of the yum/zypper-based install introduce extra dependencies on Intel® OPA packages. This could impact the script-based install because the code directly uses the rpm command for installation which is sensitive to package dependencies.
To switch to script-based install: Since you must remove the meta and alias packages first, Intel has improved the INSTALL script to handle this. Using `INSTALL -a`, or `-U`, or `-n` will switch to script-based installation. `INSTALL -u` will remove all packages include the meta and alias packages. Alternatively, you can manually remove the meta and alias packages with the yum/zypper command prior to starting the script-based install.

4.4.2 Repository Deployment into the Environment

The `IntelOPA-IFS.<OS_VERSION>-x86_64.<VERSION>.tgz` tar package contains the repository used to install the Intel® OPA software. It also includes a helper script called `opacreaterepo` that checks and rebuilds kernel rpms, creates the local repository, and generates the recommended packages to install on each compute, management, and service nodes. Intel recommends using this script to create the local repository ensuring that the correct version kernel rpms and the proper repository is created for the GPUDirect* (CUDA) support requirement.



For example, if an OPA_PKGS_CUDA repository already exists and you want to replace it with a repository for OPA_PKGS, the script will back up and remove the OPA_PKGS_CUDA to prevent the CUDA version packages from installing unintentionally.

Once a local repository has been successfully created, you can transfer it to an enterprise repository, based on organization needs, to allow sharing it among nodes.

The following shows the usage information for `opacreaterepo`.

```
Usage: opacreaterepo [-G]
       opacreaterepo -i
       opacreaterepo --help

Create a local repo for Intel OPA IFS packages.

Options:
  -G      create a repo with GPU Direct support (to install it
         must have NVidia driver installed)
  -i      display information about the repo it will create

  --help  produce full help text

Examples:
  opacreaterepo
  opacreaterepo -G
  opacreaterepo -i
```

Once the script executes successfully, it will recommend the installation commands. To use these commands, you copy and paste them into a command prompt, or save them to be edited prior to executing.

The example below shows the output for RHEL:

```
Repo IntelOPA-IFS was successfully created.

Please use the following component metapackages to install Intel OPA software

opameta_fastfabric      : Intel Omni-Path Meta Package for FastFabric
opameta_fastfabric_userspace : Intel Omni-Path Meta Package for FastFabric (user space only)
opameta_intel_hfi       : Intel Omni-Path Meta Package for Intel HFI Components
opameta_intel_hfi_userspace : Intel Omni-Path Meta Package for Intel HFI Components (user space only)
opameta_mpi_src         : Intel Omni-Path Meta Package for MPI Source
opameta_mpi_src_userspace : Intel Omni-Path Meta Package for MPI Source (user space only)
opameta_mvapich2_gcc_hfi : Intel Omni-Path Meta Package for MVAPICH2 (hfi,gcc)
opameta_mvapich2_gcc_hfi_userspace : Intel Omni-Path Meta Package for MVAPICH2 (hfi,gcc) (user space only)
opameta_oftools         : Intel Omni-Path Meta Package for OPA Tools
opameta_oftools_userspace : Intel Omni-Path Meta Package for OPA Tools (user space only)
opameta_opafm           : Intel Omni-Path Meta Package for OPA FM
opameta_opafm_userspace : Intel Omni-Path Meta Package for OPA FM (user space only)
opameta_opamgt_sdk      : Intel Omni-Path Meta Package for OPA Management SDK
opameta_opamgt_sdk_userspace : Intel Omni-Path Meta Package for OPA Management SDK (user space only)
opameta_opa_stack       : Intel Omni-Path Meta Package for OPA OPA Stack
opameta_opa_stack_dev   : Intel Omni-Path Meta Package for OPA OPA Development
opameta_opa_stack_dev_userspace : Intel Omni-Path Meta Package for OPA OPA Development (user space only)
opameta_opa_stack_userspace : Intel Omni-Path Meta Package for OPA OPA Stack (user space only)
opameta_openmpi_gcc_hfi : Intel Omni-Path Meta Package for OpenMPI (hfi,gcc)
opameta_openmpi_gcc_hfi_userspace : Intel Omni-Path Meta Package for OpenMPI (hfi,gcc) (user space only)

To facilitate installation, Intel provides the following aliases for common component combinations:

opanode_mgmt      : Useful for management node. Includes all components.
opanode_mgmt_userspace : Useful for container. Same as opa_mgmt expect it's using user space version components.
opanode_compute   : Useful for compute node. Includes all components except opafm and fastfabric that are for management
opanode_compute_userspace : Useful for container. Same as opa_compute expect it's using user space version components.
opanode_service   : Useful for service node. Includes all components except opafm, fastfabric and mpi components.
opanode_service_userspace : Useful for container. Same as opa_service expect it's using user space version components.
```

4.4.3 Using IFS Repository on Linux* OS

You can install the Intel® Omni-Path Software packages on Red Hat* Enterprise Linux* (RHEL*) or SUSE* Linux* Enterprise Server (SLES*) using the OS distribution included in the IFS package repository and its dependencies.



Assumptions

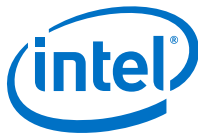
- You are logged in.
- You have your IPV4 list of IP addresses and netmasks for each interface you are going to set up.
- You have your software packages ready for installation.

Refer to *Intel® Omni-Path Fabric Software Release Notes* for supported OS packages.

Procedure

Perform the following steps to install the default Intel® Omni-Path Software configuration:

Step	Task/Prompt	Action
Set Up the IFS Repository		
1.	Create the local repository.	At the command prompt, type: <code>opacreaterepo</code> .
2.	Create the local repository on nodes that need CUDA* support. <i>Note:</i> If you install CUDA*-supported packages on nodes without NVIDIA* cards, you may see performance degradation.	At the command prompt, type: <code>opacreaterepo -G</code> . <i>Note:</i> After execution, the recommended install commands are provided.
3.	On each node, install the Intel® OPA Software. For a list of packages in specific Intel® OPA components, refer to Intel® OPA Software Components to Packages Mapping .	Copy the recommended install command (based on node type), paste it in the command prompt, and press Enter . <i>Note:</i> Alternatively, you can insert the recommended install command (based on node type) in a provision script.
Set up IPoIB IPV4 Configuration		
4.	At the command prompt, start <code>opaconfig</code> .	Type <code>opaconfig</code> .
5.	Select 2) Reconfigure OFA IP over IB.	Type 2 .
6.	Configure OFA IP over IB IPV4 addresses now? [n]:	Type y .
7.	How many OFA IP over IB interfaces would you like to configure? [1]:	Type the number of interfaces and press Enter .
8.	For each interface, provide the following information: <ul style="list-style-type: none"> • OFA IP over IB interface name • IPV4 address in dot notation (or dhcp) • IPV4 netmask in dot notation NOTE: To configure datagram mode for Accelerated IP (AIP), refer to Using RHEL* Distribution Software on page 50, Set up IPoIB IPV4 Configuration instructions.	Provide the required information and press Enter . NOTE: After setup is complete, you are returned to the main menu.
9.	Exit out of the TUI to the command prompt.	Type X .
10.	Reboot the server.	Type <code>reboot</code> and press Enter . Caution: Please do not interrupt the reboot process. Depending on your operating system, the reboot may take a few minutes.
11.	Verify the installation was successful.	Type <code>opaconfig -V</code> and press Enter .
End Task		



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](#).

4.5 Install Using the Linux* Distribution Software Packages Provided by the OS Vendor

The Intel® Omni-Path Software package is included with the Red Hat* Enterprise Linux* (RHEL*) and SUSE* Linux* Enterprise Server (SLES*) OS distributions.

This section provides the installation instructions for each OS.

4.5.1 Using RHEL* Distribution Software

You can install the Intel® Omni-Path Software package that is included with the Red Hat* Enterprise Linux* (RHEL*) 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® Omni-Path Software configuration using RHEL* OS:

Step	Task/Prompt	Action
Install OPA-Basic Software		
1.	At the command prompt, enter the installation command for opa-basic-tools.	Type yum install -y opa-basic-tools and press Enter .
2.	At the command prompt, reboot the server.	Type reboot and press Enter .
3.	Check your link using opainfo.	Type opainfo and press Enter . 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: 251 0 MB Pkts: 251 0 MB Recv Data: 251 0 MB Pkts: 251 0 MB Link Quality: 5 (Excellent)</pre>
4.	Install the rdma-core rpm.	Type yum install -y rdma-core and press Enter .
5.	On all compute nodes: install the PSM2 library.	Type yum install -y libpsm2 and press Enter .
Install Intel® Omni-Path Fabric Suite Components on the Management Node		
<i>continued...</i>		



Step	Task/Prompt	Action
6.	Install FastFabric.	Type <code>yum install -y opa-fastfabric</code> and press Enter .
7.	Install the <code>opa-address-resolution</code> rpm on all nodes.	Type <code>yum install -y opa-address-resolution</code> and press Enter .
8.	Install Fabric Manager.	Type <code>yum install -y opa-fm</code> and press Enter .
9.	Start the Fabric Manager.	Type <code>systemctl start opafm</code> and press Enter .
Set up IPoIB IPV4 Configuration		
10.	Manually edit or create the <code>ifcfg-ibX</code> file.	<p><i>Note:</i> Use the OS distribution-supplied instructions for setting up network interfaces.</p> <p>Type <code>cat /etc/network-scripts/ifcfg-ib0</code> and press Enter.</p> <p>Example output:</p> <pre>DEVICE=ib0 TYPE=infiniband BOOTPROTO=static IPADDR=10.228.200.173 BROADCAST=10.228.203.255 NETWORK=10.228.200.0 NETMASK=255.255.252.0 ONBOOT=yes</pre>
11.	Bring up the <code>ib0</code> interface.	Type <code>ifup ib0</code> and press Enter .
12.	Perform a test ping.	<p>Type <code>ping <remote IPoIB address></code> and press Enter.</p> <p>For example:</p> <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>
End Task		

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](#).

4.5.2 Using SLES* Distribution Software

You can install the Intel® Omni-Path Software package that is included with the SUSE* Linux* Enterprise Server (SLES*) OS distribution.

Assumptions

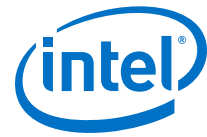
- You are logged in.
- You have your IPV4 list of IP addresses and netmasks for each interface you are going to set up.

Procedures

Perform the following steps to install the default Intel® Omni-Path Software configuration using SLES* OS:



Step	Task/Prompt	Action
Install OPA-Basic Software		
1.	At the command prompt, enter the installation command for opa-basic-tools.	Type zypper install -y opa-basic-tools and press Enter .
2.	At the command prompt, reboot the server.	Type reboot and press Enter .
3.	Check your link using opainfo.	Type opainfo and press Enter . Example output: <pre>hfil_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: 251 0 MB Pkts: 251 0 MB Recv Data: 251 0 MB Pkts: 251 0 MB Link Quality: 5 (Excellent)</pre>
4.	Install the rdma-core rpm.	Type zypper install -y rdma-core and press Enter .
5.	On all compute nodes: install the PSM2 library.	Type zypper install -y libpsm2-2 and press Enter .
Install Intel® Omni-Path Fabric Suite Components on the Management Node		
6.	Install FastFabric.	Type zypper install -y opa-fastfabric and press Enter .
8.	Install Fabric Manager.	Type zypper install -y opa-fm and press Enter .
9.	Start the Fabric Manager.	Type systemctl start opafm and press Enter .
Set up IPoIB IPV4 Configuration		
10.	Manually edit or create the ifcfg-ibX file.	<i>Note:</i> Use the OS distribution-supplied instructions for setting up network interfaces. Type cat /etc/network/ifcfg-ib0 and press Enter . Example output: <pre>BOOTPROTO=static IPADDR=192.168.0.1 BROADCAST=192.168.0.255 NETWORK=192.168.0.0 NETMASK=255.255.255.0 STARTMODE=auto</pre>
11.	Bring up the ib0 interface.	Type ifup ib0 and press Enter .
12.	Perform a test ping.	Type ping <remote IPoIB address> and press Enter . 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>
continued...		



Step	Task/Prompt	Action
(Optional) Install the Fabric Manager GUI		
13.	On one node in the fabric: install the Fabric Manager GUI.	<p><i>Note:</i> Intel recommends not to install the Fabric Manager GUI on the Management Node where the Fabric Manager is being used.</p> <p>Type zypper install -y opa-fmgrui and press Enter.</p>
End Task		

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](#).

4.6 Configure IPoIB IPV6

If required, you may configure IPoIB for IPv6 on the management and host nodes. Note that large clusters (hundreds to thousands of nodes) may experience performance penalties if IPv6 is enabled. See the *Intel® Omni-Path Fabric Performance Tuning User Guide* for more information. Instructions are dependent on the particular OS installed on the node.

Using Red Hat* Enterprise Linux*

To configure IPoIB for IPV6, perform the following:

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

```
NETWORKING_IPV6=yes
```

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

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

IPV6 address should look like the following example:

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

3. Restart the network.

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.
- Configure your IFS software.

