

Intel[®] Omni-Path Fabric Software

Release Notes for 10.5.1

September 2017



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

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

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

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

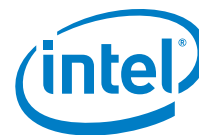
Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or by visiting: <http://www.intel.com/design/literature.htm>

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

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

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

Copyright © 2015-2017, Intel Corporation. All rights reserved.

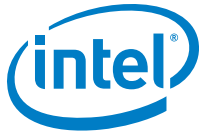


Contents

1.0 Overview of the Release	6
1.1 Introduction	6
1.2 Audience	6
1.3 Software License Agreement	6
1.4 If You Need Help	6
1.5 Enhancements and Features in this Release	6
1.6 Supported Features	7
1.7 Release Packages	8
1.8 Firmware Files	9
1.9 Operating Systems	10
1.10 Parallel File Systems	10
1.11 MPI Libraries	11
1.11.1 Supported MPI Libraries	11
1.11.2 Compiler Versions and Distributions	11
1.12 Intel Hardware	12
1.13 Documentation Versions	14
1.14 Installation Requirements	15
1.14.1 Software and Firmware Requirements	15
1.14.2 Installation Instructions	15
1.14.3 Installation Path Changes in Release 10.4 (and later) PR136906_cmt1 & PR138780	15
1.15 Product Constraints	16
1.16 Product Limitations	16
1.17 RHEL* 6.7 and CentOS* 6.7 Support	17
1.17.1 RHEL* 6.7 and CentOS* 6.7 Limitations	17
1.17.2 RHEL* 6.7 and CentOS* 6.7 Installation Prerequisites	17
1.17.3 RHEL* 6.7: Building Lustre* Kernel Modules for Intel® Omni-Path SupportPR136650: stays w/6.7 support	17
1.18 Accelerated RDMA Information	18
2.0 Issues	19
2.1 Introduction	19
2.2 Resolved Issues	19
2.3 Open Issues	31
3.0 Related Information	40
3.1 Documentation	40

Tables

1-1 Release Compatibility Matrix	9
1-2 Firmware Files	9
1-3 Supported Operating Systems	10
1-4 Supported MPI Libraries	11
1-5 Compiler Versions and Distributions	12
1-6 Supported Hardware	12
1-7 Supported Documentation Versions	14
2-1 Issues resolved in this release	19
2-2 Issues resolved in prior releases	24
2-3 Issues resolved in 10.3 release	27
2-4 Open Issues	31
3-1 Intel® Omni-Path Documentation Library	40





1.0 Overview of the Release

1.1 Introduction

This document provides a brief overview of the changes introduced into the Intel® Omni-Path Software by this release. References to more detailed information are provided where necessary. The information contained in this document is intended as supplemental information only; it should be used in conjunction with the documentation provided for each component.

These Release Notes list the features supported in this software release, open issues, and issues that were resolved during release development.

1.2 Audience

The information provided in this document is intended for installers, software support engineers, service personnel, and system administrators.

1.3 Software License Agreement

This software is provided under license agreements and may contain third-party software under separate third-party licensing. Please refer to the license files provided with the software for specific details.

1.4 If You Need Help

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.

1.5 Enhancements and Features in this Release

The following enhancements and features are new in the 10.5.1 release:

- Compatibility with Intel® HPC Orchestrator.
- Legacy BIOS Boot Mode Enhancements to support boot over fabric, custom board descriptions, and pre-boot platform configuration data for AOC support.
- Topology-aware job scheduling, which is enabled by the opamgt library and allows developers to write code that interfaces to the SA/PA. See the *Intel® Omni-Path Architecture Management API Programmer's Guide* for details.
- Multi-endpoint functionality. See the *Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide* for details.
- SNMP MIB support. See the *Intel® Omni-Path Fabric Switches Release Notes* for details.
- Support for Power Class 2 active optical cables (AOC). See [Product Constraints](#) for more information.



- Product Constraints described in [Section 1.15](#).

1.6 Supported Features

- The list of supported operating systems is in [Table 1-2](#).
- The list of supported hardware is in [Table 1-5](#).
- Active Optical Cables (AOC). For details, see the Cable Matrix at: <http://www.intel.com/content/www/us/en/high-performance-computing-fabrics/omni-path-cables.html>
- Support for the Enhanced Hypercube Routing Engine is outside the scope of Intel® OPA support. However, Intel partners may offer such support as part of their solutions. In addition there is an open source community who may be able to answer specific questions and provide guidance with respect to the Enhanced Hypercube Routing Engine.
- Support for Accelerated RDMA, also called Token ID (TID) RDMA, which is a Verbs protocol extension. See [Section 1.18](#) for details.
- Support for active optical cables (AOC) on server platforms using integrated HFI for OPA (commonly known as "-F").
- Support for GPUDirect* RDMA, which is a technology that enables a direct path for data exchange between a graphics processing unit (GPU) and a third-party peer device using standard features of PCI Express.
- Support for OpenFabrics Interfaces (OFI), a framework that includes libraries (including libfabric) and applications used to export fabric communication services to applications.
- Updated Accelerated RDMA support to include RHEL* 6.7. See [Section 1.18](#) for details.
- Signed Kernel Modules, as required to support UEFI Secure Boot
- Support for NVMe over Fabric Protocol
- Support for IBM* Platform MPI and IBM* Spectrum MPI. See [Table 1-3](#) for details.
- Virtual Fabric creation has been enhanced to better support advanced topologies, including the ability to place multicast traffic on a separate SL from unicast traffic. For details, see the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide*, section 2.
- In Release 10.4, changes were made to the installation path for all Intel® Omni-Path Software files. See [Section 1.14.3](#) for details and user action.

1.7 Release Packages

There are two Intel® Omni-Path Fabric Software packages:

- Basic for compute nodes
- IFS for management nodes

The Basic package includes:

- Software that installs the following packages to the distribution OpenFabrics Alliance* (OFA):
 - libibumad is based on the RHEL* or SLES* release package. It contains Intel patches that support Intel® Omni-Path Architecture (Intel® OPA) technology.
 - ibacm is the latest upstream code applied with RHEL* patches.
 - hfi1-firmware, hfi1-psm, hfi1-diagtools-sw, libhfi1verbs



- Open MPI and MVAPICH2. See [Section 1.11](#) for details.
 - mpitests
 - mpi-selector
 - GASnet
 - openSHMEM
 - srptools (includes the latest upstream code)
 - Firmware files listed in [Table 1-1](#).
- compat-rdma which delivers kernel changes based on the OFA version. The components installed are the hfi1 driver and Intel-enhanced versions of other kernel packages. See the *Building Lustre* Servers with Intel® Omni-Path Architecture Application Note* for details.

Note: In the Intel® Omni-Path Software package for RHEL* 7.2, the hfi1 driver and ifs-kernel-updates are supplied as a smaller package.
 - Firmware installation tools, including hfi1_eprom and TMM update tools.

The IFS package includes the Basic package plus:

- Fabric Manager, which allows comprehensive control of administrative functions using a mature Subnet Manager. Fabric Manager simplifies subnet, fabric, and individual component management, easing the deployment and optimization of large fabrics.
- Fabric Suite FastFabric Toolset, which enables rapid, error-free installation and configuration of Intel® OPA host software and management software tools, as well as simplified installation, configuration, validation, and optimization of HPC fabrics. For details, refer to the Fabric Suite FastFabric documentation.

1.8 Firmware Files

This release of the Intel® Omni-Path Software contains the firmware files listed in [Table 1-1](#). Intel provides UEFI firmware for discrete Intel® Omni-Path HFI cards and includes a platform file specific to Intel® Omni-Path HFI cards.

