

Intel[®] Omni-Path Fabric Software

Release Notes for 10.1.1

June 2016



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

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

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

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

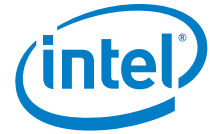
Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or by visiting: <http://www.intel.com/design/literature.htm>

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, 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 © 2015-2016, Intel Corporation. All rights reserved.



Contents

1.0 Overview of the Release	4
1.1 Introduction	4
1.2 Audience	4
1.3 Software License Agreement	4
1.4 If You Need Help	4
1.5 New Features in this Release	4
1.6 Supported Features	4
1.7 Release Packages	5
1.8 Firmware Files	6
1.9 Operating Systems	6
1.10 Parallel File System	6
1.11 Compilers	7
1.11.1 MPI	7
1.11.2 MVAPICH2 and Open MPI	7
1.12 Hardware	8
1.13 Installation Requirements	8
1.13.1 Software and Firmware Requirements	8
1.13.2 Installation Instructions	8
1.14 Product Constraints	9
1.15 Product Limitations	9
1.16 Documentation	9
2.0 Issues	11
2.1 Introduction	11
2.2 Resolved Issues	11
2.3 Open Issues	12

Tables

1-1 Firmware Files	6
1-2 Operating Systems Supported	6
1-3 MPI Compilers	7
1-4 MVAPICH2 and Open MPI	7
1-5 Hardware Supported	8
1-6 Related Documentation for this Release	9
2-1 Resolved Issues	11
2-2 Open Issues	12



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 www.intel.com for additional detail.

1.5 New Features in this Release

- The following features are new for the 10.1.1 release: Additional operating systems supported. See [Table 1-2](#) for details
- Additional hardware supported. See [Table 1-5](#) for details
- Boot over Intel® Omni-Path using secure UEFI. This replaces the previously released Boot over Fabric.

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](#).
- Coexistence with Intel® True Scale Architecture. This release supports True Scale hardware serving as an InfiniBand* storage network with the Intel® Omni-Path hardware used for computing. Note that connecting a True Scale adapter card to an



Omni-Path switch, or vice-versa, is not supported. For more details on this feature, refer to *Intel® Omni-Path Fabric Host Software User Guide*.

- Supports Dual Rail: Two Intel® Omni-Path Host Fabric Interface (HFI) cards in the same server connected to the same fabric.
- Supports Dual Plane: Two HFI cards in the same server connected to separate fabrics.
- Limited validation testing performed on network storage file systems:
 - NFS over TCP/IP
- Active Optical Cables
- MPI applications are provided in a stand-alone rpm package.
- Intel® Xeon® v4 processor (codename Broadwell) support
- KNL support (B0 technology preview)
- Monitored Intel® Omni-Path Host Fabric Interface
- DHCP and LDAP supported on Intel® Omni-Path Edge Switch 100 Series and Intel® Omni-Path Director Class Switch 100 Series hardware.

1.7 Release Packages

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

- Basic for compute nodes
- IFS for the management node

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-ddiagtools-sw, libhfi1verbs
 - Open MPI built for verbs and PSM2 using gcc, Intel, and PGI compilers.
 - MVAPICH2 built for verbs and PSM2 using gcc, Intel, and PGI compilers.
 - 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 *Intel® Omni-Path Fabric Software Installation Guide* 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.

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 in [Table 1-6](#).

1.8 Firmware Files

This release of the Intel® Omni-Path Software contains the firmware files listed in [Table 1-1](#).

Table 1-1. Firmware Files

Description	File Name	Version
HF11 UEFI Option ROM	HfiPcieGen3_0x1c.efi	0x1c
UEFI UNDI	HfiUndiStub_0x1c.rom	0x1c
HF11 SMBus Microcontroller Firmware (Thermal Monitor)	hfi1_smbus.fw	10.1.0.0.101

1.9 Operating Systems

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

Table 1-2. Operating Systems Supported

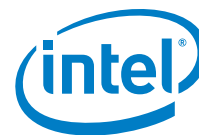
Operating System	Update/SP	Kernel Version
Red Hat* Enterprise Linux* (RHEL*) 7.1 X86_64	Update 1	3.10.0-229.el7.x86_64
Red Hat* Enterprise Linux* (RHEL*) 7.2 X86_64	Update 2	3.10.0-327.el7.x86_64
CentOS* 7.1 X86_64	N/A	3.10.0-229.el7.x86_64
CentOS* 7.2 X86_64	N/A	3.10.0-327.el7.x86_64
Scientific Linux* 7.1 X86_64	N/A	3.10.0-229.el7.x86_64
Scientific Linux* 7.2 X86_64	N/A	3.10.0-327.el7.x86_64
SUSE* Linux* Enterprise Server (SLES*) 12 X86_64	N/A	3.12.28-4-default
SUSE* Linux* Enterprise Server (SLES*) 12.1 X86_64	Service Pack 1	3.12.49-11.1-default