Related Links

[Configuring the Software](#) on page 93

4.7 Reinstall Software After a Kernel Update

If a kernel update occurs due to a distro upgrade, remove OPA-IFS if it is already installed, then download and install the latest OPA-IFS package for the new distro. Intel recommends that you use the following command to uninstall OPA-IFS:

```
opaconfig -u
```

If a kernel update comes from installing a patch for the current distro, such as installing a security patch, you may need to reinstall the Intel® Omni-Path Software to ensure the Intel® OPA kernel modules are rebuilt against the new kernel. The following provides instructions about rebuild and reinstall Intel® OPA kernel modules.

NOTE

You only need to reinstall the component OPA Stack.

For OPA-IFS user space only installation, such as Intel® OPA in a container, the OPA kernel module rebuild is not required because no OPA kernel module is installed.

To allow rebuilding of the kernel module, you must first install all of the OS kernel-related packages (RPM filenames starting with `kernel`) for the new kernel. Refer to the *Intel® Omni-Path Fabric Software Release Notes* for the list of RPMs required for each supported OS in this release.

4.7.1 Reinstall Using TUI Menus

You can reinstall the OPA Stack component using the Intel OPA Software TUI menu.

When a kernel version is changed, the TUI sets the action for related components to `Install`. Reinstalling the component OPA Stack will also prompt for the installation of OPA Stack-dependent components as well. You do not need to reinstall them. Only OPA Stack is required.

New Kernel is Active

If the new kernel is active, follow the instructions described in [Install Using the TUI Menus](#) on page 43.

New Kernel is Inactive

If the new kernel is not active, perform the following steps:



1. Launch the INSTALL TUI with the argument `-R <kernel_version>`. For example:

```
./INSTALL -R 3.10.0-1062.12.1.el7.x86_64
```

2. Continue following the instructions described in [Install Using the TUI Menus](#) on page 43.

4.7.2 Reinstall Using CLI Commands

You can reinstall the OPA Stack component using the `./INSTALL` command.

New Kernel is Active

If the new kernel is active, follow the instructions described in [Install Using CLI Commands](#) on page 45, with the following modification:

In step 2, enter the following command to reinstall the component `opa_stack`:

```
./INSTALL -i opa_stack
```

NOTE

For CUDA support, include the CUDA argument `-G`:

```
./INSTALL -G -i opa_stack
```

New Kernel is Inactive

If the new kernel is not active, follow the instructions described in [Install Using CLI Commands](#) on page 45, with the following modification:

In step 2, enter the CLI command with the argument `-R <kernel_version>`. For example:

```
./INSTALL -R 3.10.0-1062.12.1.el7.x86_64 -i opa_stack
```

NOTE

For CUDA support, include the CUDA argument `-G`:

```
./INSTALL -G -R 3.10.0-1062.12.1.el7.x86_64 -i opa_stack
```

4.7.3 Reinstall Using Linux Distribution Software Packages Provided by Intel

You can reinstall the OPA module using the OS distribution included in the IFS package repository and its dependencies.



New Kernel is Active

If the kernel is active, perform the following steps:

1. Run `opacreaterepo` to update the Intel® OPA software repository.
This script will check kernel version and rebuild OPA kernel modules.
2. If an enterprise repository is used, update it with the new updated local repository.
3. On each node, update the module packages:
 - For RHEL systems (using the `yum` command): `kmod-ifs-kernel-updates` and `ifs-kernel-updates-devel`
 - For SLES systems (using the `zypper` command): `ifs-kernel-updates-kmp-default` and `ifs-kernel-updates-devel`

New Kernel is Inactive

NOTE

`opacreaterepo` supports only current active kernel.

If the new kernel is not active, perform the following steps to update Intel® OPA software repository:

1. Change directory to the location of the installation software package:

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

2. Rebuild the OPA kernel module:

```
./INSTALL -R <kernel_version> -i opa_stack
```

NOTE

For CUDA support:

```
./INSTALL -G -R <kernel_version> -i opa_stack
```

3. Navigate to the OPA software repository:

```
cd repos/OPA_PKGS/RPMS
```

NOTE

For CUDA support:

```
cd repos/OPA_PKGS_CUDA/RPMS
```



4. Remove OPA module rpm in the OPA software repository:

```
rm -f *kernel*.rpm
```

5. Link the new OPA module rpm to the OPA software repository:

- For RHEL system:

```
ln -s ../../../../IntelOPA-OFS_DELTA.<DISTRO>-x86_64.<VERSION>/RPMS/<distro>/
kmod-ifs-kernel-updates-<kernel_version>-<build_number>.x86_64.rpm
ln -s ../../../../IntelOPA-OFS_DELTA.<DISTRO>-x86_64.<VERSION>/RPMS/<distro>/
ifs-kernel-updates-devel-<kernel_version>-<build_number>.x86_64.rpm
```

- For SLES system:

```
ln -s ../../../../IntelOPA-OFS_DELTA.<DISTRO>-x86_64.<VERSION>/RPMS/<distro>/
ifs-kernel-updates-kmp-default-<kernel_version>-<build_number>.x86_64.rpm
ln -s ../../../../IntelOPA-OFS_DELTA.<DISTRO>-x86_64.<VERSION>/RPMS/<distro>/
ifs-kernel-updates-devel-<kernel_version>-<build_number>.x86_64.rpm
```

NOTES

For CUDA support:

- For RHEL system:

```
ln -s ../../../../IntelOPA-OFS_DELTA.<DISTRO>-x86_64.<VERSION>/RPMS/<distro>/
CUDA/kmod-ifs-kernel-updates-<kernel_version>-
<build_number>cuda.x86_64.rpm
ln -s ../../../../IntelOPA-OFS_DELTA.<DISTRO>-x86_64.<VERSION>/RPMS/<distro>/
CUDA/ifs-kernel-updates-devel-<kernel_version>-
<build_number>cuda.x86_64.rpm
```

- For SLES system:

```
ln -s ../../../../IntelOPA-OFS_DELTA.<DISTRO>-x86_64.<VERSION>/RPMS/<distro>/
CUDA/ifs-kernel-updates-kmp-default-<kernel_version>-
<build_number>cuda.x86_64.rpm
ln -s ../../../../IntelOPA-OFS_DELTA.<DISTRO>-x86_64.<VERSION>/RPMS/<distro>/
CUDA/ifs-kernel-updates-devel-<kernel_version>-
<build_number>cuda.x86_64.rpm
```

6. Rebuild the repository metadata:

```
createrepo --update ..
```

7. If an enterprise repository is used, update it with the new updated local repository.

8. On each node, update the module packages:

- For RHEL systems (using the yum command): `kmod-ifs-kernel-updates` and `ifs-kernel-updates-devel`
- For SLES systems (using the zypper command): `ifs-kernel-updates-kmp-default` and `ifs-kernel-updates-devel`



5.0 Install and Upgrade Standalone Firmware

This section provides information and procedures to install and upgrade the Intel® Omni-Path standalone firmware.

The standalone firmware consist of the following rpms:

- **hfi1-uefi-<version>.rpm**: Contains HFI1 UEFI Option ROM and the Intel® Omni-Path Boot EFI application
- **hfi1-tmm-<version>.rpm**: Contains the Thermal Management Microchip (TMM) firmware
- **hfi1-firmware-tools-<version>.rpm**: Contains tools required to manage the HFI UEFI and TMM firmware:
 - `hfi1_eprom`: Update/query images on the Intel® Omni-Path HFI Adapter EPROM
 - `opatmmtool`: Manages and updates the firmware on the Thermal Management Microchip (TMM).

5.1 Before You Begin

Before installation or upgrading the standalone firmware, perform the following:

- Refer to the Release Notes for the list of compatible operating systems.
- You have installed the Intel® Omni-Path Software.

5.2 Download the Firmware

Download the firmware using the following procedures.

1. Using a web browser, go to <https://www.intel.com/content/www/us/en/design/products-and-solutions/networking-and-io/fabric-products/omni-path/downloads.html>.
2. Click on Latest Release Library.
3. In the table, click the drop-down arrow for Software and Drivers to locate the files you need.
Or, use the search box above the table to enter your search criteria.
The results are displayed.
4. Select your target firmware and click the **Download** button.
5. Review the Intel Software License Agreement.
6. Click **Accept**.

The zipped firmware is saved to your computer.



5.3 Installing and Upgrading rpms

This section provides information for installing or upgrading firmware rpms.

- To install the rpms, use `rpm -ivh <rpm name>`.
- To upgrade the rpms, use `rpm -Uvh <rpm name>`.

Important Information for Firmware Tools

- When trying to install `hfil-firmware-tools.rpm` on OPA-IFS prior to V10.7, the installation fails due to an existing version of these firmware tools owned by other rpms.

To resolve this issue, you can safely force the installation of the rpm using the following command:

```
rpm -ivh --force hfil-firmware-tools-<version>.rpm
```

- When trying to install a pre-10.7 version of OPA-IFS on systems where the `hfil-firmware-tools.rpm` is already installed, the `hfil-diagtools-sw` and `opa-basic-tools` will fail to install due to existing versions owned by the `hfil-firmware-tools`.

To resolve this issue, you can

- Uninstall the existing `hfil-firmware-tools.rpm` prior to the OPA-IFS installation.
- If desired, you can safely force the installation of the rpm using the following command:

```
rpm -ivh --force hfil-firmware-tools-<version>.rpm
```

Important Information for TMM Tools

- When trying to install `hfil-tmm rpm` on OPA-IFS prior to V10.4, the installation fails due to an existing version of TMM firmware owned by other rpm i.e., `hfil-firmware rpm`.

To resolve this issue, you can safely force the installation of the rpm using the following command:

```
rpm -ivh --force hfil-tmm-<version>.rpm
```

- When trying to install a pre-10.4 version of OPA-IFS on systems where the `hfil-tmm rpm` is already installed, the `hfil-firmware` will fail to install due to existing versions owned by the `hfil-tmm`.

To resolve this issue, you can

- Uninstall the existing `hfil-tmm rpm` prior to the OPA-IFS installation.
- If desired, you can safely force the installation of the rpm using the following command:

```
rpm -ivh --force hfil-tmm-<version>.rpm
```



- When trying to install `hfil-tmm rpm` on OPA-IFS versions between V10.4 and the current release (or vice versa), installation will succeed and the user can see that the TMM firmware is owned by two rpms (`hfil-tmm` and `hfil-firmware`).

This is an expected behavior as TMM firmware in both the rpms are identical.

```
— rpm -qf /lib/firmware/updates/hfil_smbus.fw  
  
hfil-tmm-<version>.noarch  
  
hfil-firmware-<version>.noarch
```

Related Links

[Upgrade the HFI UEFI Firmware](#) on page 60

5.3.1 Upgrade the HFI UEFI Firmware

The HFI UEFI firmware for the Intel® Omni-Path Host Fabric Interface (HFI) discrete adapters is available as a separate rpm. For integrated HFIs, UEFI firmware is integrated with a BIOS image and can only be changed by updating the BIOS.

Pre-Requisites

You must download the standalone firmware rpms below prior to performing the steps.

- `hfil-uefi-<version>.rpm`
- `hfil-firmware-tools-<version>.rpm`

Steps

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

1. Install the HFI UEFI rpm.

```
# rpm -i hfil-uefi-version.rpm
```

2. Install the HFI Firmware Tools rpm.

```
# rpm -i hfil-firmware-tools-version.rpm
```

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

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

4. Update the HFI UEFI firmware. The basic command is:

```
# hfil_eprom [See Additional Options] -u ./HfiPcieGen3Loader_X.X.X.X.rom ./HfiPcieGen3_X.X.X.X.efi
```



Alternatively you can use a wildcard to update the HFI UEFI firmware using the file from specific directory:

```
# cd /usr/share/opa/bios_images/
# hfi1_eprom [See Additional Options] -u ./*
```

Example output of update:

```
Updating driver file with "./HfiPcieGen3_1.7.0.0.0.efi"
Updating loader file with "./HfiPcieGen3Loader_1.7.0.0.0.rom"
Using device: /sys/bus/pci/devices/0000:05:00.0/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 `-u` option.

```
# hfi1_eprom -d all -u HfiPcieGen3Loader_X.X.X.X.X.rom
HfiPcieGen3_X.X.X.X.X.efi
```

- To update each HFI separately (using `-d` only), perform the following steps:
 - List all available HFI discrete cards in the system:

```
# hfi1_eprom -d
0: /sys/bus/pci/devices/0000:05:00.0/resource0
1: /sys/bus/pci/devices/0000:81:00.0/resource0
```

- Update specific card using card index:

```
# hfi1_eprom -d 1 -u HfiPcieGen3Loader_X.X.X.X.X.rom
HfiPcieGen3_X.X.X.X.X.efi
Updating loader file with "HfiPcieGen3Loader_X.X.X.X.X.rom"
Updating driver file with "HfiPcieGen3_X.X.X.X.X.efi"
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
```

- Read UEFI firmware version of all HFIs in the system.

```
# hfi1_eprom -d all -V
```

Example output:

```
Using device: /sys/bus/pci/devices/0000:05:00.0/resource0
loader file version: X.X.X.X.X
driver file version: X.X.X.X.X
config file version: HFI_TYPE1 vX.X.X.X
```

- Reboot the server.