Table 1-1. Firmware Files

Description	File Name	Version
HFI1 UEFI Option ROM	HfiPcieGen3_1.5.2.0.0.efi	1.5.2.0.0
UEFI UNDI Loader	HfiPcieGen3Loader_1.5.2.0.0.rom	1.5.2.0.0
HFI1 SMBus Microcontroller Firmware (Thermal Monitor)	hfi1_smbus.fw	10.4.0.0.146
Intel® Omni-Path HFI platform file Note: If you have a non-Intel HFI, contact the manufacturer's support team for details.	hfi1_platform.dat	HFI_TYPE1 v1.0.1.0

1.9 Operating Systems

This release of the Intel® Omni-Path Software supports the operating systems listed in [Table 1-2](#).



Table 1-2. Supported Operating Systems

Operating System	Update/SP	Kernel Version
Red Hat* Enterprise Linux* (RHEL*) 6.7 X86_64	Update 7	2.6.32-573.el6.x86_64
CentOS* 6.7 X86_64	Update 7	2.6.32-573.el6.x86_64
Red Hat* Enterprise Linux* (RHEL*) 7.2 X86_64	Update 2	3.10.0-327.el7.x86_64
Red Hat* Enterprise Linux* (RHEL*) 7.3 X86_64	N/A	3.10.0-514.el7.x86_64
CentOS* 7.2 X86_64	N/A	3.10.0-327.el7.x86_64
CentOS* 7.3 X86_64	N/A	3.10.0-514.el7.x86_64
Scientific Linux* 7.2 X86_64	N/A	3.10.0-327.el7.x86_64
Scientific Linux* 7.3 X86_64	N/A	3.10.0-514.el7.x86_64
SUSE* Linux* Enterprise Server (SLES*) 12.1 X86_64	Service Pack 1	3.12.49-11.1-default
SUSE* Linux* Enterprise Server (SLES*) 12.2 X86_64	Service Pack 2	4.4.21-69-default

Note: The Intel® Xeon Phi™ Processor x200 product family (Knights Landing CPU-based servers) supports the following operating systems: RHEL* 7.3, CentOS* 7.3, and SLES* 12 SP2.

Note: PSM2 GPUDirect* RDMA is supported on RHEL* 7.2, RHEL* 7.3, and SLES* 12 SP2.

1.10 Parallel File Systems

The following parallel file systems have been tested with this release of the Intel® Omni-Path Software:

- Intel® Enterprise Edition for Lustre* software v3.1
 - RHEL* versions supported by Intel® Omni-Path Software.
- IBM* Spectrum Scale version 4.2.1.0
 - RHEL* 7.2.

Refer to the *Intel® Omni-Path Fabric Performance Tuning User Guide* for details on optimizing parallel file system performance with Intel® Omni-Path Software.

1.11 MPI Libraries

1.11.1 Supported MPI Libraries

The table below lists the different MPI libraries supported by Intel® Omni-Path Fabric Software with the corresponding version, fabric support, and compiler used. Note that the second column indicates whether the MPI library is included in the Intel® Omni-Path Software package or not.

Table 1-3. Supported MPI Libraries (Sheet 1 of 2)

MPI Implementation	Included in Basic package?	Runs Over	Compiled With
Open MPI 1.10.4	Yes	Verbs	GCC
	Yes	PSM2	GCC, Intel

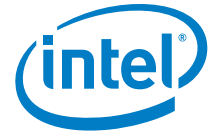


Table 1-3. Supported MPI Libraries (Sheet 2 of 2)

MPI Implementation	Included in Basic package?	Runs Over	Compiled With
Open MPI 1.10.4-cuda	No	Verbs	N/A
	Yes	PSM2	GCC, Intel
MVAPICH2-2.2	Yes	Verbs	GCC
	Yes	PSM2	GCC, Intel
Intel® MPI Library 2017 Update 3	No	Verbs	N/A
	No	PSM2	N/A
IBM* Platform* MPI version 9.1.4.3	No	Verbs	N/A
	No	PSM2	N/A
IBM* Spectrum* MPI version 10.1.0	No	Verbs	N/A
	No	PSM2	N/A

1.11.2 Compiler Versions and Distributions

The MPI libraries listed in the preceding table that are included in the release and built with PSM2 support were built with the following compiler versions:

Table 1-4. Compiler Versions and Distributions

Compiler	Linux* Distribution	Compiler Version
(GNU) gcc	RHEL* 7.2	gcc (GCC) 4.8.5 20150623 (Red Hat* 4.8.5-4)
(GNU) gcc	RHEL* 7.3	gcc (GCC) 4.8.5 20150623 (Red Hat* 4.8.5-11)
(GNU) gcc	SLES* 12 SP 1	gcc (SUSE* Linux*) version 4.8.5
(GNU) gcc	SLES* 12 SP 2	gcc (SUSE* Linux*) version 4.8.5
(Intel) icc	RHEL* 7.2	icc (ICC) 2017.0.4.196
(Intel) icc	RHEL* 7.3	icc (ICC) 2017.0.4.196
(Intel) icc	SLES* 12 SP 1	icc (ICC) 2017.0.4.196
(Intel) icc	SLES* 12 SP 2	icc (ICC) 2017.0.4.196

Note: Refer to the *Intel® Omni-Path Fabric Host Software User Guide* for set up information when using Open MPI with the SLURM PMI launcher and PSM2.

1.12 Intel Hardware

Table 1-5 lists the Intel hardware supported in this release. The table does not include OEM-specific hardware, such as custom adapters and switches.

Note: The Intel® PSM2 implementation has a limit of four (4) HFIs.

Table 1-5. Supported Hardware (Sheet 1 of 2)

Hardware	Description
Intel® Xeon® Processor E5-2600 v3 product family	Haswell CPU-based servers
Intel® Xeon® Processor E5-2600 v4 product family	Broadwell CPU-based servers
Intel® Xeon® Scalable Processor	Skylake CPU-based servers

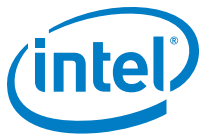
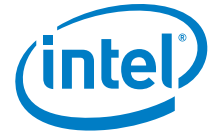


Table 1-5. Supported Hardware (Sheet 2 of 2)

Hardware	Description
Intel® Xeon Phi™ Processor x200 product family	Knights Landing CPU-based servers
Intel® Omni-Path Host Fabric Interface 100HFA016 (x16)	Single Port Host Fabric Interface (HFI)
Intel® Omni-Path Host Fabric Interface 100HFA018 (x8)	Single Port Host Fabric Interface (HFI)
Intel® Omni-Path Switch 100SWE48Q	Managed 48-port Edge Switch
Intel® Omni-Path Switch 100SWE48U	Externally-managed 48-port Edge Switch
Intel® Omni-Path Switch 100SWE48UFH	Externally-managed 48-port Edge Switch, hot-swap power and fans
Intel® Omni-Path Switch 100SWE48QFH	Managed 48-port Edge Switch, hot-swap power and fans
Intel® Omni-Path Switch 100SWE24Q	Managed 24-port Edge Switch
Intel® Omni-Path Switch 100SWE24U	Externally-managed 24-port Edge Switch
Intel® Omni-Path Director Class Switch 100SWD24	Director Class Switch 100 Series, up to 768 ports
Intel® Omni-Path Director Class Switch 100SWD06	Director Class Switch 100 Series, up to 192 ports

Note: For RHEL* 6.7 and CentOS* 6.7, only the following processors are supported:

- Intel® Xeon® Processor E5-2600 v3 product family
- Intel® Xeon® Processor E5-2600 v4 product family



1.13 Documentation Versions

Table 1-6 lists the end user document versions supported by this release.

