Intel® Omni-Path Software

Release Notes for V10.10.1

Rev. 1.0

January 2020

Doc. No.: K82053, Rev.: 1.0

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.

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

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

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

Copyright © 2020, Intel Corporation. All rights reserved.

Contents

1.0 Overview of the Release..... 5
1.1 Audience..... 5
1.2 Software License Agreement..... 5
1.3 If You Need Help..... 5
1.4 Updated Features..... 5
1.5 Product Improvements..... 6
1.6 Deprecated and Removed Features..... 8
1.7 Release Packages..... 8
1.8 Release Compatibility..... 9
1.9 Operating Systems..... 9
1.10 CUDA Support..... 9
1.11 Kernel Modules for hfi1 Driver..... 10
1.12 Parallel File Systems..... 11
1.13 MPI Libraries..... 11
1.13.1 Supported MPI Libraries..... 11
1.13.2 Compiler Versions and Distributions..... 12
1.14 Intel Hardware..... 12
1.15 Switch Firmware..... 13
1.16 Document Versions..... 13
1.17 Installation Requirements..... 14
1.17.1 Best Practices..... 14
1.17.2 Software and Firmware Requirements..... 14
1.17.3 OS RPMs Installation Prerequisites..... 14
1.17.4 Installation Instructions..... 20
1.17.5 Installation Path Changes in Release 10.4 (and later)..... 20
1.18 Product Constraints..... 21
1.19 Product Limitations..... 22
1.20 Accelerated RDMA Information..... 22

2.0 Issues..... 23
2.1 Resolved Issues..... 23
2.1.1 Issues Resolved in this Release..... 23
2.1.2 Issues Resolved in Prior Releases..... 24
2.2 Open Issues..... 26
2.2.1 Third Party Open Issues..... 28

3.0 Related Information..... 30
3.1 Intel® Omni-Path Documentation Library..... 30
3.1.1 How to Search the Intel® Omni-Path Documentation Set..... 32

Tables

1 Supported Operating Systems....9
2 CUDA Drivers Tested....10
3 Kernel Modules for hfi1 Driver....10
4 Supported MPI Libraries....12
5 Compiler Versions and Distributions....12
6 Supported Hardware....13
7 Supported Document Versions....13
8 RHEL* 7.6 Distribution RPMs....15
9 RHEL* 7.7 Distribution RPMs....15
10 RHEL* 8 Distribution RPMs....16
11 RHEL* 8.1 Distribution RPMs....17
12 SLES* 12.4 Distribution RPMs....17
13 SLES* 12.5 Distribution RPMs....18
14 SLES* 15 Distribution RPMs....18
15 SLES* 15.1 Distribution RPMs....19
16 Issues Resolved in this Release....23
17 Issues Resolved in Prior Releases....24
18 Open Issues....26
19 Third Party Open Issues....28

1.0 Overview of the Release

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.

1.1 Audience

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

1.2 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.

1.3 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.

1.4 Updated Features

The following table lists the updated features included in this release.

Feature	First Introduced
Intel® Parallel Studio XE 2019, Update 5	10.10.1
Intel® C/C++ Compiler 19.5	10.10.1
libfabric OFI version 1.7.1	10.10.0
NVIDIA* CUDA* 9.2 and 10.1	10.10.0
Open MPI version 3.1.4	10.10.0
Accelerated IPoFabric (AIP) is available in all supported OS releases. *Note:* **Accelerated IPoFabric (AIP) works only with Datagram Mode, is on by default, and requires additional tuning.** See the *Intel® Omni-Path IP and LNet Router Design Guide* and *Intel® Omni-Path Fabric Performance Tuning User Guide* for more details.	10.10.0

1.5 Product Improvements

The following improvements are included in this release:

- (Introduced in 10.10.1) Installation Enhancements
  - Meta-packages are available to facilitate yum/zypper-based IFS installation. Each meta-package corresponds to one component in the IFS INSTALL script. Install, upgrade, or downgrade using a meta-package will have the same effect as using the INSTALL script.
  - Alias packages are introduced to further facilitate IFS installation. Each alias package represents a typical HPC node (management node, compute node, and service node) IFS installation. Install using an alias package will install a full set of rpms for the corresponding HPC node.
- (Introduced in 10.10.0) Top 10 Class Storage Performance Tuning