1.10 Parallel File System

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

- IBM* General Parallel File System (GPFS) version 4.0.1
 - RHEL* 7.1 and 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 Compilers

1.11.1 MPI

This release supports the following MPI implementations:

Table 1-3. MPI Compilers

MPI Implementation	Runs Over	Compiled With
Open MPI 1.10.2	Verbs	GCC
	PSM2	GCC, Intel
MVAPICH2-2.1	Verbs	GCC
	PSM2	GCC, Intel
Intel® MPI 5.1.3	Verbs	GCC
	PSM2	GCC, Intel

1.11.2 MVAPICH2 and Open MPI

MVAPICH2 and Open MPI have been compiled for PSM2 to support the following versions of the compilers:

Table 1-4. MVAPICH2 and Open MPI

Compiler	Linux* Distribution	Compiler Version
(GNU) gcc	RHEL* 7	gcc (GCC) 4.8.2 20140120
(GNU) gcc	RHEL* 7.1	gcc (GCC) 4.8.3 20140911
(GNU) gcc	RHEL* 7.2	gcc (GCC) 4.8.5 20150623 (Red Hat* 4.8.5-4)
(GNU) gcc	SLES* 12	gcc (SUSE* Linux*) 4.8.3 20140627
(GNU) gcc	SLES* 12 SP 1	gcc (SUSE* Linux*) version 4.8.5
(Intel) icc	RHEL* 7	icc (ICC) 15.0.1
(Intel) icc	RHEL* 7.1	icc (ICC) 15.0.1
(Intel) icc	RHEL* 7.2	icc (ICC) 15.0.1
(Intel) icc	SLES* 12	icc (ICC) 15.0.1
(Intel) icc	SLES* 12 SP 1	icc (ICC) 15.0.1

See ID [132345](#) for expected behavior that may be observed when using Open MPI with the SLURM PMI launcher and PSM2.



1.12 Hardware

Table 1-5 lists the hardware supported in this release.

Table 1-5. Hardware Supported

Hardware	Description
Intel® Xeon® v3 Processor	Haswell CPU-based servers
Intel® Xeon® v4 Processor	Broadwell 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 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

1.13 Installation Requirements

1.13.1 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.

1.13.2 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 10.1.1 release.

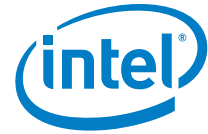
Refer to the *Intel® Omni-Path Fabric Software Installation Guide* for related software requirements and complete installation procedures. Refer to the *Intel® Omni-Path Fabric Hardware Installation Guide* for related firmware requirements.

1.13.2.1 Disable SRP Daemon Autostart

By default, RHEL* 7.1 is configured to autostart the SRP daemon on every node. This daemon performs a fabric search for SRP target devices every minute, which may impact the performance of HPC applications, especially in larger fabrics.

Intel recommends you disable the SRP daemon on all compute nodes and any service nodes. Typically, the SRP daemon only needs to be enabled on filesystem server nodes that are directly accessing block storage devices.

When installing either the IntelOPA-Basic or IntelOPA-IFS packages, there is a prompt: Enable SRP initiator autoload? Intel recommends you use the default option no on all compute nodes and any service nodes.



When installing using other means (such as provisioning systems or direct installation of the rpms), you must manually disable the SRP daemon autostart. After installing the `rdma-*noarch.rpm` package, edit the `/etc/rdma/rdma.conf` file. The default setting is: `SRP_LOAD=yes`. Intel recommends you change this to `no`.

1.14 Product Constraints

None.

1.15 Product Limitations

This release has the following product limitations:

- The UNDI 0x1c firmware resident in this release is not compatible in server configurations that include dual HFI PCIe cards.
- The embedded version of the Fabric Manager supports a maximum of 100 nodes within a fabric. This is due to the limited memory and processing resources available in the embedded environment.
- PXE boot can be accomplished using UEFI BIOS. PXE boot is not supported in legacy boot mode. Refer to the Intel® Omni-Path Fabric Software Installation Guide, “HFI UEFI PXE Installation and Configuration” Appendix for more information.
- PA Failover should **not** be enabled with FMs running on differing software versions. PA Failover is enabled via `configuration:<PM>/<ImageUpdateInterval> > 0`

1.16 Documentation

Table 1-6 lists the end user documentation for the current release.