Table 1-6. Supported Documentation Versions

Title	Doc. Number	Revision
Key:		
Shading indicates the URL to use for accessing the particular document.		
<ul style="list-style-type: none"> Intel® Omni-Path Switches Installation, User, and Reference Guides: http://www.intel.com/omnipath/SwitchPublications Intel® Omni-Path Software Installation, User, and Reference Guides (includes HFI documents): http://www.intel.com/omnipath/FabricSoftwarePublications Drivers and Software (including Release Notes): http://www.intel.com/omnipath/Downloads 		
<i>Intel® Omni-Path Fabric Quick Start Guide</i>	J57479	2.0
New title: <i>Intel® Omni-Path Fabric Setup Guide</i> (Old title: <i>Intel® Omni-Path Fabric Staging Guide</i>)	J27600	6.0
<i>Intel® Omni-Path Fabric Switches Hardware Installation Guide</i>	H76456	6.0
<i>Intel® Omni-Path Host Fabric Interface Installation Guide</i>	H76466	5.0
<i>Intel® Omni-Path Fabric Software Installation Guide</i>	H76467	7.0
<i>Intel® Omni-Path Fabric Switches GUI User Guide</i>	H76457	7.0
<i>Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide</i>	H76458	7.0
<i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i>	H76469	7.0
<i>Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide</i>	H76472	7.0
<i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i>	H76468	7.0
<i>Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide</i>	H76471	7.0
<i>Intel® Omni-Path Fabric Host Software User Guide</i>	H76470	7.0
<i>Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide</i>	H76473	7.0
<i>Intel® Omni-Path Fabric Performance Tuning User Guide</i>	H93143	9.0
<i>Intel® Omni-Path IP and Storage Router Design Guide</i>	H99668	5.0
<i>Building Lustre* Servers with Intel® Omni-Path Architecture Application Note</i>	J10040	1.0
<i>Building Containers for Intel® Omni-Path Fabrics using Docker* and Singularity* Application Note</i>	J57474	3.0
<i>Intel® Omni-Path Architecture Management API Programmer's Guide</i>	J68876	1.0
<i>Intel® Omni-Path Fabric Software Release Notes</i>	J75208	2.0
<i>Intel® Omni-Path Fabric Manager GUI Release Notes</i>	J75209	2.0
<i>Intel® Omni-Path Fabric Switches Release Notes</i> (includes managed and externally-managed switches)	J75207	2.0

For details on which document to use for a particular task, see Table 3-1 on page 26.

1.14 Installation Requirements

1.14.1 Software and Firmware Requirements

Table 1-2 lists the operating systems supported by this release. Refer to the *Intel® Omni-Path Fabric Software Installation Guide* for the required packages.



1.14.2 Installation Instructions

There are two Intel® Omni-Path Fabric Software packages:

- IntelOPA-IFS.<distro>-x86_64.<version>.tgz for the management node.
- IntelOPA-Basic.<distro>-x86_64.<version>.tgz for compute nodes.

The packages in the tgz file are RPMs. Installing individual RPMs is not supported in the 10.5.1 release.

Refer to the *Intel® Omni-Path Fabric Software Installation Guide* for related software requirements and complete installation procedures. Refer to the *Intel® Omni-Path Fabric Hardware Installation Guide* for related firmware requirements.

1.14.3 Installation Path Changes in Release 10.4 (and later)

If you are upgrading from a Intel® Omni-Path Fabric Software installation (Release 10.3 or earlier), Intel recommends that you perform the following steps before upgrading, due to changes in installation paths for RPMs and configuration files.

Run `./INSTALL -u` to uninstall existing packages.

Run `./INSTALL -a` to complete the installation.

Pre-existing configuration files are automatically saved by the RPM as `.rpmsave` files. (RPM will notify you about these files during removal.) If you want to keep these configuration files, you should move them to their new locations. A mapping of old configuration file locations to new locations is shown in the following table.

Old Location (Release 10.3 and earlier)	New Location (Release 10.4 and later)
/etc/sysconfig/opafm.xml	/etc/opa-fm/opafm.xml
/etc/sysconfig/allhosts	/etc/opa/allhosts
/etc/sysconfig/chassis	/etc/opa/chassis
/etc/sysconfig/esm_chassis	/etc/opa/esm_chassis
/etc/sysconfig/hosts	/etc/opa/hosts
/etc/sysconfig/opafastfabric.conf	/etc/opa/opafastfabric.conf
/etc/sysconfig/opaff.xml	/etc/opa/opaff.xml
/etc/sysconfig/opamon.conf	/etc/opa/opamon.conf
/etc/sysconfig/ports	/etc/opa/ports
/etc/sysconfig/switches	/etc/opa/switches

1.15 Product Constraints

- If you are upgrading from Release 10.5, refer to Technical Advisory #23 ([TA0023-10.5_Fabric_Manager_Workaround.pdf](#)) for detailed instructions on how to remove and restore normal FM performance for this release.
- Power class 2 AOC are supported and require 10.5 level firmware on both ends of the cable. Specifically, 10.5 host software and 1.5 level UEFI are required for proper operation. Integrated HFI (-F) requires a specific BIOS level to support power class 2 AOC; contact your BIOS vendor for more information.
- On certain systems with more than 1 HFI, the ports can come up in an unexpected way. This can lead to the wrong HFI being associated with `ib0` or `ib1`. To ensure ports come up in the expected order each time, use the module parameter `hf11`



port_reorder=1 and if you are loading the driver during initramfs boot, then ensure that you rebuild initramfs.

1.16 Product Limitations

This release has the following product limitations:

- The embedded version of the Fabric Manager supports a maximum of 100 HFI ports involving less than 20 switch ASICs. Calculate the number of switch ASICs in your fabric as follows:
 - 1 ASIC per Intel® Omni-Path Edge Switch 100 Series
 - 2 ASICs per Intel® Omni-Path Director Class Switch 100 Series Leaf module
 - 2 ASICs per Intel® Omni-Path Director Class Switch 100 Series Spine module
- Performance Administration (PA) Failover should **not** be enabled with FMs running on differing software versions.
To disable PA failover, edit the `/etc/sysconfig/opafm.xml` file and in the `<Pm>` section, change `<ImageUpdateInterval>` to 0.
- Enabling UEFI Optimized Boot on some platforms can prevent the HFI UEFI driver from loading during boot. To prevent this, do not enable UEFI Optimized Boot.

1.17 RHEL* 6.7 and CentOS* 6.7 Support

1.17.1 RHEL* 6.7 and CentOS* 6.7 Limitations

RHEL* 6.7 and CentOS* 6.7 are supported in this release with the following limitations:

- Processor support:
 - Intel® Xeon® Processor E5-2600 v3 product family
 - Intel® Xeon® Processor E5-2600 v4 product family
 - File system support:
 - GPFS
 - NFS
 - Intel® Enterprise Edition for Lustre* software
- Note:** For Intel® Enterprise Edition 3.0 Clients (support RHEL* 6.7) and Intel® Enterprise Edition 3.1 Servers (support RHEL* 7.3): You cannot upgrade your Clients beyond version 3.0 until you move to a newer RHEL* version.
- MVAPICH2 and Open MPI have been compiled for PSM2 to support the following versions of the compilers:

Compiler	Linux* Distribution	Compiler Version
(GNU) gcc	RHEL* 6.7 CentOS* 6.7	gcc (GCC) 4.4.7
(Intel) icc	RHEL* 6.7 CentOS* 6.7	icc (ICC) 15.0.1

- Performance is within 2%-5% of RHEL* 7.2 performance for the following features:
 - PSM bandwidth
 - MPI latency
 - Verbs bandwidth



1.17.2 RHEL* 6.7 and CentOS* 6.7 Installation Prerequisites

Starting in Release 10.5, refer to the *Intel® Omni-Path Fabric Software Installation Guide*, OS RPMs Installation Prerequisites section for the detailed list.

1.17.3 RHEL* 6.7: Building Lustre* Kernel Modules for Intel® Omni-Path Support

The Intel® Enterprise Edition for Lustre* kernel modules are dependent on InfiniBand* core modules. In order for them to load properly in a system where the OPA stack is installed, they must have been compiled against `Module.symvers`. The IFS package installs `Module.symvers`, which is generated while building InfiniBand* core and Intel® Omni-Path kernel modules. This file contains symbol information that can be used to build higher-level kernel modules such as Lustre* that are dependent on InfiniBand* core or Intel® Omni-Path kernel modules.