  Intel® Omni-Path drivers have been further hardened and tuned to meet the needs of high-end HPC filesystems. Refer to *Intel® Omni-Path Fabric Performance Tuning User Guide*.
- (Introduced in 10.10.0) Performance Administration (PA) History Scalability

  The performance of the PA when monitoring large clusters and/or clusters with longer retention of performance history has been improved. This benefit can be seen when using `opareport` (especially with `--begin` and `--end` options), `opatop`, the FM GUI, the OPA Management API, and many other tools. In addition, the following PA history analysis capabilities have been enhanced:
  - `opatop` now permits direct access to historic PM images by inputting a specific time. When a long history is retained, this capability is more effective than simply stepping forward or backward among the images. Refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide*, Navigating PM Sweeps section.
  - FastFabric now includes a more effective way to extract all of the counters for the fabric and report the specific link and neighbor associated with them.

    The new `opaextractperf2` script permits easy analysis of information, in a spreadsheet or through other customer scripts, for the fabric. `opareport` focus options and `--begin` and `--end` options may be used to obtain the subset of needed information.

    The new `opamergeperf2` script combines two previous `opaextractperf2` runs with running counters into a csv file showing the delta between the two points in time.

    Refer to the respective sections in the *Intel® Omni-Path Fabric Suite FastFabric User Guide*.
- (Introduced in 10.10.0) Installation Enhancements

  The IFS INSTALL script and *Intel® Omni-Path Fabric Software Installation Guide* now handle installation into `chroot` images better to simplify the preparation of server boot images for provisioning systems. This includes the following:
  - OPA Install now automatically updates the boot image through `dracut` even when installing into a `chroot` image that contains a different kernel than the current running system. This can simplify procedures for creating and maintaining boot images in provisioning systems.
  - The INSTALL script now allows command line control over enabling the `arptable` tuning function.

- (Introduced in 10.10.0) Improved Fabric Debug/Diagnostics
  - The FM will now log the *LinkDownReason* in Node Disappearance log messages.
  - `opareport` now includes the link down reason logs in `-o comps -d5` reports and snapshots. `opareport` also allows focus on link down reason through the `-F ldr:` option.
  - `opareport` now includes the *LinkCRCMode* in `-o comps`.
  - `opareport -o links` now includes port counters.
  - PM logging for fabric error events is now improved to include the error information provided by the hardware. On security, routing, and some signal integrity error events, the PM will get and clear hardware *ErrorInfo* and log the information. FM configuration allows a threshold for the number of such log events per sweep. This log can provide additional information such as which `PKey` was attempting to be used or which `DLID` was unable to be routed. This information can aid sysadmins to debug faulty applications or configurations.
  - The `opaextractbadlinks`, `opaextractlids`, and `opaextractmissinglinks` tools have been enhanced to permit use of `opareport -F` focus options to limit the scope of the report

  Refer to the *Intel® Omni-Path Fabric Suite FastFabric User Guide* for more information.
- (Introduced in 10.10.0) The *Intel® Omni-Path Management API Programmer's Guide* documents the performance characteristics of PA query functions. This can permit developers of PA applications to design their PA usage for improved scalability and performance.
- (Introduced in 10.10.0) The Intel® Omni-Path Fabric Manager can now be more tolerant of non-responsive SMAs on hosts whose port remains active. This can permit applications to soak the CPU or memory of a server without the host being removed from the fabric by the FM. This is controlled via the `NonRespMaxCount` FM configuration option. Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide* for more information.
- (Introduced in 10.10.0) A new `PSM2_AVX512` environment variable has been added that allows the user to control whether PSM2 uses AVX512 code for performing larger memory copies. Disabling AVX512 in PSM2 may improve performance for some applications and traffic patterns. Refer to the *Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide*.
- (Introduced in 10.10.0) Information for customizing and rebuilding MPI Libraries and MPI Applications has been added to documentation. Refer to the *Intel® Omni-Path Fabric Setup Guide*, Customizing MPI section.
- (Introduced in 10.10.0) Prerequisite information has been added to Router Redundancy/Failover with VRRP v3. Refer to the *Intel® Omni-Path IP and LNet Router Design Guide*, Router Redundancy/Failover with VRRP v3 Prerequisites section.
- (Introduced in 10.10.0) New troubleshooting information for diagnosing issues with cables has been added to documentation. Refer to the *Intel® Omni-Path Fabric Suite Fabric Manager User Guide*, Diagnosing Cable Issues appendix.
- (Introduced in 10.10.0) Enhanced information for dual/multi-rail support has been added to documentation. Refer to the following publications:

- *Intel® Omni-Path Fabric Software Installation Guide*, Multi-Subnet Fabrics and Multi-Rail Usage sections
- *Intel® Omni-Path Fabric Host Software User Guide*, Multi-Rail Support in PSM2 section
- *Intel® Omni-Path Fabric Performance Tuning User Guide*, Lustre Multi-Rail support with OPA section

1.6 Deprecated and Removed Features

- As of 10.10.1, CongDiscards counters are not longer thresholded in the default configuration file (`/etc/opa/opamon.conf`). As a result, they are no longer reported by `opareport -o errors`, `opaextractstat`, `opaextractstat2`, `opaextractbadlinks`, and `opafabricanalysis` when using the default configuration file.
- Support for SHMEM has been removed in the 10.10.0 release.
- As of 10.10.0, native verbs support in Open MPI (openib BTL) is no longer maintained. It has been removed from the Open MPI build in IFS.

1.7 Release Packages

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

- Basic for compute, service, and login nodes
- IFS for management nodes

The Basic package includes:

- Software that installs the following packages to the distribution OpenFabrics Alliance* (OFA):
  - hfi1-firmware, libpsm2 (for RHEL*) and libpsm2-2 (for SLES*), hfi1-diagtools-sw
  - Open MPI and MVAPICH2. See MPI Libraries on page 11 for details.
  - mpitests
  - mpi-selector
  - Open Fabrics Interface (OFI) libfabric

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.
- 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.

1.8 Release Compatibility

This release is backward compatible with the most recent minor release version. For example, Release 10.N is backward compatible with Release 10.N-1 and Release 10.N-1.x.

1.9 Operating Systems

This release of the Intel® Omni-Path Software supports the operating systems listed in the following table.

Table 1. Supported Operating Systems

Operating System	Update/SP	Base Kernel Version
Red Hat* Enterprise Linux* (RHEL*) 7.6 X86_64	Update 6	3.10.0-957.el7.x86_64
Red Hat* Enterprise Linux* (RHEL*) 7.7 X86_64	Update 7	3.10.0-1062.el7.x86_64
Red Hat* Enterprise Linux* (RHEL*) 8 X86_64		4.18.0-80.el8.x86_64
Red Hat* Enterprise Linux* (RHEL*) 8.1 X86_64	Update 1	4.18.0-147.el8.x86_64
CentOS*-7 (1810) X86_64 (corresponds to RHEL* 7.6)	(1810)	3.10.0-957.el7.x86_64
CentOS*-7 (1908) X86_64 (corresponds to RHEL* 7.7)	(1908)	3.10.0-1062.el7.x86_64
CentOS*-8 (1905) X86_64 (corresponds to RHEL* 8)	(1905)	4.18.0-80.el8.x86_64
Scientific Linux* 7.6 X86_64	Update 6	3.10.0-957.el7.x86_64
Scientific Linux* 7.7 X86_64	Update 7	3.10.0-1062.el7.x86_64
SUSE* Linux* Enterprise Server (SLES*) 12.4 X86_64	Service Pack 4	4.12.14-94.41_default
SUSE* Linux* Enterprise Server (SLES*) 12.5 X86_64	Service Pack 5	4.12.14-120_default
SUSE* Linux* Enterprise Server (SLES*) 15 X86_64		4.12.14-23_default
SUSE* Linux* Enterprise Server (SLES*) 15.1 X86_64	Service Pack 1	4.12.14-195_default