```
# reboot
```



Related Links

[Download the Firmware](#) on page 58

5.4 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*.

5.4.1 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.

5.4.2 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](#) section in the [Setting Up PXE Boot for Supported Operating Systems](#) 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.

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.

Limitations

When UEFI Secure Boot mode is enabled, the following Intel® Omni-Path Linux* utilities are not available:

- hfil_eprom
- hfidiaqs



5.5 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:

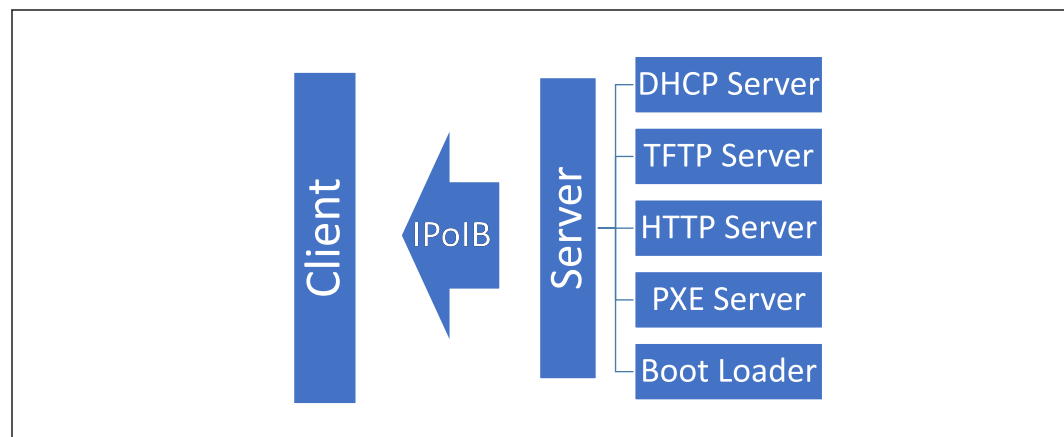
1. Client requests to DHCP Server and receives the IP address of TFTP Server and boot image (the name of bootloader).
2. Client contacts the TFTP Server and receives the boot image.
3. Client executes the boot image and downloads from PXE Server all other needed files like kernel and initial root file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



5.5.1 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.

5.5.1.1 Configure PXE Server with RHEL*

5.5.1.1.1 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).

NOTE

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
```

5.5.1.1.2 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
```

5.5.1.1.3 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/shim.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 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/shim.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
```

5.5.1.1.4 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
```

5.5.1.1.5 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
```

5.5.1.1.6 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/`.

5.5.1.1.7 Configure the Boot Loader

To configure a PXE server to use the **Grub2** boot loader, perform the following steps.



NOTE

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

1. Copy the `shim.efi` file from the shim package and the `grubx64.efi` file from the grub2-efi package in the ISO image file. Enter the following commands as root:

```
# mount -t iso9660 /path_to_image/name_of_image.iso /mount_point -o loop,ro
# cp -pr /mount_point/Packages/shim-version-architecture.rpm /
publicly_available_directory
# cp -pr /mount_point/Packages/grub2-efi-version-architecture.rpm /
publicly_available_directory
# umount /mount_point
```

2. Extract the packages.

```
# rpm2cpio shim-version-architecture.rpm | cpio -dimv
# rpm2cpio grub2-efi-version-architecture.rpm | cpio -dimv
```

3. Copy the EFI boot images from your boot directory.

```
# cp publicly_available_directory/boot/efi/EFI/redhat/shim.efi /var/lib/
tftpboot/uefi
# cp publicly_available_directory/boot/efi/EFI/redhat/grubx64.efi /var/lib/
tftpboot/uefi
```

4. Create the Grub2 configuration file `/var/lib/tftpboot/uefi/grub.cfg`.

The following is an example for RHEL* 7.x using `vmlinuz` kernel file and `initrd.img` as a minimal file system image.

```
set timeout=60
menuentry 'RHEL 7' {
    linuxefi uefi/vmlinuz
    initrdefi uefi/initrd.img
}
```

5.5.1.2 Configure PXE Server with SLES*

5.5.1.2.1 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).



NOTE

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
```

5.5.1.2.2 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
```

5.5.1.2.3 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/shim.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:

```
00:00:00:00:00:02:00:00:01:57:
```

```
[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/shim.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
```

5.5.1.2.4 Install a TFTP Server

For SLES 12 SPX

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
```

For SLES 15 and above

To install a TFTP server, perform the following steps:

1. Install TFTP package using the command:

```
zypper install tftp
```

2. Configure TFTP configuration in `/etc/sysconfig/tftp` if needed.
3. Start TFTP:

```
systemctl start tftp.socket
```

4. Enable TFTP to be run always with OS start:

```
systemctl enable tftp.socket
```

For additional information, refer to *Section 15.2.1, Installing TFTP Server* on <https://documentation.suse.com/sles/15-SP1/html/SLES-all/cha-deployment-prep-pxe.html>

5.5.1.2.5 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/
```

5.5.1.2.6 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/`.



5.5.1.2.7 Configure the Boot Loader

To configure a PXE server to use the **Grub2** boot loader, perform the following steps.

NOTE

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

1. Copy the EFI boot files from the ISO image file. Enter the following commands as root.

```
# mount -t iso9660 /path_to_image/name_of_image.iso /mount_point -o loop,ro
# cd /mount_point/EFI/BOOT
# umount /mount_point
```

2. Copy the EFI boot images from your boot directory.

```
# cp -a bootx64.efi grub.efi MokManager.efi /var/lib/tftpboot/uefi
```

3. Create the Grub2 configuration file: /var/lib/tftpboot/uefi/grub.cfg

The following is an example for SUSE* Linux 12.x using vmlinuz kernel file and initrd.img as a minimal file system image:

```
set timeout=60
menuentry 'SUSE Linux 12' {
    linuxefi uefi/linux
    initrdefi uefi/initrd.img
}
```

5.5.2 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 1. 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 Boot from Ethernet	Customized by UEFI driver	Customized by UEFI driver
UEFI	Legacy BIOS	Yes (BIOS that runs UEFI driver in this mode)	Achieved pre boot	Provided by UEFI driver	Boot local Boot from Ethernet	Default board description	Customized by UEFI driver
UEFI	Legacy BIOS	No (BIOS does not run UEFI driver)	Achieved by host driver post boot	Provided by driver post boot (with limitations)	Boot local Boot from Ethernet	Default board description	Default subsystem values
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
		in this mode – e.g. Intel UEFI BIOS)					
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.

5.5.2.1 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* 7.7 `initrd.img` on a machine that has Linux* RHEL* 7.7 running on it.

5.5.2.2 Install RHEL* 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* 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/hfi1_dc8051_d.fw
/lib/firmware/updates/hfi1_dc8051.fw
/lib/firmware/updates/hfi1_fabric_d.fw
/lib/firmware/updates/hfi1_fabric.fw
/lib/firmware/updates/hfi1_pcie_d.fw
/lib/firmware/updates/hfi1_pcie.fw
/lib/firmware/updates/hfi1_sbus_d.fw
/lib/firmware/updates/hfi1_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` where `xx` is the version number.

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

default=rhelxx
image=vmlinuz
    label=rhelxx
    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/RHELxx"
```

5.5.2.3 Install SLES* over Intel® OP HFI

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

1. Log in as root.
2. Mount the SLES* `.iso` file (where `xx-SPx` is the SLES* version number).

```
sudo mkdir /mnt/iso
sudo mount -o loop SLE-xx-SPx-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 grub.cfg to include the line `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.

```
set timeout=60
menuentry 'SUSE Linux xx' {
    linuxefi uEFI/linux insmod=ib_ipoib netdevice=ib0 netsetup=dhcp
    install=http://192.168.100.1/netinstall/SLES_XXX
    initrdefi uEFI/initrd-new.img
}
```



6.0 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.

6.1 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.

6.1.1 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) 1.8
- Ethernet card/local network access
- Minimum disk space requirement is 100 MB. This amount includes the disk space needed for the application binaries and the application cache for topology information.
- 2 GB or greater of RAM
- 1280 x 800 resolution (65 K color depth)

6.1.2 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) 1.8
- libgnome
- X Window System
- Ethernet card/local network access
- Minimum disk space requirement is 100 MB. This amount includes the disk space needed for the application binaries and the application cache for topology information.
- 2 GB or greater of RAM
- 1280 x 800 resolution (65 K color depth)

6.1.3 Download the Fabric Manager GUI Software

1. Using a web browser, go to <https://www.intel.com/content/www/us/en/design/products-and-solutions/networking-and-io/fabric-products/omni-path/downloads.html>.
2. Click on Latest Release Library.
3. In the table, click the drop-down arrow for Software and Drivers to locate the files you need for the version you want to install.

Or, use the search box above the table to enter your search criteria.

NOTE

You can download either the installation executable file or the source files to build your Intel® OPA Fabric Manager GUI jar file.

This document only addresses installing through the executable file.

The results are displayed.

4. Select your target software and click the **Download** button.
5. Review the Intel Software License Agreement.
6. Click **Accept**.

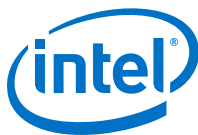
The zipped firmware is saved to your computer.

6.2 Install the Fabric Manager GUI Software on Windows*

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

Assumption

- You have downloaded the installation executable for Windows* to your desktop.
- You have installed JRE version 1.8.



Pre-requisites

The Java Runtime Environment (JRE) version 1.8 must be installed for Fabric Manager GUI to execute. If JRE1.8 is not installed, the Fabric Manager GUI installer will do it for you. If this happens, you will need to update the environment variable JAVA_HOME to the path where the JRE is installed; for example, C:\Program Files\Java\jre1.8.0_181. This section is intended to guide you through the configuration steps to ensure the software functions properly.

IMPORTANT

Use of any JRE version other than 1.8 may result in unexpected behavior.

Step	Task/Prompt	Action
1.	Open the Windows* Control Panel and set up the environment.	For versions pre-Win10: 1. Click the Windows Start icon . 2. Select Control Panel . 3. Click the System icon . 4. Select "Advanced system settings". For Win10: 1. Click the Windows Search icon . 2. Type <code>Advanced system settings</code> and press Enter .
2.	Access the System Environment Variables.	In the System Properties dialog, click Environment Variables .
3.	Under System variables, create or edit the JAVA_HOME variable.	1. Click New to create a new variable named "JAVA_HOME", or select "JAVA_HOME" and click Edit to modify. 2. Set the variable value to the path where the JRE can be found, for example, C:\Program Files\Java\jre1.8.0_181. 3. Click OK .
4.	Under System variables, edit the Path variable.	1. Select "Path" and click Edit . 2. Click New . 3. Type <code>%JAVA_HOME%\bin</code> and click Enter . 4. Move the entry to the top of the list, by selecting it and clicking Move Up until it is at the top. 5. Click OK .
5.	To accept the changes to the Environment Variables.	Click OK .
6.	To accept changes to the System Properties.	Click OK .
	End Task	

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 IntelOPA-FMGUI.windows-x.x.x.x.x.exe file on the desktop
continued...		



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

Next Steps

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



6.3 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.

Pre-requisites

The Java Runtime Environment (JRE) version 1.8 must be installed for Fabric Manager GUI to execute. If JRE1.8 is not installed, the Fabric Manager GUI installer will do it for you. If this happens, you will need to update the environment variable `JAVA_HOME` to the path where the JRE is installed; for example, `/usr/lib/jfm/java-1.8<build-id>/jre`. This section is intended to guide you through the configuration steps to ensure the software functions properly.

IMPORTANT

Use of any JRE version other than 1.8 may result in unexpected behavior.

NOTE

These instructions assume a Bash shell environment is being used

Step	Task/Prompt	Action
1.	Set up the environment using one of two methods: Locally - Edit the <code>.bashrc</code> file in your account. Globally - Create a shell script in <code>/etc/profile.d</code>	During runtime the 'fmgui' script checks the environment variable <code>JAVA_HOME</code> . If it is not set, it checks the system path for 'java'. This is the order of precedence that can be used to permit two versions of the JRE to run simultaneously on a given system. <i>Note:</i> If it is necessary for the JRE on the system path to be greater than 1.8, then <code>JAVA_HOME</code> MUST be set to point to JRE1.8.
2.	Update the script with the following entries:	<ul style="list-style-type: none">If there is no JRE on the system path, the Fabric Manager GUI installer will install it for you and it will be available on the system path. No action is required.If the version of JRE on the system path is greater than 1.8, then enter: <code>export JAVA_HOME=<path-to-jre-8></code> Example: <code>export JAVA_HOME=/usr/lib/jvm/java-1.8.0-openjdk-1.8.0.161-2.b14.el7.x86_64/jre</code>If the version of JRE on the system path is 1.8, then no action is required
3.	Save the file and exit.	
	End Task	

Procedure

To install the software, perform the following:



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

Next Steps

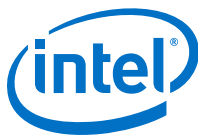
- To prepare the Fabric Manager for Fabric Manager GUI, refer to [Prepare Fabric Manager for Fabric Manager GUI](#).