It is installed in:

```
/lib/modules/<kernel_ver>/include-ifs-kernel/Module.symvers
```

Build Lustre* modules using the following command:

```
# KBUILD_EXTRA_SYMBOLS="/lib/modules/<kernel_ver>/  
include-ifs-kernel/Module.symvers" rpmbuild -rebuild -without  
servers <pathtomyrpm>
```

1.18 Accelerated RDMA Information

Accelerated RDMA is a Verbs protocol extension to improve the performance of RDMA write and RDMA read operations on Intel® Omni-Path hardware.

This extension improves the efficiency of large message transfers to provide performance benefits for storage protocols and other Verbs-based protocols. The benefits include increased achievable bandwidth with reduced CPU utilization. The Token ID (TID) RDMA protocol accelerates the OpenFabrics Alliance* (OFA) Verbs API with no changes required to API consumers. The acceleration technique is performed by the host driver and the application running over the OFA Verbs API does not need to make any code change.

Accelerated RDMA is off by default.

To enable it, add `cap_mask=0x4c09a01cbba` to the `/etc/modprobe.d/hfi1.conf` file. Instructions on how to do this are in the *Intel® Omni-Path Fabric Performance Tuning User Guide*, "Setting HFI1 Driver Parameters" section.

Note: Accelerated RDMA must be enabled on all nodes to function. Mixing of enabled and disabled nodes will not show performance benefits.



2.0 Issues

2.1 Introduction

This section lists the resolved and open issues in the Intel® Omni-Path Software.

2.2 Resolved Issues

Table 2-1 lists issues that are resolved in this release.

Table 2-1. Issues resolved in this release (Sheet 1 of 3)

ID	Description	Resolved in Release
140199	In some scenarios involving device reboot, down links, fabrics with spare ports or fabrics with DCS; that the SM may skip setting some important SMA attributes. This can result in ports which are Active but unable to pass data, resulting in errors and failures from assorted applications depending on which nodes are communicating with each other.	10.5.1
132160	In Release 10.5, the Open MPI pre-compiled packages were built with a dependency on numactl being available at runtime.	10.5
133380	In Release 10.5, the PM has been updated allowing you to change the weight and threshold of PA categories. This enables you to recalculate values using already stored port data. Also, PA query time, memory usage, and disk space usage will decrease with the new PM History version. To change thresholds and weights, edit the opafm.xml file and restart the FM. Note: The FM no longer supports the previous short term history (STH) file after this change. The old files do not need to be removed as they will age out normally.	10.5
133596	When running the install script and installing all available packages, the installer may return a <code>prereq not installed</code> error.	10.5
134409	In links exhibiting a high error rate, a rare <code>PortRcvError</code> is possible, resulting in a link down event. Such links should retrain and return to operation without user interaction. In cases where the Link Quality is less than or equal to 3, the interconnect in the link should be evaluated for possible replacement to prevent future <code>PortRcvErrors</code> from occurring.	10.5
135259	In links exhibiting a high error rate, a rare <code>PortRcvError</code> is possible, resulting in a link down event. Such links should retrain and return to operation without user interaction. In cases where the Link Quality is less than or equal to 3, the interconnect in the link should be evaluated for possible replacement to prevent future <code>PortRcvErrors</code> from occurring.	10.5
135360	On a system running RHEL* 7.2 or 7.3, if two <code>kmem_cache_creates()</code> occur with the same name, a kernel panic may result when the caches are deleted. The kernel panic is caused by <code>hfil_user_sdma_free_queues</code> .	10.5



Table 2-1. Issues resolved in this release (Sheet 2 of 3)

ID	Description	Resolved in Release
135390	<p>Very old HFI adapters may be programmed with an obsolete version of the AOC platform configuration file.</p> <p>In these cases, errors such as the following may be observed:</p> <pre>[26.903186] hfi1 0000:d5:00.0: hfi1_0: parse_platform_config:Bad config file [26.903186] hfi1 0000:d5:00.0: hfi1_0: parse_platform_config:File claims to be larger than read size [27.351555] hfi1 0000:d5:00.0: hfi1_0: tune_serdes: Unknown port type</pre> <p>Update the platform configuration file on the HFI to the current version. For details, see the <i>Intel® Omni-Path Fabric Software Installation Guide</i>, section B.1.</p>	10.5
135929	<p>Intel® Omni-Path Boot nodes occasionally dropped from fabric when switching master SM from one node to another.</p>	10.5
135963	<p>Cannot install IFS software on RHEL* 7.3 using the command: <code>./INSTALL -vv -a</code></p>	10.5
136049	<p>The expected width of a card is not showing up correctly in <code>opaverifyhosts</code>.</p>	10.5
136436	<p>On SLES* 12.2, <code>node_desc</code> is not populated with the host name when system is booted up.</p>	10.5
136437	<p>When using RHEL* 7.2, the default generic PXE boot image does not work due to missing driver and firmware files.</p>	10.5
136567	<p>The Intel® OPA software does not support SSH passphrases. In earlier releases, if a passphrase was created, an error message was displayed to remove the passphrase and proceed.</p> <p>In Release 10.5, when SSH keys are created, the passphrase is automatically left empty.</p>	10.5
136727	<p>Initialization of PSM2 library fails with the following error message:</p> <pre>Error: PSM is in the finalized state</pre>	10.5
136821	<p>When performing boot over fabric, links may take up to 6 minutes to become active.</p>	10.5
136901	<p>Occasionally, nodes may be dropped by the Fabric Manager while they are in a pre-boot mode. This can occur when the node has multiple HFIs on a single socket.</p>	10.5
136971	<p>When using the Accelerated RDMA feature (TID RDMA), certain Verbs Multi-PPN tests may cause error messages.</p>	10.5
136996	<p>In Release 10.5, the <code>opaconfig</code> tool has been moved into a new top level RPM.</p>	10.5
137054	<p>Pinging an Intel® OPA UEFI permanent IP address from a DHCP server fails on subsequent reboots unless the corresponding network interface has first been initialized in the UEFI network stack.</p>	10.5
137123	<p>In Release 10.5, the Fabric Manager is not compatible with older versions of the FM GUI. You must use the same version of both Fabric Manager and FM GUI.</p>	10.5
137364	<p>The node description of a node may change after rebooting. This issue has been seen on RHEL* and SLES* and requires manual enabling and starting of the <code>rdma-ndd</code> service.</p>	10.5
137372	<p>Packets may be stuck in kernel when attempting writes to file system via IPoIB interface.</p>	10.5
137499	<p>HFI links may occasionally take several minutes to reach link up.</p>	10.5
137577	<p><code>opatmmtool</code> does not provide a correct error message if it is run on a system that does not have a TMM.</p>	10.5
137616	<p>When booting in legacy BIOS boot mode on RHEL* 7.x and SLES* 12.x, the following message is present in the kernel:</p> <pre>Request for unknown module key 'Intel Corporation: Intel(R) Omni-Path HFI UEFI: 719ebaa125172ba69ad01b850b7458f85c89bb07' err -11</pre>	10.5
137744	<p>In Release 10.4.2 and earlier, the values for <code>MinInitial</code> and <code>MinTail</code> were reported in flits by the <code>opareport</code>, <code>opasmaquery</code>, and <code>opasaquery</code> tools.</p> <p>In Release 10.5, this output is now converted to bytes, and is displayed in decimal.</p> <p>See the <i>Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide</i> for details.</p>	10.5
137791	<p>In Release 10.5, changes to settings in the Preemption section of the FM configuration file are only updated on HFI or switch ports when the port is bounced.</p> <p>See the <i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i> for details.</p>	10.5