Documents are available at the following URLs:

- Intel® Omni-Path Switches Installation, User, and Reference Guides
<http://www.intel.com/SUPPORT/NETWORK/OMNI-SWTCH100/SB/CS-035856.HTM>
- Intel® Omni-Path Host Fabric Interface Installation, User, and Reference Guides (includes software documents)
<http://www.intel.com/support/network/omni-adptr/sb/CS-035857.htm>
- Drivers and Software (including Release Notes)
<https://downloadcenter.intel.com/>

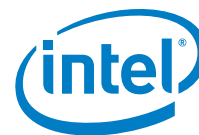
Table 1-6. Related Documentation for this Release (Sheet 1 of 2)

Document Title
Hardware Documents
<i>Intel® Omni-Path Fabric Switches Hardware Installation Guide</i>
<i>Intel® Omni-Path Fabric Switches GUI User Guide</i>
<i>Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide</i>
<i>Intel® Omni-Path Edge Switch Platform Configuration Reference Guide</i>
<i>Intel® Omni-Path Fabric Managed Switches Release Notes</i>
<i>Intel® Omni-Path Fabric Externally-Managed Switches Release Notes</i>



Table 1-6. Related Documentation for this Release (Sheet 2 of 2)

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



2.0 Issues

2.1 Introduction

This section provides a list of 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. Resolved Issues

ID	Component	Description	Resolved in Release
133145	HFI Host Driver	hfi1_eprom has options to write, read, and erase information from fabric EPROM. However, it does not have an option to display current versions of the OptionROM version, Bootcode version, and Configuration file version. It is important for users to have this option to verify successful update.	10.1.1
133241	HFI Host Driver	Under some circumstances, an HFI adapter returns QSFP memory read errors and is unable to bring up the physical link. (AOC link does not come up.)	10.1.1
134012	HFI Host Pre-boot Software	The UEFI driver may not work on systems with the HFI card plugged into a PCIe slot that is not directly connected to the PCIe root complex.	10.1.1
134039	Software Man Pages	Man pages contain extra Unicode and other characters that do not belong in the text.	10.1.1



2.3 Open Issues

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

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

ID	Component	Description	Workaround
129563	HFI Host Fabric Software/ MPI	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 /etc/security/limits.conf: * hard memlock unlimited * soft memlock unlimited Also, you must increase the lkey_table_size: 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 Intel® Omni-Path Fabric Performance Tuning User Guide.
130233	HFI Host Pre-boot Software	Intel® Omni-Path DHCP fails to ID system based on hardware ID.	Workaround not available.
130336	HFI Host Fabric Software/ Tools	hfi1stats cannot be run at user level due to mount-point privileges	The administrator can provide "sudo access" to "hfi1stats" or provide root access to users.
130435	HFI Host Driver	Packet traffic stops after reconnecting a Fabric Manager HFI port to a management allowed port that was previously marked "management not allowed".	Reboot the Fabric Manager node.
131745	HFI Host Fabric Software/ MPI	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.	Stopping and resuming the "hung" orted process allows the job to finish normally. To find the hung process, run the ps and find a node with several job zombie processes. In that same node, identify the orted process ID and send a stop signal (kill -19 <PID>) and a continue signal (kill -18 <PID>).
132207	Open Software	Kernel crash caused by the ib_srpt module.	Install this kernel patch: https://git.kernel.org/cgit/linux/kernel/git/torvalds/linux.git/commit/?id=51093254bf879bc9ce96590400a87897c7498463
132345	Host Fabric Software/ MPI	When using Open MPI with the SLURM PMI launcher and PSM2, you may see the following error: Error obtaining unique transport key from ORTE (orte_precondition_transports not present in the environment). This is expected behavior. For details, see the Intel® PSM2 Programmer's Guide, in the description of psm2_ep_open() function, unique_job_key parameter.	To resolve this error, set the Open MPI environment variable OMPI_MCA_orte_precondition_transports in every node running the job. The format is 16 digit hexadecimal characters separated by a dash. For example, OMPI_MCA_orte_precondition_transports=13241234acffedeb-abcdefabcdef1233 This key is used by the PSM2 library to uniquely identify each different job end point used on the fabric. If two MPI jobs are running on the same node sharing the same HFI and using PSM2, each one should have a different key.
132714	Host Fabric Software/MPI	The latest mvapich2 requires autoconf 2.67/automake 1.15/libtool 2.4.3 or newer to rebuild.	To rebuild the package from srpm, download and install the latest autoconf/automake/libtool onto the build machine.
132718	Software Installation	The Intel® Omni-Path SLES 12.1 package contains infiniband-diags version 1.6.4 which does not include the rdma-ndd service.	Install infiniband-diags version 1.6.5. Make sure the rdma-init-kernel is run before rdma-ndd.