6.4 Prepare Fabric Manager for Fabric Manager GUI

Fabric Manager GUI requires a running fabric with FE available. This task shows 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.	NOTE: The following steps contain the necessary edits to <code>opa_fm.xml</code> .
2.	Ensure that the FE is enabled.	In the <code><FE></code> section, "Start" is 1. <pre><Fe> <Start>1</Start><!-- default FE startup for all instances --></pre>
3.	Ensure that FE has proper security setup.	
	<ul style="list-style-type: none">• If you want to run FE without SSL, ensure that the <code>SslSecurityEnable</code> is 0. Also, when you run Fabric Manager GUI, set up FE host as unsecured (clear the "Secure" checkbox).	<pre><!-- OpenSSL FE network security parameters --> <SslSecurityEnable>0</SslSecurityEnable></pre>
	<ul style="list-style-type: none">• If you want to run FE with SSL, ensure that you set <code>SslSecurityEnable</code> to 1. To set up keys and <code>opa_fm.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>:<ul style="list-style-type: none">— "Out-of-Band Security"— "Core-Level Public Key Infrastructure (PKI) Best Practices Guidelines"— "Advanced-Level Public Key Infrastructure Best Practices Guidelines"— "SSL Key Creation for Fabric Manager GUI"And, when you run Fabric Manager GUI, set up FE host as "Secure" with the key files (select the "Secure" checkbox).	<pre><!-- OpenSSL FE network security parameters --> <SslSecurityEnable>1</SslSecurityEnable></pre>
4.	Ensure that PM is enabled.	In the <code><Pm></code> section, "Start" is 1. <pre><Pm> <!-- ESM does not support Start via XML configuration. Use CLI commands --> <Start>1</Start> <!-- default PM startup for all instances --></pre>
5.	If you made any changes...	Save and close the file.
6.	Restart the Fabric Manager so that your changes take effect.	Type <code>systemctl restart opa_fm</code> .
	End Task	

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*](#).



6.5 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:

1. From the **Start** menu, select **All Programs**.
2. 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**.
-

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**.

6.6 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.

6.7 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.

6.7.1 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.

6.7.1.1 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.

6.7.2 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.

6.7.2.1 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
```

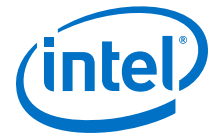
The Fabric Manager GUI cache is cleared.

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.



Part 2: Configuring the Software



7.0 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 [Software Installation Checklists](#) on page 163 to track your configuration and verification of the software.

7.1 Pre-Configuration Requirements

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

7.1.1 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.

7.1.2 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.

7.2 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 (or Q)
```

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 or Q)
```

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` or `Q` 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` or `Q` 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
```



8.0 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.

8.1 Before You Begin

Before starting the chassis configuration, perform the following:

- Be sure you have completed all [Chassis Configuration Prerequisites](#) on page 94.
- 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 138 for additional instructions.

8.2 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
Configuring the Chassis		
1.	Access the FastFabric OPA Chassis Setup/Admin menu .	
	a) If you are not already logged into Intel FastFabric OPA Tools , at the command prompt...	Type <code>opafastfabric</code> and press Enter .
	b) Access the FastFabric OPA Chassis Setup/Admin menu .	Press 1 .
	c) Select menu items.	Select all operations by entering each menu item 0 – e. NOTE: 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 P . NOTE: Each selected item is preformed in the order of the menu list.
2.	Edit the Configuration and Select/Edit Chassis File (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">• <code>FF_CHASSIS_LOGIN_METHOD</code>• <code>FF_CHASSIS_ADMIN_PASSWORD</code>.• Select the location for the result files from FastFabric with the <code>FF_RESULT_DIR</code> parameter.
continued...		



Step	Task/Prompt	Action
		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>chassis</code> 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 n and Press Enter .
	e) Continue to next step.	Press any key to continue.
3.	Verify Chassis via Ethernet Ping (menu item 1)	
	• If all chassis were found...	Press any key to continue to Update the Chassis Firmware .
	• If some chassis were not found...	Press ESC to exit the menu and review the following list for those chassis which were not found: <ul style="list-style-type: none"> • Is chassis powered on and booted? • Is chassis connected to management network? • Are chassis IP address and network settings consistent with DNS or /etc/hosts ? • Is Management node connected to the management network? • Are Management node IP address and network settings correct? • Is management network itself up (including switches, routers, and others)? • Is correct set of chassis listed in the chassis file? You may need to repeat the previous step to review and edit the file.
4.	(Optional) Update Chassis Firmware (menu item 2)	NOTE: Before continuing, refer to the <i>Intel® Omni-Path Fabric Switches Release Notes</i> for any prerequisites.
	a) For Directories, all <code>.spkg</code> 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 Enter .
	b) Would you like to run the firmware now? [n]:	Type y and press Enter . NOTE: FastFabric ensures that all chassis are running the firmware level provided, and installs and/or reboots each chassis as needed. If any chassis fails to be updated, use the View opachassisadmin Result Files option to review the result files from the update. Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more details.
	c) Continue to next step.	Press any key to continue. <i>Note:</i> This operation may take a few minutes to complete.
5.	Set Up Chassis Basic Configuration (menu item 3)	
	a) For each prompt...	Accept the defaults.
	b) Continue to next step.	Press any key to continue.
6.	Set Up Password-less ssh/scp (menu item 4)	
continued...		



Step	Task/Prompt	Action
	a) Would you like to override the default Chassis password? [n]:	Press Enter to accept the default Chassis password.
	b) Continue to next step.	Press any key to continue.
7.	Reboot the Chassis (menu item 5)	
	a) Performing Chassis Admin: Reboot Chassis Would you like to be prompted for chassis' password? [n]:	Press Enter to accept the default.
	b) Continue to next step.	Press any key to continue. <i>Note:</i> This operation may take a few minutes to complete.
8.	Get Basic Chassis Configuration (menu item 6)	
	a) Continue to next step.	Press any key to continue.
9.	Configure Chassis Fabric Manager (menu item 7)	NOTE: 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 generate and press Enter .
	b) For each prompt...	Accept the defaults.
	c) Would you like to restart the FM? [n]:	Type y and press Enter .
	d) Would you like to run the FM on slave MMs? [n]:	Choose: <ul style="list-style-type: none"> Type y. 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. Type n. 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 be started only on the master management module in the applicable chassis.
	e) Would you like to do the operation in parallel? [y]:	Press Enter .
	f) Would you like to enable FM start at boot? [n]:	Type y .
	g) Would you like to enable FM start on slave MMs at boot? [n]:	Choose: <ul style="list-style-type: none"> Type y. 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. Type n. 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.
continued...		



Step	Task/Prompt	Action
	h) Would you like to be prompted for chassis' password? [n]:	Press Enter .
	i) Are you sure you want to proceed? [n]:	Type y .
	j) Continue to next step.	Press any key to continue.
10.	Updating the Chassis FM Security Files (menu item 8)	
	a) Enter Files/Directories to use (or none):	NOTE: You must enter a file or the task ends.
	b) For subsequent prompts...	Provide the required information and press Enter .
	c) Continue to next step.	Press any key to continue.
11.	Get the Chassis FM Security File (menu item 9)	
	a) Continue to next step.	Press any key to continue.
Verifying the Chassis Configuration		
12.	Check OPA Fabric Status (menu item a)	
	a) For each prompt...	Provide the required information and press Enter .
	b) Continue to next step.	Press any key to continue.
13.	Control Chassis Fabric Manager (menu item b)	
	a) For each prompt...	Provide the required information and press Enter .
	b) Are you sure you want to proceed? [n]:	Select y .
	c) Continue to next step.	Press any key to continue.
14.	Generate All Chassis Problem Report Information (menu item c)	
	a) Would you like to be prompted for chassis' password? [n]:	Press Enter .
	b) Continue to next step.	Press any key to continue.
15.	Run a Command on All Chassis (menu item d)	
	a) If there are any other operations that need to be performed on all chassis...	Use the Run a Command on All Chassis option. Each time this is executed, a single chassis CLI command may be specified to be executed against all selected chassis. When using these commands, additional setup or verification of the chassis may be performed. Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> .
	b) Continue to next step.	Press any key to continue.
16.	View the opachassisadmin Result Files (menu item e)	
	a) About to: vi /root/punchlist.csv /root/test.res /root/test.log	Press any key to review files or press ESC to abort.
	b) Continue to next step.	Press any key to continue.
	End Task	

Next Steps

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



8.3 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.

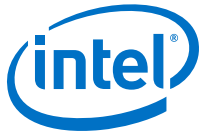
Procedures

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"> • <code>FF_CHASSIS_LOGIN_METHOD</code> • <code>FF_CHASSIS_ADMIN_PASSWORD</code>. • Select the location for the result files from FastFabric with the <code>FF_RESULT_DIR</code> parameter. 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> .
	• If all chassis were found...	Continue to the next step.
	• If some chassis were not found...	Review the following list for those chassis that were not found: <ul style="list-style-type: none"> • Is chassis powered on and booted? • Is chassis connected to management network? • Are chassis IP address and network settings consistent with DNS or <code>/etc/hosts</code> ? • Is Management node connected to the management network? • Are Management node IP address and network settings correct? • Is management network itself up (including switches, routers, and others)? • Is correct set of chassis listed in the chassis file? You may need to repeat the previous step to review and edit the file.
3.	Update the chassis firmware using either the parallel or serial method.	NOTE: 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> .
continued...		



Step	Task/Prompt	Action
	NOTES: During the firmware update, the following files are produced: <ul style="list-style-type: none"> • <code>test.res</code> – Appended with summary results of run • <code>test.log</code> – Appended with detailed results of run • <code>save_tmp/</code> – Contains a directory per failed operation with detailed logs • <code>test_tmp*/</code> – Intermediate result files while operation is running If any chassis fails to be updated, refer to the View opachassisadmin result files section to review the result files from the update. Refer to the <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 <code>opachassisadmin -S -F chassisfile configure</code> .
	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"> • Syslog server • NTP server • Timezone and DST information • Chassis link width • Configure Node Desc to match Ethernet* chassis name • Configure the Link CRC Mode
5.	Set up secure password-less SSH, such that the Management Node can securely log into all the chassis as <code>admin</code> through the management network, without requiring a password.	Type <code>opasetupssh -p -S -C -F chassisfile</code> .
6.	Reboot all the selected chassis and ensure they reboot fully, as verified through ping over the management network.	Type <code>opachassisadmin -S -F chassisfile reboot</code> .
7.	Retrieve basic information from chassis such as syslog, NTP configuration, time zone, node description, and other information.	Type <code>opachassisadmin -F chassisfile getconfig</code> .
8.	Configure the Chassis Fabric Manager for any Intel® Omni-Path Chassis 100 Series chassis using either parallel or serial method.	NOTE: 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 <code>opachassisadmin -S -F chassisfile -P package -a run fmconfig</code> .
	b) For the serial method...	Type <code>FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -P package -a run fmconfig</code> .
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.	Note: The FM security files are the private key, public key, and certificate files required by the FM, to support secure socket connection via OpenSSL. Refer to the <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 <code>opachassisadmin -S -F chassisfile -s securityFiles -a push fmsecurityfiles</code> .
	b) For the serial method...	Type <code>FF_MAX_PARALLEL=0 opachassisadmin -S -F chassisfile -s securityFiles -a push fmsecurityfiles</code> .
continued...		



Step	Task/Prompt	Action
10.	Retrieve the Chassis Fabric Manager Security Files from the chassis.	Type <code>opachassisadmin -F chassisfile fmgetsecurityfiles</code> .
Verifying the Chassis Configuration		
11.	Check the OPA fabric status.	Type <code>opalinkanalysis -U -x snapshot_suffix reports verifyall > results_file 2>&1</code> . NOTE: This step performs the following operations: <ul style="list-style-type: none"> • Perform a fabric error analysis. • Clear the error counters after generating a report. • Perform a fabric link speed error analysis. • Check for links configured to run slower than supported. • Check for links connected with mismatched speed potential. • Enter a filename for the results or save the results to the default location which is: <code>/root/ffres/linkanalysis.res</code>
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.	NOTE: There is a disruption as FMs are restarted. Using the serial method may reduce the disruption.
	• For parallel method...	Type <code>opachassisadmin -S -F chassisfile -a restartall fmcontrol</code> .
	• For serial method...	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.	NOTE: There is a disruption as FMs are restarted. Using the serial method may reduce the disruption.
	• For parallel method...	Type <code>opachassisadmin -S -F chassisfile -a restart fmcontrol</code> .
	• For serial method...	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.	
	• For parallel method...	Type <code>opachassisadmin -S -F chassisfile -a stop fmcontrol</code> .
	• For serial method...	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.	
	• For parallel method...	Type <code>opachassisadmin -S -F chassisfile -a runall fmcontrol</code> .
	• For serial method...	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.	
continued...		