Table 2-1. Issues resolved in this release (Sheet 3 of 3)

ID	Description	Resolved in Release
137869	When using GPUDirect* RDMA, you must run .INSTALL using the -G option. In Release 10.4, the INSTALL script did not pass the -G argument to remote hosts.	10.5
138047	The Open MPI implementation for MPI_Wtime() may change when using different CPU frequency drivers (intel_pstate vs acpi_freq) and turbo status of the CPU. Release 10.5 includes patches for Open MPI 1.10.x that resolve this issue.	10.5
138108	In Release 10.5, PKey handling for Active VFs was changed. To avoid disruptions when upgrading from a prior release, you must ensure that all Active VFs have explicit PKeys defined in the opafm.xml configuration file. To find PKeys that are currently assigned to each Active VF, type <code>opareport -o vfinfo</code> and press Enter . Using this information, manually edit the <VirtualFabric> section of the opafm.xml file for each VF in the list to insert the following: <Pkey>pkey_number</Pkey>	10.5
138171	During execution of an <code>opacapture</code> command, the core file may not be properly copied. This can occur on systems running the SLES* operating system when the following message appears during the capture: <code>/usr/lib/opa-fm/bin/fm_capture: line 148: [: too many arguments</code> If a core file exists in the default directory, that core file will be copied.	10.5
138183	In Release 10.4, additional fields were added to the <code>opareport -o snapshot -r</code> XML output format that were not present in Release 10.3. Therefore, Release 10.3 snapshot files reported a "Mandatory Tag Not Found" parser error using Release 10.4 Fabric Manager tools. The resolution was to regenerate any such snapshot files using the <code>opareport</code> tool in Release 10.4.	10.5
138188	Coexistence feature with Intel® True Scale HCA cards does not function on servers running RHEL* 7.3 OS.	10.5
138460	When upgrading from Release 10.3 and earlier to Release 10.4 or newer, the <code>allhost</code> path include statement was not updated with the new path which is: <code>/etc/opa/allhosts</code> This issue is resolved in Release 10.5.	10.5
139073	An internal I2C error may result in a Director leaf reinitializing (momentarily going off-line), which may disrupt fabric operation. In Release 10.5, the switch firmware now handles I2C errors such that leaf operation is not impacted.	10.5
139407	Shell history overflow caused by Intel® OPA commands run by scripts. In Release 10.5, commands run by <code>fastfabric opahostadmin</code> will now be omitted from the shell history.	10.5
139797	Switch port connected to HFI stuck in LinkTearDown state.	10.5
139931	Clarified Section 1.16, "Product Limitations" on page 13 : The embedded version of the Fabric Manager supports a maximum of 100 HFI ports involving less than 20 switch ASICs. Calculate the number of switch ASICs in your fabric as follows: <ul style="list-style-type: none"> • 1 ASIC per Intel® Omni-Path Edge Switch 100 Series • 2 ASICs per Intel® Omni-Path Director Class Switch 100 Series Leaf module • 2 ASICs per Intel® Omni-Path Director Class Switch 100 Series Spine module 	10.5

Table 2-2 lists issues that are resolved in prior releases.

Table 2-2. Issues resolved in prior releases (Sheet 1 of 3)

ID	Description	Resolved in Release
130336	<code>hfilstats</code> cannot be run at user level due to mount-point privileges.	10.4
131017	Verbs <code>ib_send_bw</code> , <code>ib_read_bw</code> , and <code>ib_write_bw</code> are not working with the -R option to use the RDMA CM API to create QPs and exch data.	10.4
134268	The Option ROM image (e.g. containing a UEFI driver) may not be executed if the BIOS configures the HFI Expansion ROM BAR with an address that is not 16MB aligned.	10.4
134353	Very infrequently, when a link goes down, the logical link state can remain stuck in the 'Init' state.	10.4



Table 2-2. Issues resolved in prior releases (Sheet 2 of 3)

ID	Description	Resolved in Release
134493	When using MVAPICH2 with Intel® Omni-Path PSM2, users will notice unexpected behavior when seeding the built-in random number generator with functions like srand or srandom before MPI_Init is called. MPI_Init re-seeds the random number generator with its own value and does not restore the seed set by the user application. This causes different MPI ranks to generate different sequences of random numbers even though they started with the same seed value.	10.4
134821	The UEFI network stack is initialized with a default network address before the driver receives a MAD packet containing an updated and actual subnet prefix. Therefore, in ARP and IP UEFI drivers the old (default) HW address is still used, causing problems with packet receiving and transmitting.	10.4
135040	You can't currently specify portions of an Intel® DCS chassis that is not populated and is not expected to be populated. If CoreFull is 1, all the internal links for that chassis are generated when run against opaxlattopology. If CoreFull is 0, none of the links are generated.	10.4
135068	Older versions of Grub 2 may not properly boot over Ethernet with the HFI UEFI driver.	10.4.2
135180	OpenMPI/PSM2 timeouts during MPI stress tests on Haswell and Intel® Xeon Phi™ mixed fabrics.	10.4
135326	Calling opasmaquery fails when called from a non-SM node to a node which has not booted to the OS.	10.4
135355	Due to changes in where the IFS packages are installed, customers using the FastFabric tools and upgrading to 10.3 from an earlier release must find each occurrence of /opt/opa in the opafastfabric.conf file and replace the string with /usr/lib/opa.	10.4
135545	A change has been made to several SA record attributes which causes incompatibilities between the Fabric tool suite and the SA.	10.4
135648	MPI applications are installed under the /usr/lib directory structure, which may be set up to be read-only overall. This causes resulting FastFabric operations to fail since mpi_apps contain source code and run scripts for sample MPI applications, test programs and benchmarks.	10.4
135711	After generating the opafm.xml file from the config_generate script, the FE is not enabled.	10.4
135873	hostverify.sh fails with RHEL* 6.7 due to the Intel P-State driver not being the default cpufreq driver.	10.4
136137	The hfi1_eprom tool man page contains incorrect information in the -d device option.	10.4
136733	Slow memory deregistration has been observed.	10.4
136902	A snapshot file with a multicast group with rate 10g will not be read properly. The following error is returned: opafabricanalysis: Port 0:0 Error: Unable to analyze fabric snapshot. See /var/usr/lib/opa/analysis/latest/fabric.0:0.links.stderr opafabricanalysis: Possible fabric errors or changes found	10.4
136945	When using the TID RDMA feature (Accelerated RDMA), certain MVAPICH over Verbs tests may cause error messages.	10.4
136985	opahfirew has output errors when the HFI driver is not installed.	10.4
136995	The opahfirew tool output uses the term "HWRev" to indicate the revision of the silicon on the card.	10.4
137015	The state and configuration of ipoib interfaces are controlled by the NetworkManager service. The NetworkManager in RHEL* 7.2 mistakenly assumes the ipoib interface is type 'ethernet' and fails to initialize it, due to a mismatch against its actual type which is 'infiniband'.	10.4
137096	The IFS package does not install all the RPMs that it contains. In particular, infiniband-diags and libibmad are not automatically installed. The absence of infiniband-diags may result in failure of node descriptions to be populated, such that all hosts have the same hfi1_0 description.	10.4
137108	When using the TID RDMA feature (Accelerated RDMA), virtual machines, and other cases where the IOMMU is enabled, do not operate correctly. This can lead to stability issues, and possibly data corruption, because the address used to receive data into will be incorrect.	10.4
137142	When using the TID RDMA feature (Accelerated RDMA), certain MPI benchmark tests may cause Kernel panic.	10.4

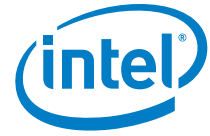


Table 2-2. Issues resolved in prior releases (Sheet 3 of 3)

ID	Description	Resolved in Release
137221	Querying for switch info with opasmaquery while using the -g option will print incorrect IPv4 addresses.	10.4
137708	Following a link bounce event, there is a possibility that a link will fail to reach the Armed/Active state. The likelihood of this issue depends largely on the link type: <ul style="list-style-type: none"> • Compute Nodes: These links are very unlikely to be affected. • FM Nodes: These links are the most exposed. If an FM link is affected and not recovered, there may be downstream effects over time. 	10.4.1
138920	Reverted to a prior version of 8051 firmware to resolve a link bring up reliability regression issue while interoperating with older link firmware.	10.4.2



2.3 Open Issues

Table 2-3 lists the open issues for this release.