1.10 CUDA Support

PSM2 GPUDirect* RDMA with CUDA is supported as shown below:

- CUDA Toolkit 10.1 is supported on RHEL* 7.6, RHEL* 8, SLES* 12 SP4, and SLES* 15.
- Intel does not officially support CUDA with SLES* 12.5, SLES* 15 SP1, RHEL* 7.7, and RHEL* 8.1 at this time.

This table below reports the CUDA driver versions that have been tested with this release of the Intel® Omni-Path Software. Note that not all combinations of OS, drivers, and Intel® OPA software are tested.

Table 2. CUDA Drivers Tested

OS Distro	CUDA Runtime	Kernel
RHEL* 7.6	CUDA 10.1 (10.1.243)	3.10.0-957.el7.x86_64
RHEL* 8		4.18.0-80.el8.x86_64
SLES* 12 SP4		4.12.14-94.41_default
SLES* 15		4.12.14-23_default

For information on compatible driver versions, refer to NVIDIA*'s CUDA Compatibility, CUDA Toolkit and Compatible Driver Versions table.

1.11 Kernel Modules for hfi1 Driver

This release of the Intel® Omni-Path Software contains the hfi1 driver kernel modules listed in the following table.

NOTE

The version number is shown in bold.

Table 3. Kernel Modules for hfi1 Driver

OS	Non-CUDA/CUDA	RPM Version
RHEL* 7.6	Non-CUDA	kmod-ifs-kernel-updates-3.10.0_957.el7.x86_64-**1917**.x86_64.rpm
	CUDA	kmod-ifs-kernel-updates-3.10.0_957.el7.x86_64-**1917cuda**.x86_64.rpm
RHEL* 7.7	Non-CUDA	kmod-ifs-kernel-updates-3.10.0_1062.el7.x86_64-**1917**.x86_64.rpm
RHEL* 8	Non-CUDA	kmod-ifs-kernel-updates-4.18.0_80.el8.x86_64-**1917**.x86_64.rpm
	CUDA	kmod-ifs-kernel-updates-4.18.0_80.el8.x86_64-**1917cuda**.x86_64.rpm
RHEL* 8.1	Non-CUDA	kmod-ifs-kernel-updates-4.18.0_147.el8.x86_64-**1917**.x86_64.rpm
SLES* 12.4	Non-CUDA	ifs-kernel-updates-kmp-default-4.12.14_94.41_default_k4.12.14_94.41-**1917**.x86_64.rpm
	CUDA	ifs-kernel-updates-kmp-default-4.12.14_94.41_default_k4.12.14_94.41-**1917cuda**.x86_64.rpm
SLES* 12.5	Non-Cuda	ifs-kernel-updates-kmp-default-4.12.14_120_default_k4.12.14_120-**1917**.x86_64.rpm
SLES* 15	Non-CUDA	ifs-kernel-updates-kmp-default-4.12.14_23_default_k4.12.14_23-**1917**.x86_64.rpm
	CUDA	ifs-kernel-updates-kmp-default-4.12.14_23_default_k4.12.14_23-**1917cuda**.x86_64.rpm
SLES* 15.1	Non-CUDA	ifs-kernel-updates-kmp-default-4.12.14_195_default_k4.12.14_195-**1917**.x86_64.rpm

1.12 Parallel File Systems

This section reports the parallel file systems that have been tested with this release of the Intel® Omni-Path Software. Note that not all combinations of OS, file system, and Intel® OPA software are tested.

The following parallel file systems have been tested:

- Lustre* Long Term Support (LTS) release:
  - Version 2.10.8 on SLES* 12 SP4
  - Version 2.12.3 on SLES* 12 SP4 (via LNET Self Test)
- IBM* Spectrum Scale[1]:
  - Version 5.0.4 on SLES* 12 SP4

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

1.13 MPI Libraries

NOTE

The binary and sources for Open MPI and MVAPICH2 are included in the software package. FastFabric provides a simple tool to rebuild these MPIs using the Intel® Compiler (not included in the software package). Refer to *Intel® Omni-Path Fabric Suite FastFabric User Guide*, Rebuilding MPI Library and Tools.