Step	Task/Prompt	Action
	<ul style="list-style-type: none"> For parallel method... 	Type <code>opachassisadmin -S -F chassisfile -a run fmcontrol</code> .
	<ul style="list-style-type: none"> For serial method... 	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> .
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> . NOTE: 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"> <code>punchlist.csv</code> <code>test.res</code> <code>test.log</code>
	End Task	

Next Steps

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



9.0 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.

9.1 Before You Begin

Before starting the switch configuration, perform the following:

- Be sure you have completed all [Switch Configuration Prerequisites](#) on page 95.
- 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 138 for additional instructions.



9.2 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
Configuring the Externally-Managed Switches		
1.	Access the FastFabric OPA Switch Setup/Admin menu .	
	a) If you are not already logged into Intel FastFabric OPA Tools , at the command prompt...	Type <code>opafastfabric</code> and press Enter .
	b) Access the FastFabric OPA Switch Setup/Admin menu .	Press 2 .
	c) Select menu items.	Select all operations by entering each menu item 0 – a. NOTE: 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 P . NOTE: Each selected item is preformed in the order of the menu list.
2.	Edit the Configuration and Select/Edit Switch File (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 n and Press Enter .
	e) Continue to next step.	Press any key to continue.
3.	Generate or Update Switch File (menu item 1)	
	a) For each prompt...	Provide the required information and press Enter .
	b) Continue to next step.	Press any key to continue.
continued...		



Step	Task/Prompt	Action
4.	Test for Switch Presence (menu item 2)	
	• If all switches were found...	Press any key to continue to Verify Switch Firmware .
	• If some switches were not found...	Press ESC to exit the menu and review the following list for those switches that were not found: <ul style="list-style-type: none"> • Is switch powered on and booted? • Is switch connected to Intel® Omni-Path Fabric? • Is Subnet Manager running? • Is Management Node's Port active? • Is Management Node connected to the correct Intel® Omni-Path Fabric? • Is FM Switch LED activated on the switch port to which the Fabric Management node is connected? For more information, refer to the "FM Switch" section in the <i>Intel® Omni-Path Fabric Switches Hardware Installation Guide</i>. • 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.
5.	Verify Switch Firmware (menu item 3)	
	a) Continue to next step.	Press any key to continue.
6.	Update Switch Firmware (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) 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 y .
	c) For subsequent prompts...	Press Enter to accept the defaults.
	d) Continue to next step.	Press any key to continue.
7.	Set Up Switch Basic Configuration (menu item 5)	
	a) For each prompt...	Press Enter to accept the defaults.
	b) Continue to next step.	Press any key to continue.
8.	Reboot Switch (menu item 6)	
	a) Continue to next step.	Press any key to continue.
9.	Report Switch Firmware & Hardware Info (menu item 7)	
	a) Continue to next step.	Press any key to continue.
10.	Get Basic Switch configuration (menu item 8)	
	a) Continue to next step.	Press any key to continue.
11.	Report Switch VPD Information (menu item 9)	
	a) Continue to next step.	Press any key to continue.
Verifying the Switch Configuration		
12.	View opaswitchadmin Result Files (menu item a)	
continued...		



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 ESC to abort.
	b) Would you like to remove test.res test.log test_tmp* and save_tmp in /root ? [n]:	Press Enter to save or type y to remove the files.
	End Task	

Next Steps

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

9.3 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. NOTE: You can also update switch names in the switches file by comparing the actual fabric to topology xml data.	Type opagenswitches -s -o switches .
3.	Test to see if each externally-managed switch is present.	Type opaswitchadmin -L switchfile ping .
	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"> Is switch powered on and booted? Is switch connected to Intel® Omni-Path Fabric? Is Subnet Manager running? Is Management Node's Port active?

continued...



Step	Task/Prompt	Action
		<ul style="list-style-type: none">Is Management Node connected to the correct Intel® Omni-Path Fabric?Is FM Switch LED activated on the switch port to which the Fabric Management node is connected? For more information, refer to the "FM Switch" section in the <i>Intel® Omni-Path Fabric Switches Hardware Installation Guide</i>.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.
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.	NOTE: 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">For the parallel method...	Type <code>opaswitchadmin -O -L switchfile -P packages -a run upgrade</code> .
	<ul style="list-style-type: none">For the serial method...	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> . NOTE: 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> .
Verifying the Switch Configuration		
11.	View the <code>opaswitchadmin</code> results files.	Type <code>editor result_dir/result_file</code> . NOTE: 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"><code>punchlist.csv</code><code>test.res</code><code>test.log</code>
	End Task	

Next Steps

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



10.0 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.

10.1 Before You Begin

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

- On RHEL 7.x systems, install parallel gzip (pigz) and make it the default compression tool prior to upgrading the software to avoid false installation failures being reported.
- 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 138 for additional instructions.

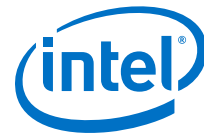
- For the following Host Setup menu item, prepare to answer questions or provide input as follows:

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

Choose MPI Directory Selection:

```
Host Setup: Build Test Apps and Copy to Hosts
MPI Directory Selection

Please Select MPI Directory:
0) /usr/mpi/gcc/mvapich2-2.3b-hfi
1) /usr/mpi/gcc/openmpi-3.1.4-cuda-hfi
2) /usr/mpi/gcc/openmpi-3.1.4-hfi
```



```
3) Enter Other Directory
X) Return to Previous Menu (or ESC or Q)
```

NOTE

Refer to the *Intel® Omni-Path Fabric Software Release Notes* for the latest supported MPI Library versions.

- 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.

10.2 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.

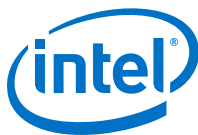
Assumption

- You are logged into the management node.
- On RHEL 7.x systems, you have installed parallel gzip (`pigz`) and made it the default compression tool.

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



Step	Task/Prompt	Action
2.	Edit Configuration and Select/Edit Host File (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"> • <code>FF_IPOIB_SUFFIX</code> • <code>FF_IPOIB_NETMASK</code> • <code>FF_IPOIB_CONFIG</code> • <code>FF_PRODUCT</code> • <code>FF_PACKAGES</code> • <code>FF_INSTALL_OPTIONS</code> • <code>FF_UPGRADE_OPTIONS</code> If you made any changes, save and close the file. Press any key to continue.
	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 n and Press Enter .
	e) Continue to next step.	Press any key to continue.
3.	Verify Hosts Pingable (menu item 1)	
	• If all hosts were found...	Press any key to continue to Set Up Password-Less SSH/SCP .
	• If some hosts were not found...	Press ESC to exit the menu and review the following list for those hosts that were not found: <ul style="list-style-type: none"> • Host powered on and booted? • Host connected to management network? • Host management network IP address and network settings consistent with DNS or <code>/etc/hosts</code>? • Management node connected to the management network? • Management node IP address and network settings correct? • Management network itself up (including switches, routers, and others)? • Correct set of hosts listed in the <code>hosts</code> file? You may need to repeat the previous step to review and edit the file.
4.	Set Up Password-Less SSH/SCP (menu item 2)	
	a) Password for <code>root</code> on all hosts:	Type the password for <code>root</code> on all hosts and press Enter .
	b) Continue to next step.	Press any key to continue.
5.	Copy /etc/hosts to All Hosts (menu item 3)	
	a) Continue to next step.	Press any key to continue.
6.	Show <code>uname -a</code> for All Hosts (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.
continued...		



Step	Task/Prompt	Action
7.	Install/Upgrade OPA Software (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 Upgrade the Software on the Remaining Servers on page 159.
	a) Do you want to use ./IntelOPA-[Basic IFS].DISTRO.VERSION.tgz? [y]:	Press Enter to accept the default.
	b) Would you like to do a fresh [i]ninstall, an [u]pgrade or [s]kip this step? [u]:	Type i and press Enter .
	c) Are you sure you want to proceed? [n]:	Type y and press Enter .
	d) Complete the installation.	Press any key to continue. <i>Note:</i> This step may take several minutes to complete.
	• If all hosts install...	Press any key to continue.
	• If any hosts fails to install...	Use the View opahostadmin Results Files menu item to review the result files from the update.
8.	Configure IPoIB IP Address (menu item 6)	
	a) Continue to next step.	Press any key to continue.
9.	Build Test Apps and Copy to Hosts (menu item 7)	
	a) For each prompt...	Provide your selections and press Enter .
	b) Continue to next step.	Press any key to continue.
10.	Reboot Hosts (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 .
Optional Tasks Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more information.		
11.	Refresh SSH Known Hosts (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.	Rebuild MPI Library and Tools (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.	Run a Command on All Hosts (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 Copy a File to All Hosts menu item to copy the appropriate files to all hosts to enable and configure NTP.
14.	Copy a File to All Hosts (menu item c)	A file on the local host may be specified to be copied to all selected hosts.
Verifying the Host Configuration		
15.	View opahostadmin Result Files (menu item d)	
<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 Enter to save or type y to remove the files.
	End Task	

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.](#)

10.3 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

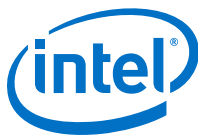
The following steps provide simplified instructions for first-time verification of the hosts. For additional details, refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.

Step	Task/Prompt	Action
1.	Access the FastFabric OPA Host Verification/Admin Menu .	
	a) If you are not already logged into Intel FastFabric OPA Tools , at the command prompt...	Type <code>opafastfabric</code> and press Enter .
	b) Access the OPA Host Verification/Admin Menu .	Press 4 .
	c) Select menu items.	Select items 0 – 8 .
	d) Start the operations.	Press P . NOTE: Each selected item is preformed in the order of the menu list.
2.	Edit Configuration and Select/Edit Host File (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">FF_TOPOLOGY_FILEFF_IPOIB_SUFFIXFF_DEVIATION_ARGS

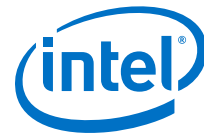
continued...



Step	Task/Prompt	Action
		<ul style="list-style-type: none"> ff_host_basename_to_ipoib ff_host_basename Review the following parameters which are used for overall fabric health checks: <ul style="list-style-type: none"> FF_ANALYSIS_DIR FF_ALL_ANALYSIS FF_FABRIC_HEALTH FF_CHASSIS_CMDS FF_CHASSIS_HEALTH FF_ESM_CMDS FF_ALL_ANALYSIS should be updated to reflect the type of SM (esm or hostsm). If you made any changes, save and close the file. Press any key to continue.
	b) Edit the ports configuration file.	Review the file. If you made any changes, save and close the file. Press any key to continue.
	b) Create or edit the allhosts configuration file.	Create the file with the Management Node's hosts name (the TCP/IP management network name, for example mgmthost) and include the hosts file. <i>Note:</i> If you have a cluster with mixed servers or HFI configurations, create a /etc/opa/*hosts 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 n and Press Enter .
	e) Continue to next step.	Press any key to continue.
3.	Summary of Fabric Components (menu item 1)	
	a) After the operation completes...	Review the results against the expected configuration of the cluster. NOTE: 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.	Verify Hosts Pingable, SSHable and Active (menu item 2)	
	a) For each prompt...	Provide the required information and press Enter .
	b) After completion of the tests, you are prompted: Would you like to now use /etc/opa/good as Host File? [y]:	Press Enter to use the file or n to discard the file.
5.	Perform Single Host Verification (menu item 3)	
	a) Would you like to edit /root/hostverify.sh 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 /root/hostverify.sh file If you made any changes, save and close the file.
continued...		



Step	Task/Prompt	Action
	b) For each prompt...	Provide the required information and press Enter .
	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 Perform Single Host Verification 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 <i>*host</i> file
	f) Continue to next step.	Press any key to continue.
6.	Verify OPA Fabric Status and Topology (menu item 4)	
	a) For each prompt...	Provide the required information and press Enter .
	a) Continue to next step.	Press any key to continue.
7.	Verify Hosts See Each Other (menu item 5)	
	a) Continue to next step.	Press any key to continue.
8.	Verify Hosts Ping via IPoIB (menu item 6)	
	a) Continue to next step.	Press any key to continue.
9.	Refresh SSH Known Hosts (menu item 7)	
	a) Continue to next step.	Press any key to continue.
10.	Check MPI Performance (menu item 8)	NOTE: This test identifies nodes whose performance is not consistent with others in the fabric. It is not intended as a benchmark of fabric latency and bandwidth. This test intentionally uses techniques to reduce test runtime.
	a) For each prompt...	Press Enter to select the defaults.
	• If all hosts pass...	Continue to the next step.
	• If any hosts fail...	<ul style="list-style-type: none">Carefully examine the failing hosts to verify the HFI models, PCIe slot used, BIOS settings, and any motherboard or BIOS settings related to devices on PCIe buses or slot speeds.Also verify that the HFI and any riser cards are properly seated. 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.
Optional Tasks Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more information.		
11.	Check Overall Fabric Health (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 Configure and Initialize Health Check Tools Using FastFabric CLI Commands on page 132 for more information.
continued...		