Table 2-3. Open Issues (Sheet 1 of 6)

ID	Description	Workaround
129563	Memory allocation errors with MVAPICH2-2.1/Verbs.	When running MVAPICH2 jobs with a large number of ranks (for example, between 36 and 72 ranks), you must set the following parameters in <code>/etc/security/limits.conf</code> : * hard memlock unlimited * soft memlock unlimited Also, you must increase the <code>lkey_table_size</code> : LKEY table size in bits (2^n , $1 \leq n \leq 23$) from its default of 16 to 17. For instructions on setting module parameters, refer to Appendix A in the <i>Intel® Omni-Path Fabric Performance Tuning User Guide</i> .
131745	When running OpenMPI 1.10.0 on SLES* 12 with large number of ranks per node (over 40), it may happen that the ORTE daemon (orted) "hangs" during the finalization of job. This is an issue in Open MPI with the version of glibc used in SLES* 12. It is being researched by the Open MPI community in issue: https://github.com/open-mpi/ompi/issues/1136	Stopping and resuming the "hung" orted process allows the job to finish normally. To find the hung process, run the <code>ps</code> and find a node with several job zombie processes. In that same node, identify the orted process ID and send a stop signal (<code>kill -19 <PID></code>) and a continue signal (<code>kill -18 <PID></code>).
132207	Kernel crash caused by the <code>ib_srpt</code> module.	Install this kernel patch: https://git.kernel.org/cgit/linux/kernel/git/torvalds/linux.git/commit/?id=51093254bf879bc9ce96590400a87897c7498463
133604	Bonding driver shows incorrect hardware address of IPoIB interfaces.	This workaround applies to RHEL* 7.3 (and earlier) and SLES* 12 SP2 (and earlier) distributions: Use the <code>opa.info</code> command to retrieve the PortGUID and <code>ip addr show ib0</code> to get the correct 20-byte hardware address of OPA network interface.



Table 2-3. Open Issues (Sheet 2 of 6)

ID	Description	Workaround
133633	OpenMPI and MVAPICH2 compiles fail to link properly when using the Intel compilers. This issue is caused by a bug in Intel Compiler 2015.	<p>The workaround is to rebuild using Intel Composer 2017.</p> <p>For Open MPI: <code>export PATH=/opt/intel/bin:\$PATH</code> <code>export</code> <code>LD_LIBRARY_PATH=/opt/intel/lib/intel64/:\$LD_LIBRARY_PATH</code> <code>opaconfig #</code> Follow the prompts to rebuild MPI Library and Tools.</p> <p>For MVAPICH2: <code>mkdir /tmp/osu_mvapich2 && cd /tmp/osu_mvapich2</code> <code>wget http://mvapich.cse.ohio-state.edu/download/mvapich/mv2/mvapich2-2.2.tar.gz</code> <code># extract the tgz</code> <code># copy the modified dist_tarball.sh and wfr-mvapich2.spec.in into mvapich2-2.2.tar.gz</code> <code># Note the VERSION and RELEASE variables in the dist_tarball.sh were updated to 2.2 and davi0, respectively</code> <code>cd mvapich2-2.2/</code> <code># run dist_tarball.sh</code> <code>./dist_tarball.sh</code></p> <p><code># create an rpmbuild workspace and copy the source tarball to the rpmbuild/SOURCES directory</code> <code>cd ..</code> <code>mkdir -p rpmbuild/{BUILD,RPMS,SOURCES,SPECS,SRPMS}</code> <code>cp mvapich2-2.2/mvapich2-2.2-davi0.tar.gz rpmbuild/SOURCES</code> <code># build the srpm</code> <code>cd rpmbuild</code> <code>rpmbuild -ts --define "_topdir \$PWD" --nodeps SOURCES/mvapich2-*.tar.gz</code></p> <p><code># copy the resulting srpm to /opt/opa/src/MPI/</code> <code>cp SRPMS/mvapich2-2.2-davi0.src.rpm /opt/opa/src/MPI/</code> <code># rename any pre-existing mvapich2 srpms in /opt/opa/src/MPI/</code> <code>mv /opt/opa/src/MPI/mvapich2-2.1-10.src.rpm /opt/opa/src/MPI/_mvapich2-2.1-10.src.rpm</code></p> <p><code>#launch FF and follow the prompts to rebuild MVAPICH2 (captured in osu_mvapich2-2.2_FF_composer2017.res).</code></p>
134471	The HFI UEFI driver cannot boot via PXE using Grub 2.	Contact Intel Customer Support for assistance.
134494	Open MPI uses srand() family functions at MPI_Init() time. Therefore, if the user sets srand() before calling MPI_Init(), the values will be altered.	a) Fixed in Open MPI 2.0.1. b) Call srand() functions family after calling MPI_Init().
134819	In KNL-F EFI shell, the command <code>ifconfig -l</code> does not correctly display the IP address after being assigned via DHCP.	Launch a newer version of the EFI shell from the embedded shell.
134904	Legacy PXE boot using IPXE while the HFI UEFI driver is loaded may cause a hang.	Configure PXE operation to boot using UEFI boot mode.



Table 2-3. Open Issues (Sheet 3 of 6)

ID	Description	Workaround
135028	NVMe over Fabric Protocol is only supported on Intel® OPA with Linux* kernel 4.5 and later versions.	To use NVMe functionality on Intel® OPA, you must patch the kernel.
135084	In rare circumstances, the HFI may not appear in the PCI config space after a power cycle.	Reboot or power cycle the platform.
135390	<p>Very old HFI adapters may be programmed with an obsolete version of the AOC platform configuration file. In these cases, errors such as the following may be observed:</p> <pre>[26.903186] hfi1 0000:d5:00.0: hfi1_0: parse_platform_config:Bad config file [26.903186] hfi1 0000:d5:00.0: hfi1_0: parse_platform_config:File claims to be larger than read size [27.351555] hfi1 0000:d5:00.0: hfi1_0: tune_serdes: Unknown port type</pre>	Update the platform configuration file on the HFI to the current version. For details, see the <i>Intel® Omni-Path Fabric Software Installation Guide</i> , section B.1.
135975	After performing an OPA software configuration update, some unmanaged switches do not update the settings for LinkWidth and LinkWidthDnGrade enables.	A reboot is required for configuration changes made to an externally managed switch to become active.
136160	<p>On some Intel® Xeon Phi™ with integrated Intel® Omni-Path fabric platforms, the second integrated HFI is discovered first and is subsequently identified as the first HFI device. As a result, when issuing Intel® Omni-Path commands, the second HFI appears first in the results. In Linux* and various Intel® Omni-Path tools, the HFI reporting order may be the opposite of the order appearing on the Intel® Xeon Phi™ with integrated Intel® Omni-Path fabric cable/faceplate.</p>	<p>You can identify the second integrated HFI by inspecting the Node GUID or Port GUID/Port GUID reported by <code>opainfo</code> or other commands such as <code>hfi1_control -i</code>. Note that bit 39 of the PortGUID, the most significant bit, is set for the second HFI, and is clear for the first HFI.</p> <p>Keep in mind that when issuing various Intel® Omni-Path CLI commands targeted at a specific HFI using the <code>-h</code> option, <code>-h 1</code> correlates to the device that is listed as <code>hfi1_0</code>. As a result, the issued command affects the second HFI instance in cases where the second HFI port instance appears first.</p> <p>By default, ports are ordered as enumerated by the kernel. There is a new module parameter called <code>port_reorder</code>. When set, the HFI1 ports on the same ASIC will be enumerated in increasing order.</p> <p>To enable this feature, use the command: <code>modprobe hfi1 port_reorder=1</code></p>
136419	When running SLES* 12.2 with inbox OPA drivers installed, the state may not change from "Offline" to "Physical Linkup (Init)" as expected.	Add a <code>platform.dat</code> file in <code>/lib/firmware/updates</code> , then restart.
136432	Certain perftest tools such as <code>ib_write_bw</code> do not work on RHEL* 7.3 when using the RDMA CM with UD QPs.	Roll back the perftest package to the level found in RHEL* 7.2, which is perftest-2.4. Then install this package on RHEL* 7.3.
136500	RDMA perftests can hang on start on a client side when RDMA CM (-R option) is used.	Intel recommends that you use the same version of perftests across your fabric. Obtain the latest perftests version from the upstream repository.