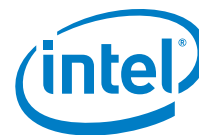


Table 2-2. Open Issues (Sheet 2 of 3)

ID	Component	Description	Workaround
133038	HFI Host Driver	PSM fails to acquire a context despite the kernel having contexts available. The following error message is returned: assign_context command failed: Device or resource busy	Reload hfi1 driver or reboot node.
133269	Host Software/Verbs	GFP_NOIO error message using GFP_KERNEL when using IPoIB on OPA in connected mode.	Workaround not available.
133377	HFI Host Driver	irqbalance settings are not being honored correctly after a reboot.	Restart irqbalance after the hfi1 driver has loaded.
133408	HFI Hardware/PCIe card	In rare scenarios, loading the hfi1 host driver causes syslog messages including text such as "kernel: Uhhuh. NMI received for unknown reason 39 on CPU 0." On some systems, this may also be reported as "PCIe Fat" or as "CPU ERR2".	The error reporting is a transitory condition that arises on driver load. You can ignore the error and continue. Alternatively, ensure that the UEFI driver is being executed. This will perform the initialization before the OS boots and will avoid this error reporting. (For details, see the Intel® Omni-Path Fabric Host Software User Guide.)
133607	Fabric Management Tools/FastFabric	opaverifyhosts fails to run on RHEL7.2	The user can modify the /opt/opa/samples/hostverify.sh before running the pcicfg, pcispeed and hfi_pkt tests. In line 263 where it says: pci_id=\$(grep -o -m1 "24f[01]" pcicfg.stdout) Change grep to egrep. Same for lines 311 and 498.
133676	Fabric Management Software/Fabric Manager	Creating a virtual fabric with standby enabled causes opapaqueries to fail.	Do not enable standby on virtual fabrics.
133709	HFI Host Fabric Software/Tools	Transmit Discard (xmit discard) counts are captured at the port level through the opapmaquery command. opapmaquery additionally shows per VL counts for xmit discard (based on tool options), but currently shows zero counts.	The VL level counts can be found using the hfi1stats command.
133820	HFI Host Pre-boot Software	The UEFI driver may not work on some systems. The system freezes when the UEFI driver starts.	There are no known workarounds.
133918	HFI Host Pre-boot Software	Booting over Omni-Path HFI occasionally fails with "buffer too small, need netbuf_size = 16777216".	Retry boot
133987	Fabric Management Software/Fabric Manager	SM crash with general protection fault during compounding of PM sweep image	Set ImagesPerComposite to 1 in opafm.xml
134118	HFI Host Fabric Software/PSM2	PSM2_MQ_RNDV_HFI_WINDOW stalls HPL on bin1 KNL nodes	Some large memory multi-node MPI applications running on a cluster of pre-release Intel® Xeon Phi™ CPU 7200 nodes with OPA may require reducing the PSM2_MQ_RNDV_HFI_WINDOW environment variable from the default of 4194304 to a lower value, such as 262144 or 393216. Please only adjust this parameter if you see that your workload is hanging or stalling.



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

ID	Component	Description	Workaround
134124	Fabric Management Software/SM	HFI port stuck in INIT state due to SM failure to set pkeys.	Bounce the link
134131	HFI Host Pre-boot Software	Servers with dual HFI PCIe cards become un-bootable after flashing the UNDI 0x1c firmware to one or both of the HFI PCIe cards.	Do not use the UNDI 0x1c firmware on servers with dual HFI PCIe cards. Some servers function in non-legacy mode in this configuration.
134135	HFI Host Fabric Software/ HFI Host Driver	When running communication-intensive workloads with 10KB MTU, it is possible to encounter node and/or job failures.	Change MTU default size in configuration file from 10KB MTU to 8KB MTU (hfi1 driver module parameter max_mtu=8192). Refer to Section A.3, "Setting HFI1 Driver Parameters" in the Intel® Omni-Path Performance Tuning User Guide.

§ §