NOTE

Intel recommends that you run the Intel® MPI mpivars.sh script when using Intel® MPI with Intel® OPA and OFI.

If you are not using Intel® MPI with Intel® OPA and OFI, do not run this script.

1.13.1 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 if the MPI library is included in the Intel® Omni-Path Software package.

NOTE

As of 10.10.0, the Open MPI build in IFS no longer includes native verbs support (openib BTL).

1 Formerly known as General Parallel File System (GPFS)

Table 4. Supported MPI Libraries

MPI Implementation	Included in Basic Package?	Runs Over	Compiled With
Open MPI 3.1.4	Yes	PSM2, OFI	GCC, Intel (ICC)
Open MPI 3.1.4-cuda	Yes	PSM2	GCC, ICC
MVAPICH2-2.3B	Yes	PSM2	GCC, ICC
Intel® MPI Library 2019, Update 5	No	PSM2 (via OFI)	N/A
IBM* Spectrum* MPI version 10.1.1.0	No	PSM2	N/A

1.13.2 Compiler Versions and Distributions

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

Table 5. Compiler Versions and Distributions

Compiler	OS Distribution	Compiler Version
(GNU) gcc	RHEL* 7.6	gcc version 4.8.5 20150623 (Red Hat* 4.8.5-36) (GCC)
(GNU) gcc	RHEL* 7.7	gcc version 4.8.5 20150623 (Red Hat 4.8.5-39) (GCC)
(GNU) gcc	RHEL* 8	gcc version 8.2.1 20180905 (Red Hat* 8.2.1-3) (GCC)
(GNU) gcc	RHEL* 8.1	gcc version 8.3.1 20190920 (Red Hat* 8.3.1-0) (GCC)
(GNU) gcc	SLES* 12 SP4	gcc version 4.8.5 20181207 (SUSE* Linux)
(GNU) gcc	SLES* 12 SP5	gcc version 4.8.5 20191104 (SUSE* Linux)
(GNU) gcc	SLES* 15	gcc version 7.3.1 20180323 (SUSE* Linux*)
(GNU) gcc	SLES* 15 SP1	gcc version 7.4.1 20190424 [gcc-7-branch revision 270538] (SUSE* Linux*)

NOTE

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

1.14 Intel Hardware

The following table 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 6. Supported Hardware

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 Processors	Skylake CPU-based servers
2nd Generation Intel® Xeon® Scalable Processors	Cascade Lake CPU-based servers
Intel® Xeon Phi™ 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

1.15 Switch Firmware

The following firmware is supported for Intel® Omni-Path switches:

- Intel® Omni-Path Switch Firmware 10.8.x revision (managed and externally-managed switches)
- Intel® Omni-Path Switch Firmware 10.7.x revision (managed and externally-managed switches)

Refer to the *Intel® Omni-Path Fabric Switches Release Notes* for more information.

1.16 Document Versions

The following table lists the end user document versions supported by this release.

Table 7. Supported Document Versions

Title	Doc. Number	Revision
Intel® Omni-Path Fabric Quick Start Guide	J57479	8.0
Intel® Omni-Path Fabric Setup Guide	J27600	12.0
Intel® Omni-Path Fabric Switches Hardware Installation Guide	H76456	12.0
Intel® Omni-Path Host Fabric Interface Installation Guide	H76466	7.0
Intel® Omni-Path Fabric Software Installation Guide	H76467	16.0
Intel® Omni-Path Fabric Switches GUI User Guide	H76457	12.0

continued...