Step	Task/Prompt	Action
12.	Start or Stop Bit Error Rate Cable Test (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.	Generate All Hosts Problem Report Info (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.	Run a Command on All Hosts (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.
Review the Host Verification		
15.	View opahostadmin Result Files (menu item d)	A file on the local host may be specified to be copied to all selected hosts.
	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 Enter to save or type y to remove the files.
	End Task	

Next Steps

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

10.4 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.

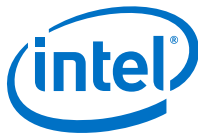
Assumption

- You are logged into the management node.

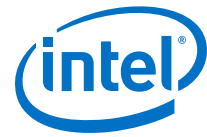
Procedures

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"> FF_IPOIB_SUFFIX FF_IPOIB_NETMASK FF_IPOIB_CONFIG FF_PRODUCT FF_PACKAGES FF_INSTALL_OPTIONS FF_UPGRADE_OPTIONS If you made any changes, save and close the file.
continued...		



Step	Task/Prompt	Action
	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"> If all hosts were found... 	Continue to the next step.
	<ul style="list-style-type: none"> If some hosts were not found... 	Review the following list for those hosts that were not found: <ul style="list-style-type: none"> Host powered on and booted? Host connected to management network? Host management network IP address and network settings consistent with DNS or <code>/etc/hosts</code>? Management node connected to the management network? Management node IP address and network settings correct? Management network itself up (including switches, routers, and others)? Correct set of hosts listed in the hosts file? You may need to repeat the previous step to review and edit the file.
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. NOTE: 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. Note: An initial installation uninstalls any existing OFA Delta or IFS software. Initial installs must be performed when installing on a clean system or on a system that has stock OFA installed. To upgrade the fabric, refer to Upgrade the Software on the Remaining Servers on page 159.	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.
	<ul style="list-style-type: none"> If all hosts install... 	Continue to the next step.
	<ul style="list-style-type: none"> If any hosts fails to install... 	Use the View opahostadmin 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> . NOTE: The file is created with a statically assigned IPv4 address. The IPoIB IP address for each host is determined by the resolver (Linux* <code>host</code> command). If not found through the resolver, <code>/etc/hosts</code> on the given host is checked.
continued...		



Step	Task/Prompt	Action
9.	Build the MPI 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 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</code> <code>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.
Verifying the Host Configuration		
15.	View <code>opahostadmin</code> result files.	Type <code>editor result_dir/result_file.</code> NOTE: 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">• <code>test.res</code>• <code>test.log</code>
	End Task	

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](#).

10.5 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.

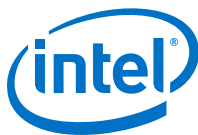
Assumption

- You are logged into the management node.

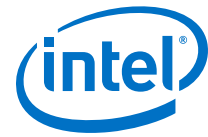
Procedures

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:
<i>continued...</i>		



Step	Task/Prompt	Action
		<ul style="list-style-type: none"> FF_TOPOLOGY_FILE FF_IPOIB_SUFFIX FF_DEVIATION_ARGS ff_host_basename_to_ipoib ff_host_basename <p>Review the following parameters which are used for overall fabric health checks:</p> <ul style="list-style-type: none"> FF_ANALYSIS_DIR FF_ALL_ANALYSIS FF_FABRIC_HEALTH FF_CHASSIS_CMDS FF_CHASSIS_HEALTH FF_ESM_CMDS <p>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.</p>
	b) Edit the <code>/etc/opa/ports</code> configuration file.	<p>Review the file. If you made any changes, save and close the file. Press any key to continue.</p>
	b) Create or edit the <code>/etc/opa/allhosts</code> configuration file.	<p>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.</p> <p><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</p> <p>If you made any changes, save and close the file.</p>
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 opafabricinfo .
	a) After the operation completes...	<p>Review the results against the expected configuration of the cluster.</p> <p>NOTE: If components are missing, or degraded or omitted links are found, they should be corrected. Subsequent steps aid in locating any such links.</p>
3.	(Optional) Verify each host is pingable.	Type opapingall -p -f hostfile .
	<ul style="list-style-type: none"> If all hosts were found... 	Continue to the next step.
	<ul style="list-style-type: none"> If some hosts were not found... 	<p>Review the following list for those hosts that were not found:</p> <ul style="list-style-type: none"> Host powered on and booted? Host connected to management network? Host management network IP address and network settings consistent with DNS or <code>/etc/hosts</code>? Management node connected to the management network? Management node IP address and network settings correct? Management network itself up (including switches, routers, and others)? Correct set of hosts listed in the <code>hosts</code> file? You may need to repeat the previous step to review and edit the file.
continued...		



Step	Task/Prompt	Action
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</code> order with all duplicates removed in the <code>OPA_CONFIG_DIR/</code> directory: <ul style="list-style-type: none"> • good • alive • running • active • bad • quarantined
5.	Perform a single host test on all hosts.	Type <code>opaverifyhosts -k -c -u hostverify.res -T timelimit -f hostfile test</code> . Note: 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.
6.	Verify OPA Fabric status and topology.	Type <code>opalinkanalysis -U -x snapshot_suffix all verifyall > \$FF_RESULT_DIR/linkanalysis.res 2>&1</code> . NOTE: 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. Note: 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. NOTE: This test identifies nodes whose performance is not consistent with others in the fabric. It is not intended as a benchmark of fabric latency and bandwidth. This test intentionally uses techniques to reduce test runtime.	Type <code>opahostadmin -f hostfile mpipeperfdeviation</code> .
	• If all hosts pass...	Continue to the next step.
continued...		



Step	Task/Prompt	Action
	<ul style="list-style-type: none">If any hosts fail...	<ul style="list-style-type: none">Carefully examine the failing hosts to verify the HFI models, PCIe slot used, BIOS settings, and any motherboard or BIOS settings related to devices on PCIe buses or slot speeds.Also verify that the HFI and any riser cards are properly seated. <p>Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more information.</p>
Optional Tasks		
11	Baseline the present fabric configuration for use in future fabric health checks. NOTE: This should be performed after configuring any additional Management Nodes.	Type <code>opaallanalysis -b</code> .
12.	Perform host and/or ISL cable testing. NOTE: The test allows for starting and stopping an extended Bit Error Rate test. The system prompts to clear hardware counters. NOTE: 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> .
Review the Host Verification		
13.	View <code>opahostadmin</code> result files.	Type <code>editor result_dir/result_file</code> . NOTE: 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"><code>test.res</code><code>test.log</code>
	End Task	

Next Steps

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



11.0 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 138.

11.1 Before You Begin

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

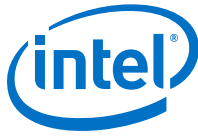
- If required, upgrade the software to add additional IFS components using the procedure documented in [Upgrade from OPA-Basic to OPA-IFS Software Package](#) on page 161.
 - The Management Node must have at least Intel® Omni-Path Fabric Suite FastFabric, the Intel® Omni-Path Fabric Stack, and should 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.

11.2 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 opafastfabric .
3.	From the main menu, access the FastFabric OPA Host Setup menu .	Press 3 .
	a) Edit Configuration and Select/Edit Host File	Select 0 and press P 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 X to return to the main menu.
4.	Access the OPA Host Verification/Admin Menu .	Press 4 .
	a) Edit Configuration and Select/Edit Host File	Select 0 and press P 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 X 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 FastFabric OPA Host Setup menu .	Press 3 .
	a) (Linux) Set Up Password-Less SSH/SCP	Select 2 and press P to start the operation.
	b) Password for <code>root</code> on all hosts:	Type the password for <code>root</code> on all hosts and press Enter .
	c) Upon completing the task, you are returned to the FastFabric OPA Host Setup menu .	
	d) Refresh SSH Known Hosts	Select 9 and press P to start the operation.
	e) Complete the operation.	Press any key to end the operation.
	End Task	

Next Steps

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



11.3 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> .
	End Task	

Next Steps

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



12.0 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.

12.1 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*.

12.2 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*.

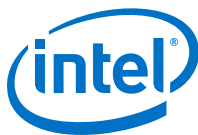
Assumption

- 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 FastFabric OPA Host Verification/Admin Menu .	
	a) If you are not already logged into Intel FastFabric OPA Tools , at the command prompt...	Type <code>opafastfabric</code> and press Enter .
	b) Access the OPA Host Verification/Admin Menu .	Press 4 .
2.	Edit Configuration and Select/Edit Host File (menu item 0)	Select item 0 and press P .
	a) Edit <code>opafastfabric.conf</code> file.	Review the following parameters: <ul style="list-style-type: none"> 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
3.	Check Overall Fabric Health (menu item 9)	Select item 9 and press P .
	a) Performing Host Admin: Check Overall Fabric Health Baseline present configuration? [n]:	Press Enter .
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) Check Overall Fabric Health (menu item 9)	Select item 9 and press P .
	b) Performing Host Admin: Check Overall Fabric Health Baseline present configuration? [n]:	Press y and press Enter .
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.
	End Task	

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](#).

12.3 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">• 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
continued...		



Step	Task/Prompt	Action
		<ul style="list-style-type: none"> FF_CHASSIS_CMDS FF_CHASSIS_HEALTH FF_ESM_CMDS
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 <code>opaallanalysis -e</code> .
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"> 1. Resolve the errors. 2. Rerun 3. and 4. until a clean run occurs.
5.	Create a cluster configuration baseline. <i>Note:</i> This may also be done using the FastFabric TUI menu by selecting Check Overall Fabric Health and answering y to the question: Baseline present configuration? [n]:	Type <code>opaallanalysis -b</code> .
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.
	End Task	

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](#).



13.0 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`.

13.1 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 funs:

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

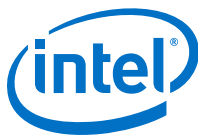
For example, `./config_hpl2 256s` selects the configuration and, subsequently, `./run_hpl2 256` runs the configuration.

13.2 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 <number of processes><problem size></code> . For example, <code>./config_hpl2 32s</code>
4.	Run the configuration.	Type <code>./run_hpl2 <number of processes><problem size></code> For example, <code>./run_hpl2 32</code>
5.	Review the results. NOTE: 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 <number of processes><problem size></code> . For example, <code>./config_hpl2 9216m</code>
8.	Run the configuration.	Type <code>./run_hpl2 <number of processes><problem size></code> For example, <code>./run_hpl2 9216</code>
9.	Review the results.	
	End Task	

Next Steps

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

13.3 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 Enter .
	b) # of cores per node [72]?	Type a number and press Enter .
	c) # of RAM per node (in MB) [64156]?	Type a number and press Enter .
	d) Memory pressure (range between 0.1 and 0.9) [0.3]?	Type a decimal number and press Enter .
	e) From the results, note the number of processes. This will be added to the run command: <code>./run_hpl2 <number of processes></code> . HPLinpack benchmark input file - 9216 processes , 0.5 memory size Generated by <code>hpl_dat_gen.sh</code> . HPL.out output file name (if any) 6 device out (6=stdout,7=stderr,file) 1 # of problems sizes (N) 733488 Ns 1 # of NBs 168 NBS 0 PMAP process mapping (0=Row-,1=Column-major) 1 # of process grids (P x Q) 72 Ps 128 Qs 16.0 threshold 1 # of panel fact 1 PFACTs (0=left, 1=Crout, 2=Right) 1 # of recursive stopping criterium 4 NBMINs (>= 1) 1 # of panels in recursion 2 NDIVs 1 # of recursive panel fact. 2 RFACTs (0=left, 1=Crout, 2=Right) 1 # of broadcast 1 BCASTs (0=1rg,1=1rM,2=2rg,3=2rM,4=Lng,5=LnM) 1 # of lookahead depth 1 DEPTHs (>=0) 2 SWAP (0=bin-exch,1=long,2=mix) 168 swapping threshold 0 L1 in (0=transposed,1=no-transposed) form 0 U in (0=transposed,1=no-transposed) form 1 Equilibration (0=no,1=yes) 8 memory alignment in double (> 0)	Continue to the next step.
3.	Run the configuration.	Type <code>./run_hpl2 <number of processes></code> . For example: <code>./run_hpl2 9216</code>
4.	Review the results.	
	End Task	

Next Steps

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

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 QSFP-DD Cables

If your hardware configuration uses QSFP-DD cables, you will need to perform the following steps to be able to query the cable information using command line tools.

1. On the Host, open the `/etc/opa-fm/opa_fm.xml` file for editing.
2. Search on "<CableInfoPolicy>".
3. Uncomment the line and change the value to `ByPort` as shown below.

```
<CableInfoPolicy>ByPort</CableInfoPolicy>
```
4. Save the `opa_fm.xml` file.
5. Restart the Fabric Manager using `systemctl restart opa_fm`.
6. Copy the `opa_fm.xml` file to other hosts.
7. Restart the Fabric Manager on the other hosts using `systemctl restart opa_fm`.

14.2 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. No further instructions are required from this section.

- **The subnets are primarily independent.**

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. To configure independent subnets, follow the instructions described in this section.



- **The subnets are overlapping.**

Multiple components are common to more than one subnet, such as FastFabric nodes, servers, and others. This includes dual rail fabrics with dual subnets. To configure overlapping subnets, follow the instructions described in this section.

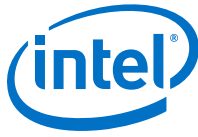
14.2.1 Setting Up Subnets

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 subnets.