Table 2-3. Open Issues (Sheet 4 of 6)

ID	Description	Workaround
136728	<p>If hundreds of links are bouncing while the FM is sweeping, the FM sweep time may be significantly extended. This can result in unexpected delays in FM responsiveness to fabric changes or host reboots. (The issue is that active links bounce between the time FM discovers one side of the link versus the other side of the link.)</p> <p>In Release 10.3.1 a change was made to improve the FM responsiveness in large fabrics of >1000 nodes when numerous links bounce (or hosts are rebooted) at once.</p>	<p>The following workarounds are recommended:</p> <ul style="list-style-type: none"> When rebooting nodes on a production cluster, perform reboots in batches of 300 nodes or less. During cluster deployment, carefully follow the procedures in the <i>Intel® Omni-Path Fabric Setup Guide</i> and use FastFabric to check signal integrity and placement of all cables. Correct or disable any problematic links before starting production use of the cluster. When replacing or expanding a production cluster, repeat the procedures in the <i>Intel® Omni-Path Fabric Setup Guide</i> to verify the new hardware. Correct or disable any problematic links before resuming production use of the cluster. Use the PM, FM logs, FM GUI, FastFabric, and other tools to monitor signal integrity and link stability. Correct or disable any problematic links when discovered.
136822	The Intel UEFI driver contained in the server BIOS must be executed for proper support of Active Optical Cables (AOC) in an integrated HFI environment. Some BIOS do not execute the UEFI in Legacy BIOS Boot mode, and there are BIOS configuration settings that may prevent the UEFI from executing in any mode.	<p>Avoid the use of Legacy BIOS boot mode if your platform does not execute the HFI driver in that mode.</p> <p>Avoid BIOS settings or other configuration settings that do not execute the HFI driver during boot.</p>
137106	When running SLES* 12.2 with inbox OPA drivers installed, the state may not change from "Offline" to "Physical Linkup (Init)" as expected.	Add a <code>platform.dat</code> file in <code>/lib/firmware/updates</code> , then restart.
137212	The RHEL* 6.7 base version of the <code>perftest</code> package includes a <code>ib_send_lat</code> utility that may cause a segmentation fault when run with the <code>-z</code> option.	Run the utility without using the <code>-z</code> option. Alternatively, install and run <code>perftest-3.0</code> .
137409	When using DHCP from an Intel OPA HFI, the DHCP client-identifier field (option 61) must be used to set up DHCP static leases. However, what is sent in the DHCP discover packet is different for PXE boot attempts versus a DHCP boot attempt.	The DHCP client-identifier can be explicitly specified in DHCP client configuration or NetworkManager scripts.
137951	In the HFI BIOS screen for Advanced NIC Configuration, a warning message about incorrect custom <code>P_Key</code> value is not completely displayed.	The valid range for custom <code>P_Key</code> value is <code>0x8001</code> to <code>0xFFFFE</code> .
138909	Installation fails due to a build error in <code>IntelOPA-Tests</code> , which is caused by a version mismatch between the ICC libraries in the IFS package and the ICC libraries available in the system.	When installing IFS software and using the pre-built MPIS for ICC, the ICC runtime version installed must match the version of ICC used to compile the MPIS in IFS. See Section 1.11, "MPI Libraries" on page 8 for version details.
139397	IPoIB traffic stalls during reboot testing.	None.
139550	Infrequently, an AOC may exhibit an unexpectedly high local link integrity error rate after the link comes up, relative to the error rate on previous link up occasions. This can be determined by observing a link quality of <5. These links may eventually experience a link width downgrade.	Bounce the link.
139613	The Subsystem Vendor and Subsystem Device ID in the PCI configuration space of Intel® Omni-Path discrete HFI cards may not indicate the correct OEM vendor and device. As a result, the <code>lspci</code> command may show incorrect Subsystem Vendor and Device ID information. This issue affects Intel server boards for Intel® Xeon® Processor v3 and v4 Product Family configured in Legacy OS boot mode.	Reconfigure the system from Legacy OS boot mode to UEFI boot mode.
139650	In rare cases, AOCs with mismatched firmware on each end of the link may experience longer than expected link-up times.	If link-up times are unacceptably long, ensure both ends of links are running the same firmware version.



Table 2-3. Open Issues (Sheet 5 of 6)

ID	Description	Workaround
139660	Following a boot it is possible, although rare, that the IpoB interface will fail to come up. Hosts attempting to ping this host will get no response.	Reboot or reload the driver.
139692	On SLES* 12 SP2, when installing the Release 10.5 software with this command: rpm -i hfdiags-0.8-66.x86_64.rpm The following error is returned: error: Failed dependencies: libreadline.so.7()(64bit) is needed by hfdiags-0.8-66.x86_64	Use this command to install: rpm -i hfdiags-0.8-81.x86_64.rpm --nodeps
139743	While running uperf test, timeouts on ib0 Transmit Queue may be seen.	None.
139834	When using the FastFabric TUI to run "Perform Single Host Verification", the test hangs during operation.	Run the hostverify.sh script manually using the opaverifyhosts command.
139981	When Accelerated RDMA (TID RDMA) is enabled, GPFS file system can hang.	None.
139995	When installing Intel® Omni-Path Software on a clean system (with no prior versions installed), the following error message may be displayed: cat: //etc/opa/version_delta: No such file or directory	This message can be safely ignored. The installer is looking for an IFS version file before it has been created on the system. The installation is not impacted.
140073	Error message when running MPI tests: Unable to initialize PSM2 CUDA support In Release 10.5, if IFS is installed with CUDA* support, then the NVIDIA* CUDA* Toolkit is also required to be installed in order to run any MPI application using PSM2 transport. Release 10.5 IFS installations without CUDA* support do not require NVIDIA* CUDA* Toolkit to be installed.	Download and install the NVIDIA* toolkit here: https://developer.nvidia.com/cuda-downloads or Install the IFS without NVIDIA* CUDA* support.
140151	On RHEL* 6.7, during IFS 10.5 installation, the following warning messages are displayed: /etc/sysconfig/irqbalance: line 29: --hintpolicy=exact: command not found Usage: /etc/init.d/irqbalance {start stop status restart reload condrestart force-reload} /etc/sysconfig/irqbalance: line 29: --hintpolicy=exact: command not found	After installing the software, perform the following: 1. Edit /etc/sysconfig/irqbalance and change the line from: IRQBALANCE_ARGS= --hintpolicy=exact to: IRQBALANCE_ARGS=--hintpolicy=exact (remove space after =) 2. Run the command: chkconfig --level 2345 irqbalance on This command enables the irqbalance service to run automatically when rebooted. 3. Run the command: service irqbalance restart and make sure there is no error.