Procedures

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

Step	Task/Prompt	Action
1.	Installation Prerequisites	
	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) Fabric Setup Prerequisites	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.	Install the Intel® Omni-Path Software	As stated.
3.	Configure the Software	
	a) Configure the Chassis	<ul style="list-style-type: none"> • 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. • 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.
	b) Configure an FM instance for each subnet.	Each subnet requires an FM instance and a unique Subnet Prefix. <ul style="list-style-type: none"> • If the FM nodes are connected to both subnets, refer to Configuring Dual Rails for Dual Subnets. • If the FM nodes are each connected to a single subnet, refer to Setting Up Dual Rails for a Single Subnet. Set up each FM with a unique subnet prefix.
	c) Configure the Externally-Managed Switches	<ul style="list-style-type: none"> • 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.
<i>continued...</i>		



Step	Task/Prompt	Action
		<ul style="list-style-type: none"> 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.
	d) Install the Host Software on the Remaining Hosts Using the FastFabric TUI Menu	<ul style="list-style-type: none"> 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. For hosts that are connected to more than one subnet, IPoIB must be configured such that each subnet is an independent IPoIB network interface, typically with different IP subnets. Refer to the <i>Intel® Omni-Path Fabric Host Software User Guide</i> for more information on configuring IPoIB.
	e) Verify the Host Software on the Remaining Servers Using the FastFabric TUI Menu	<ul style="list-style-type: none"> Edit Configuration and Select/Edit Host File 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. Verify Hosts See Each Other 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). Check MPI Performance 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).
	f) Configure Additional Management Nodes Using FastFabric TUI Menus	<p>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.</p> <p><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.</p>
	g) Configure and Initialize Health Check Tools Using FastFabric CLI Commands	<p>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.</p> <p>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.</p>
continued...		



Step	Task/Prompt	Action
	h) 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.
	i) 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	

14.3 Multi-Rail Usage

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

14.3.1 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.

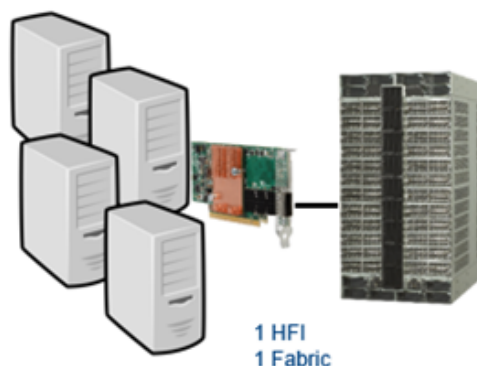
The multi-rail feature can be applied to a single subnet or multiple subnets. By enabling multi-rail, a process can use multiple network interface cards (HFIs) to transfer messages.

TERMINOLOGY:

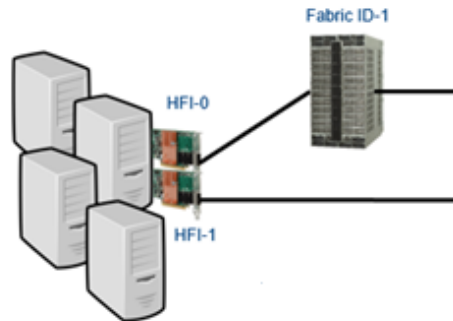
- Subnets can also be referred to as *planes* or *fabrics*.
 - Hosts can also be referred to as *nodes*.
 - HFIs can also be referred to as *rails*.
-

Three basic scenarios include:

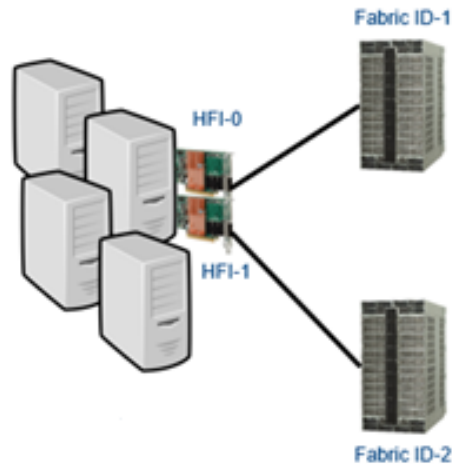
- Single-rail in a single subnet: This scenario, shown in the following figure, consists of one HFI in a server connected to one subnet. This is the default configuration during installation. This configuration provides the performance required by most applications in use today.



- Dual-rail in 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.

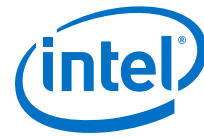


- Dual-rail in dual subnets: This scenario, shown in the following figure, consists of two HFIs in the same server connected to separate subnets. Depending on the platform, this configuration may provide improved MPI message rate, latency, and bandwidth to the node.



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.



14.3.2 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. In most scenarios, no adjustments to `opafm.xml` are required for this configuration.

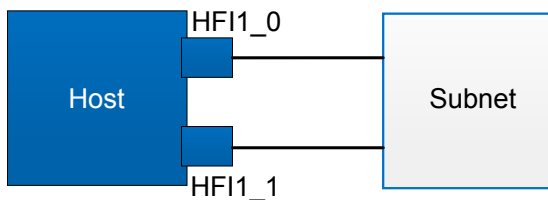
If `opafm.xml` is not modified, some APIs may not function correctly or may issue a warning (such as older versions of Open MPI) as shown below; however, both rails participate in the MPI application correctly.

WARNING: There are more than one active ports on host 'a', but the default subnet GID prefix was detected on more than one of these ports. If these ports are connected to different physical IB networks, this configuration will fail in Open MPI. This version of Open MPI requires that every physically separate IB subnet that is used between connected MPI processes must have different subnet ID values.

In this case, in order to prevent this warning, 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:



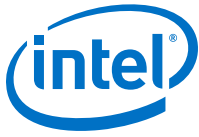
Important Information

Below is an overview of the steps you should take:

1. Stop all standby Fabric Manager using `systemctl stop opafm`.
2. Stop the master Fabric Manager using `systemctl stop opafm`.
3. Update the `opafm.xml` file of the master Fabric Manager as described below.
4. Update the `opafm.xml` file of all standby Fabric Manager as described below.
5. Restart the master Fabric Manager using `systemctl restart opafm`.
6. Restart the standby Fabric Managers using `systemctl restart opafm`.
7. Run `service opafm status` to verify that the Fabric Managers are running.

Procedures

Perform the following steps to set the SubnetPrefix:



1. Open the /etc/opa-fm/opafm.xml file for editing.
2. Search for "<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**.

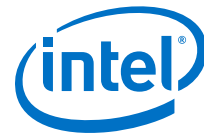
An example of the change is shown below.

```
<!-- 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. Save the opafm.xml file.

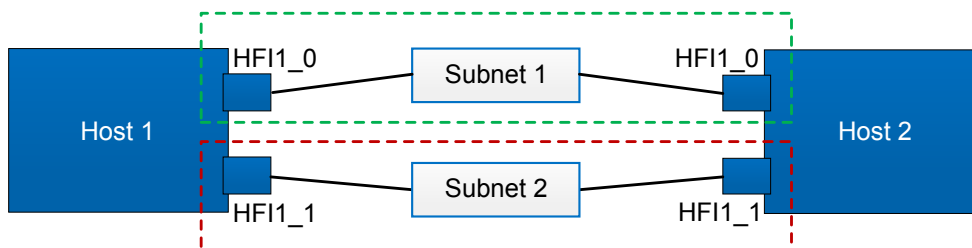


14.3.3 Configuring Dual Rails for Dual Subnets

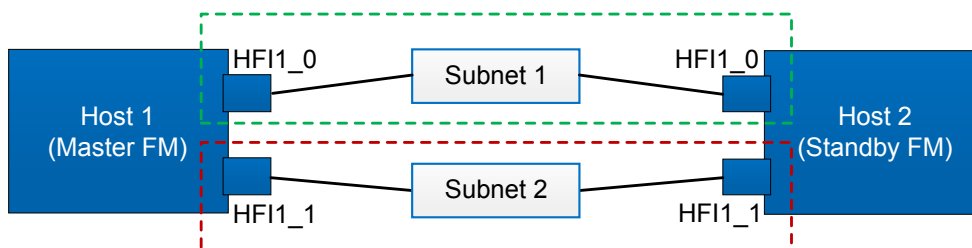
A common multi-rail, multi-subnet scenario is to set up dual rails for dual subnets where you configure two HFIs in each host to two different subnets to provide failover and load-balancing functions.

Assumptions

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



- Similarly, HFIs in the FMs have been cabled as shown below:



Important Information

Below is an overview of the steps you should take:

1. Stop all standby Fabric Manager using `systemctl stop opafm`.
2. Stop the master Fabric Manager using `systemctl stop opafm`.
3. Update the `opafm.xml` file of the master Fabric Manager as described below.
4. Update the `opafm.xml` file of all standby Fabric Manager as described below.
5. Restart the master Fabric Manager using `systemctl restart opafm`.
6. Restart the standby Fabric Managers using `systemctl restart opafm`.
7. Run `service opafm status` to verify that the Fabric Managers are running.

Procedures

Perform the following steps to set up the master Fabric Manager and standby Fabric Manager:

1. On the host, open the `/etc/opa-fm/opafm.xml` file for editing.



2. Search for "<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**.

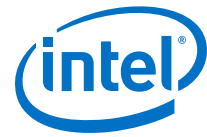
An example of the change is shown below.

```
<!-- 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>0xfe80000000001001</SubnetPrefix> <!-- should be unique -->
  <!-- Overrides of the Common.Shared or Fm.Shared parameters if desired -->
  <!-- <LogFile>/var/log/fm1_log</LogFile> --> <!-- log for this instance -->
  <!-- <StartupRetries>5</StartupRetries> -->
  <!-- <StartupStableWait>10</StartupStableWait> -->
</Shared>
```

5. Edit **fm1** settings as shown below:
 - Start = **1** (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>0xfe80000000001001</SubnetPrefix> <!-- should be unique -->
  <!-- Overrides of the Common.Shared or Fm.Shared parameters if desired -->
  <!-- <LogFile>/var/log/fm1_log</LogFile> --> <!-- log for this instance -->
  <!-- <StartupRetries>5</StartupRetries> -->
  <!-- <StartupStableWait>10</StartupStableWait> -->
</Shared>
```

7. Save the **opafm.xml** file.

14.3.4 IP Routing Guidance

In scenarios where two IPoFabric interfaces on the same host (for example, **ib0/ib1**) are configured on the same subnet, care must be taken to define the Linux routing tables for the desired behavior. This is because Linux may always use **ib0** to send return packets even if the communication request is initiated over **ib1**. This also should be configured for proper VRRP operation (refer to the *Intel® Omni-Path IP and LNet Router Design Guide* for details on setting up VRRP).

```
HostA:
ib0: 192.168.1.1
ib1: 192.168.2.1

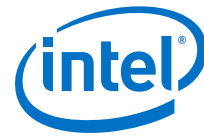
HostB:
ib0: 192.168.1.2
ib1: 192.168.2.2
```



If a netmask, such as 255.255.0.0, is defined for the above hosts, communication over ib1 between the hosts may not work correctly. In this scenario, define the following routes on both hosts:

```
ip route add 192.168.1.0/24 via ib0  
ip route add 192.168.2.0/24 via ib1
```

Alternatively, you can add these definitions to route-ib0/ib1 files. Refer to the *Intel® Omni-Path IP and LNet Router Design Guide*, Adding Static Routes for more detail.



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/neigh/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 OPA 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, if applicable, 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 *Intel® Omni-Path Fabric Software Release Notes*, OS RPMs Installation Prerequisites.

16.2 Download the Intel® Omni-Path Software

If the OS you installed did not include the OPA RPMs, download the software package from the Intel Resource & Design Center using the following procedures.

1. Using a web browser, go to <https://www.intel.com/content/www/us/en/design/products-and-solutions/networking-and-io/fabric-products/omni-path/downloads.html>.
2. Click on Latest Release Library.
3. In the table, click the drop-down arrow for Software and Drivers to select the files you need for the OS you have installed on your fabric.

Or, use the search box above the table to enter your search criteria.

- For OPA-Basic software, select:

```
Intel® Omni-Path Basic Software - DISTRO.VERSION - OPA release
```

- For OPA-IFS software, select:

```
Intel® Omni-Path IFS Software - DISTRO.VERSION - OPA release
```

NOTE

DISTRO.VERSION refers to the distribution and CPU.

4. Locate your target software and click the **Download** button.
5. Review the Intel Software License Agreement.
6. Click **Accept**.

The zipped software is saved to your computer.

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 153.
- You have downloaded and extracted the software package per [Upgrade Getting Started](#) on page 153.
- If your `opafm.xml` configuration file contains custom changes that you want to carry forward, make a backup copy of it.
- To upgrade the IFS software, you must first stop the standby Fabric Managers, then stop the master Fabric Manager. See [Stopping the Fabric Manager](#) on page 157.

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.

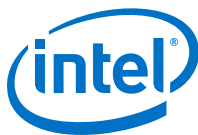
Assumption

- 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 Enter :
continued...		