Table 2-3. Open Issues (Sheet 6 of 6)

ID	Description	Workaround
140208	<p>The node description of a node may change after rebooting. This issue has been seen on RHEL* and SLES* distributions and requires restart of the rdma-ndd service.</p>	<p>Complete one of the following workarounds:</p> <ul style="list-style-type: none"> Using the TUI, set the rdma-ndd service to Autostart after rebooting. When the Intel OPA Autostart Menu displays, review the items. Change the option for RDMA NDD from the default value of [Disable] to [Enable]. To do this, type 3. Run the OPA Autostart operations by typing P. Manually enable the rdma-ndd daemon with the command: # <code>systemctl start rdma-ndd</code>
140281	<p>On SLES* 12 SP2, during IFS 10.5 installation, the following warning message is displayed: <code>IRQBALANCE_ARGS="" --hintpolicy=exact</code></p>	<p>After installing the software, perform the following:</p> <ol style="list-style-type: none"> Edit <code>/etc/sysconfig/irqbalance</code> and change the line from: <code>IRQBALANCE_ARGS= --hintpolicy=exact</code> to: <code>IRQBALANCE_ARGS=--hintpolicy=exact</code> (remove space after =) Run the command: <code>systemctl enable irqbalance</code> This command enables the irqbalance service to run automatically when rebooted. Run the command: <code>systemctl restart irqbalance</code> and make sure there is no error.
140691	<p>When running <code>opaswitchadmin</code> against multiple hosts simultaneously, it will send schedule requests in parallel to those hosts, and it's possible that some hosts may intermittently fail due to timeouts at high levels of parallelism.</p>	<p>The number of hosts that we query in parallel can be limited by setting the <code>FF_MAX_PARALLEL</code> value in <code>/etc/opa/opafastfabric.conf</code>. Results may vary by fabric. Intel recommends reducing this to 2.</p>



3.0 Related Information

3.1 Documentation

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

- Intel® Omni-Path Switches Installation, User, and Reference Guides
www.intel.com/omnipath/SwitchPublications
- Intel® Omni-Path Fabric Software Installation, User, and Reference Guides
www.intel.com/omnipath/FabricSoftwarePublications
- Drivers and Software (including Release Notes)
www.intel.com/omnipath/downloads

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

Table 3-1. Intel® Omni-Path Documentation Library (Sheet 1 of 3)

Task	Document Title	Description
Key: Shading indicates the URL to use for accessing the particular document.		
• Intel® Omni-Path Switches Installation, User, and Reference Guides: http://www.intel.com/omnipath/SwitchPublications		
• Intel® Omni-Path Software Installation, User, and Reference Guides (includes HFI documents): http://www.intel.com/omnipath/FabricSoftwarePublications		
• Drivers and Software (including Release Notes): http://www.intel.com/omnipath/Downloads		
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. It outlines the most basic steps for getting your Intel® Omni-Path Architecture (Intel® OPA) cluster installed and operational.
Setting up an Intel® OPA cluster	New title: <i>Intel® Omni-Path Fabric Setup Guide</i> (Old title: <i>Intel® Omni-Path Fabric Staging Guide</i>)	Provides a high level overview of the steps required to stage a customer-based installation of the Intel® Omni-Path Fabric. Procedures and key reference documents, such as Intel® Omni-Path user guides and installation guides are provided to clarify the process. Additional commands and BKM's are defined to facilitate the installation process and troubleshooting.
Installing hardware	<i>Intel® Omni-Path Fabric Switches Hardware Installation Guide</i>	Describes the hardware installation and initial configuration tasks for the Intel® Omni-Path Switches 100 Series. This includes: Intel® Omni-Path Edge Switches 100 Series, 24 and 48-port configurable Edge switches, and Intel® Omni-Path Director Class Switches 100 Series.
	<i>Intel® Omni-Path Host Fabric Interface Installation Guide</i>	Contains instructions for installing the HFI in an Intel® OPA cluster. A cluster is defined as a collection of nodes, each attached to a fabric through the Intel interconnect. The Intel® HFI utilizes Intel® Omni-Path switches and cabling.



Table 3-1. Intel® Omni-Path Documentation Library (Sheet 2 of 3)

Task	Document Title	Description
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 User Interface (TUI) to guide you through the installation process. You have the option of using command line interface (CLI) commands to perform the installation or install using the Linux* distribution software.
Managing a switch using Chassis Viewer GUI Installing switch firmware (managed switches)	<i>Intel® Omni-Path Fabric Switches GUI User Guide</i>	Describes the Intel® Omni-Path Fabric Chassis Viewer graphical user interface (GUI). It provides task-oriented procedures for configuring and managing the Intel® Omni-Path Switch family. Help: GUI Online Help.
Managing a switch using the CLI Installing switch firmware (managed switches)	<i>Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide</i>	Describes the command line interface (CLI) task information for the Intel® Omni-Path Switch family. Help: -help for each CLI.
Managing a fabric using FastFabric	<i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i>	Provides instructions for using the set of fabric management tools designed to simplify and optimize common fabric management tasks. The management tools consist of TUI menus and command line interface (CLI) commands.
	<i>Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide</i>	Describes the command line interface (CLI) for the Intel® Omni-Path Fabric Suite FastFabric. Help: -help and man pages for each CLI. Also, all host CLI commands can be accessed as console help in the Fabric Manager GUI.
Managing a fabric using Fabric Manager	<i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i>	The Fabric Manager uses a well defined management protocol to communicate with management agents in every Intel® Omni-Path Host Fabric Interface (HFI) and switch. Through these interfaces the Fabric Manager is able to discover, configure, and monitor the fabric.
	<i>Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide</i>	Provides an intuitive, scalable dashboard and set of analysis tools for graphically monitoring fabric status and configuration. It is a user-friendly alternative to traditional command-line tools for day-to-day monitoring of fabric health. Help: Fabric Manager GUI Online Help.
Configuring and administering Intel® HFI and IPoIB driver 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 both cluster administrators and Message-Passing Interface (MPI) application programmers, who have different but overlapping interests in the details of the technology.
Writing and running middleware that uses Intel® OPA	<i>Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide</i>	Provides a reference for programmers working with the Intel® PSM2 Application Programming Interface (API). The Performance Scaled Messaging 2 API (PSM2 API) is a low-level user-level communications interface.
Optimizing system performance	<i>Intel® Omni-Path Fabric Performance Tuning User Guide</i>	Describes BIOS settings and parameters that have been shown to ensure best performance, or make performance more consistent, on Intel® Omni-Path Architecture. If you are interested in benchmarking the performance of your system, these tips may help you obtain better performance.
Designing an IP or storage router on Intel® OPA	<i>Intel® Omni-Path IP and Storage Router Design Guide</i>	Describes how to install, configure, and administer an IPoIB router solution (Linux* IP or LNet) for inter-operating between Intel® Omni-Path and a legacy InfiniBand* fabric.
Building a Lustre* Server using Intel® OPA	<i>Building Lustre* Servers with Intel® Omni-Path Architecture Application Note</i>	Describes the steps to build and test a Lustre* system (MGS, MDT, MDS, OSS, OST, client) from the HPDD master branch on a x86_64, RHEL*/CentOS* 7.1 machine.



Table 3-1. Intel® Omni-Path Documentation Library (Sheet 3 of 3)

Task	Document Title	Description
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 Architecture Management API Programmer's Guide</i>	Contains a reference for programmers working with the Intel® Omni-Path Architecture Management (Intel OPAMGT) Application Programming Interface (API). The Intel OPAMGT API is a C-API permitting in-band and out-of-band queries of the FM's Subnet Administrator and Performance Administrator.
Learning about new release features, open issues, and resolved issues for a particular release	<i>Intel® Omni-Path Fabric Software Release Notes</i>	
	<i>Intel® Omni-Path Fabric Manager GUI Release Notes</i>	
	<i>Intel® Omni-Path Fabric Switches Release Notes</i> (includes managed and externally-managed switches)	