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

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.

Optimally, to upgrade the OPA-IFS software, you perform the following sequence:

1. [Stop](#) all standby Fabric Managers.
2. [Stop](#) the master Fabric Manager.
3. [Upgrade](#) the master Management Node.
4. [Update](#) the `opa_fm.xml` file, as needed.
5. [Reboot](#) the master node. (See note below.)
6. [Upgrade](#) the standby Management Nodes.
7. [Update](#) the `opa_fm.xml` file, as needed.
8. [Reboot](#) the standby nodes. (See note below.)
9. [Upgrade](#) the remaining servers.
10. [Reboot](#) the remaining servers.



NOTE

If you did not enable autostart during the upgrade, you must manually restart the Fabric Manager at this time. Refer to [Starting the Fabric Manager](#) on page 159.

17.4.1 Stopping the Fabric Manager

This section provides instructions for stopping the Fabric Manager using CLI commands.

NOTE

You can also stop the Fabric Manager using the Chassis Viewer. Refer to *Intel® Omni-Path Fabric Switches GUI User Guide*, Stopping the Fabric Manager.

To stop the Fabric Manager, perform the following steps:

1. Log into the Fabric Manager system as `root` or as a user with `root` privileges.
2. Stop the Fabric Manager:

```
systemctl stop opafm
```

17.4.2 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 have stopped all standby FMs, followed by the master FM.
- 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 Enter : <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 Enter .
3.	Select 1) Install/Uninstall Software.	Type 1 .
4.	Review the items to be upgraded.	Accept the defaults. Type N to go to the next page. NOTE: 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 P to perform the actions.
6.	For each system prompt...	Accept the defaults by pressing Enter to continue.
continued...		



Step	Task/Prompt	Action
7.	When the Intel OPA Autostart Menu displays, review the items.	Intel recommends leaving all of the Autostart selections set to the default values. NOTE: 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 P .
9.	For each system prompt, "Hit any key to continue..."	Press any key. NOTE: When the installation completes, you are returned to the main menu.
10.	Exit the installation menu.	Type X .
11.	Reboot the server. <i>Note:</i> If your next step is to update the <code>opafm.xml</code> configuration file, do not reboot the server at this time.	Type <code>reboot</code> and press Enter . <i>Note:</i> If OPA Autostart was disabled, you must restart the master FM, followed by the standby FMs. Refer to Starting the Fabric Manager .
	End Task	

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](#).

17.4.3 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> . NOTE: The resulting differences will highlight your previous customizations. 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.
continued...		



Step	Task/Prompt	Action
3.	Update the new opafm.xml with customizations necessary for your fabric.	Use the resulting differences file from Step 1 to locate the customizations and add them to the new opafm.xml.
4.	Update /etc/opa-fm/opafm.xml.	Use the new file from Step 3 to replace your /etc/opa-fm/opafm.xml file.
5.	Restart the Fabric Manager	<ul style="list-style-type: none"> If you previously upgraded the management node, type <code>reboot</code> and press Enter. <i>Note:</i> If OPA Autostart was disabled, you must restart the master FM, followed by the standby FMs. Refer to Starting the Fabric Manager. If you only updated the opafm.xml file, type <code>systemctl restart opafm</code> and press Enter.
	End Task	

Next Steps

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

17.4.4 Starting the Fabric Manager

This section provides instructions for starting the Fabric Manager using CLI commands.

NOTE

You can also start the Fabric Manager using the Chassis Viewer. Refer to *Intel® Omni-Path Fabric Switches GUI User Guide*, Starting the Fabric Manager

To start the Fabric Manager, perform the following steps:

- Log into the Fabric Manager system as `root` or as a user with `root` privileges.
- Start the Fabric Manager:

```
systemctl start opafm
```

- Verify that all the tasks are up and running:

```
systemctl status opafm
```

17.4.5 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
Configuring the Hosts		
1.	Access the FastFabric OPA Host Setup menu.	
	a) If you are not already logged into Intel FastFabric OPA Tools , at the command prompt...	Type <code>opafastfabric</code> and press Enter .
	b) Access the FastFabric OPA Host Setup menu.	Press 3 .
	c) Select menu items.	<ul style="list-style-type: none"> Select items 0 (optional) and 5.
	d) Start the operations.	Press P . NOTE: Each selected item is preformed in the order of the menu list.
2.	(Optional) Edit Configuration and Select/Edit Host File (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.
	c) Do you want to edit/review/change the files? [y]:	Type n and Press Enter .
	e) Continue to next step.	Press any key to continue.
3.	Install/Upgrade OPA Software (menu item 5)	
	a) Do you want to use <code>./IntelOPA-[Basic IFS].DISTRO.VERSION.tgz</code> ? [y]:	Press Enter to accept the default.
	b) Would you like to do a fresh [i]nstall, an [u]pgrade or [s]kip this step? [u]:	Press Enter to accept the default.
	c) Are you sure you want to proceed? [n]:	Type y and press Enter .
	d) When the Intel OPA Autostart Menu displays, review the items.	Intel recommends leaving all of the Autostart selections set to the default values. NOTE: 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 P .
	f) For each system prompt, "Hit any key to continue..."	Press any key. NOTE: When the installation completes, you are returned to the main menu.
	g) Complete the installation.	Press any key to continue.
	<ul style="list-style-type: none"> If all hosts install... 	Press any key to continue.
	<ul style="list-style-type: none"> If any hosts fails to install... 	Use the View opahostadmin Result Files menu item to review the result files from the update.
Optional Tasks		
Refer to the <i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i> for more information.		
4.	Reboot Hosts (menu item 8)	(Linux) 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.
continued...		



Step	Task/Prompt	Action
5.	Run a Command on All Hosts (menu item b)	<p>(Linux) 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.</p> <p><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.</p>
6.	Copy a File to All Hosts (menu item c)	<p>This menu item specifies a file on the local host to be copied to all selected hosts.</p> <p><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.</p>
7.	View opahostadmin Result Files (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 Enter to save or type y to remove the files.
	End Task	

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.

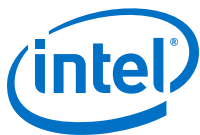
Assumptions

- 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 Enter : cd IntelOPA-IFS.DISTRO.VERSION where <i>DISTRO.VERSION</i> is the distribution and CPU.
2.	At the command prompt, start the install script.	Type ./INSTALL and press Enter .
3.	Select 1) Install/Uninstall Software.	Type 1 .
4.	Review the items to be upgraded. NOTE: Ensure that FastFabric and OPA_FM show Install and all other selections show Up To Date.	Type N to go to the next page. NOTE: If you need to change any item, enter the alphanumeric character associated with the item to toggle between Upgrade or Don't Install.
continued...		



Step	Task/Prompt	Action
5.	Start the upgrade.	Type P to perform the actions.
6.	For each system prompt...	Accept the defaults by pressing Enter 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. NOTE: 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 P .
9.	For each system prompt, "Hit any key to continue..."	Press any key. NOTE: When the installation completes, you are returned to the main menu.
10.	Exit the installation menu.	Type X .
11.	Reboot the server.	Type <code>reboot</code> and press Enter .
	End Task	

Next Steps

- If you are setting up additional Management Nodes, go to [Set Up Additional Management Nodes](#) on page 127.



Appendix A Software Installation Checklists

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

Print the appropriate checklists to monitor your progress.

A.1 OPA-Basic Checklists

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

A.1.2 Install the OPA-Basic Software Checklist

Step	Description	Complete
1.	Download and Extract the OPA-Basic Software Package. Refer to Installation Getting Started .	
2.	Install the Intel® Omni-Path Software .	



A.1.3 Upgrade the OPA-Basic Software Checklist

Step	Description	Complete
1.	Complete the steps in the Upgrade Prerequisites .	
2.	Download and Unpack the new OPA-Basic Software Package per Download the Intel® Omni-Path Software and Unpack the Tar File , respectively.	
3.	Upgrade the OPA-Basic Software on each compute node per Upgrade the OPA-Basic Software .	

A.2 OPA-IFS Checklists

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

A.2.2 Install and Configure the OPA-IFS Software Checklist

Step	Description	Complete
1.	Download and extract the OPA-IFS Software Package. Refer to Installation Getting Started .	
2.	Install the Intel® Omni-Path Software .	
3.	Configure the Chassis .	
4.	Configure the Externally-Managed Switches .	
5.	Install the Host Software on the Remaining Hosts Using the FastFabric TUI Menu .	
continued...		



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

A.2.3 Upgrade the OPA-IFS Software Checklist

Step	Description	Complete
1.	Complete the steps in the Upgrade Prerequisites .	
2.	Download and extract the new OPA-IFS Software Packages per Upgrade Getting Started .	
3.	Upgrade the Software on the Management Nodes.	
4.	Upgrade the Software on the Remaining Servers.	
5.	Verify the Host Software on the Remaining Servers Using the FastFabric TUI Menu.	



Appendix B Intel® OPA Software Components to Packages Mapping

To help you choose the proper packages for installation, the following table maps Intel® OPA Software components to their respective RPM packages.

Table 2. Intel® OPA Software Components Mapping

Component	Package	Description
	opaconfig	Configuration tool for the Intel® Omni-Path Fabric Suite FastFabric
opa_stack	opa-scripts	User space initialization scripts for the Intel® Omni-Path Fabric
	kmod-ifs-kernel-updates	Kernel modules built for the Linux* kernel family of processors
opa_stack_dev	ifs-kernel-updates-devel	Development headers for Intel HFI1 driver interface
oftools	opa-basic-tools	Basic tools for fabric management necessary on all compute nodes
	opa-address-resolution	ibacm-distributed SA provider (dsap) for name and address resolution on Intel® OPA platform. It also contains the library and tools to access the shared memory database exported by dsap.
intel_hfi	hfi1-firmware	Firmware required by the Gen1 HFI driver
	hfi1-firmware_debug	Debug signed firmware required by the Gen1 HFI driver
	libpsm2	PSM2 Messaging API, or PSM2 API, is the low-level, user-level communications interface for the Intel® OPA family of products. PSM2 users are enabled with mechanisms necessary to implement higher level communications interfaces in parallel environments.
	libpsm2-devel	Intel® PSM2, psm2*.h, headers and libpsm2.so files necessary for developing software using libpsm2
	libpsm2-compat	Support for MPIs linked with PSM versions less than 2. This will allow software compiled to use Intel® True Scale PSM libinfinipath to run with Intel® OPA PSM2 libpsm2.
	libfabric	User-space API to access high-performance fabric services, such as RDMA
	libfabric-devel	Development files for the libfabric library
continued...		



Component	Package	Description
	libfabric-psm	PSM provider as a <i>plugin</i> to an existing libfabric installation. This plugin will override older PSM provider functionality in the existing libfabric installation.
	libfabric-psm2	PSM2 provider as a <i>plugin</i> to an existing libfabric installation. This plugin will override older PSM2 provider functionality in the existing libfabric installation.
	libfabric-verbs	Verbs provider as a <i>plugin</i> to an existing libfabric installation. This plugin will override older verbs provider functionality in the existing libfabric installation.
	hfi1-diagtools-sw	User utilities for the Intel® Omni-Path Host Fabric Interface
	hfi1diags	Diagnostic and debug utilities for the Intel® Omni-Path Host Fabric Interface
fastfabric	opa-fastfabric	Tools for managing the fabric on a management node
	opa-mpi-apps	Applications and source for testing MPI performance in conjunction with opa-fastfabric or as standalone
delta_ipoib		
opa_fm	opa_fm	Intel Intel® Omni-Path fabric management applications. This includes: the Subnet Manager, Baseboard Manager, Performance Manager, Fabric Executive, and some fabric management tools.
	opa_fm-debuginfo	Debug information for package opa_fm. Debug information is useful when developing applications that use this package or when debugging this package.
opamgt_sdk	opa-libopamgt	Library necessary to build applications that interface with an Intel® Omni-Path Fabric Suite Fabric Manager
	opa-libopamgt-devel	Necessary headers for opamgt development
mvapich2_gcc_hfi	mvapich2_gcc_hfi	MPI-3 implementation that includes all MPI-1 features. It is based on MPICH2 and MVICH.
	mpitests_mvapich2_gcc_hfi	Set of popular MPI benchmarks: <ul style="list-style-type: none"> • IMB • OSU benchmarks
openmpi_gcc_hfi	openmpi_gcc_hfi	Tools necessary to compile, link, and run Open MPI jobs
	mpitests_openmpi_gcc_hfi	Set of popular MPI benchmarks: <ul style="list-style-type: none"> • IMB • OSU benchmarks
continued...		



Component	Package	Description
openmpi_gcc_cuda_hfi	openmpi_gcc_cuda_hfi	Tools necessary to compile, link, and run Open MPI jobs
	mpitest_openmpi_gcc_cuda_hfi	Set of popular MPI benchmarks: <ul style="list-style-type: none">• IMB• OSU benchmarks
mpisrc	mpi-selector	Tool that allows system administrators to set a site-wide default for which MPI implementation is to be used, but also allows users to set their own defaults MPI implementation, thereby overriding the site-wide default
delta_debug		