Title	Doc. Number	Revision
Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide	H76458	12.0
Intel® Omni-Path Fabric Suite FastFabric User Guide	H76469	16.0
Intel® Omni-Path Fabric Suite Fabric Manager User Guide	H76468	15.0
Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide	H76471	15.0
Intel® Omni-Path Fabric Host Software User Guide	H76470	16.0
Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide	H76473	14.0
Intel® Omni-Path Fabric Performance Tuning User Guide	H93143	18.0
Intel® Omni-Path IP and LNet Router Design Guide (Old title: *Intel® Omni-Path IP and Storage Router Design Guide*)	H99668	10.0
Building Containers for Intel® Omni-Path Fabrics using Docker and Singularity* Application Note*	J57474	9.0
Intel® Omni-Path Management API Programmer's Guide	J68876	8.0
Configuring Non-Volatile Memory Express (NVMe*) over Fabrics on Intel® Omni-Path Architecture Application Note*	J78967	2.0
Intel® Omni-Path Fabric Software Release Notes	K82053	1.0
Intel® Omni-Path Fabric Manager GUI Release Notes	K69636	2.0
Intel® Omni-Path Fabric Switches Release Notes (includes managed and externally-managed switches)	K82050	2.0
Intel® Omni-Path Fabric Unified Extensible Firmware Interface (UEFI) Release Notes	K50782	2.0
Intel® Omni-Path Fabric Thermal Management Microchip (TMM) Release Notes	K38341	2.0
Intel® Omni-Path Fabric Firmware Tools Release Notes	K50784	2.0

Related Links

Intel Omni-Path Documentation Library on page 30

1.17 Installation Requirements

This section provide installation requirements for this release.

1.17.1 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.

1.17.2 Software and Firmware Requirements

Table 1 on page 9 lists the operating systems supported by this release. For the required packages, refer to OS RPMs Installation Prerequisites.

1.17.3 OS RPMs Installation Prerequisites

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

- In addition to normal OS installation options, OS RPMs must be installed before you can install the Intel® Omni-Path software.
- Refer to the applicable section below to verify that all required RPMs are installed for the specific version of your OS distribution.
- Depending on the packages you choose, there may be additional prerequisites. For additional information, refer to the Release Notes for your specific release and installation type.

Red Hat* Enterprise Linux* OS RPMs

The tables below list the RPMs for each supported RHEL* release.

NOTES

- Some RHEL* rpms are available in a Server-Optional repository. Please install the RHEL* Server-Optional version from Red Hat* which contains additional, required development packages.
- Rdma-ndd is part of rdma-core.

Table 8. RHEL* 7.6 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	atlas	libstdc++-devel	bash	bison
infinipath-psm	bc	ncurses-libs	irqbalance	expat-devel
libibumad	coreutils	numactl-libs	kernel	flex
libibverbs	createrepo	openssl	kmod	libnl3-devel
libnl3	expat	openssl-devel	libgcc	libpfm
librdmacm	expect	openssl-libs	perl	libuuid-devel
opensm-libs	gcc-gfortran	pciutils	perl-Getopt-Long	ncurses-devel
perftest	glibc	redhat-rpm-config	perl-PathTools	numactl-devel
qperf	kernel-devel	rpm-build	perl-Socket	opensm-libs
rdma-core	libatomic	sysfsutils	pkgconfig	openssl-devel (1.0.1 or higher)
rdma-core-devel	libgfortran	tcl	python	tcl-devel
	libgomp	tcsh	systemd	valgrind-devel
	libquadmath	zlib	systemd-libs	
	libstdc++			

Table 9. RHEL* 7.7 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	atlas	libstdc++-devel	bash	bison
infinipath-psm	bc	ncurses-libs	irqbalance	expat-devel
libibumad	coreutils	numactl-libs	kernel	flex

continued...

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
libibverbs	createrepo	openssl	kmod	libnl3-devel
libnl3	expat	openssl-devel	libgcc	libpfm
librdmacm	expect	openssl-libs	perl	libuuid-devel
opensm-libs	gcc-gfortran	pciutils	perl-Getopt-Long	ncurses-devel
perftest	glibc	redhat-rpm-config	perl-PathTools	numactl-devel
qperf	kernel-devel	rpm-build	perl-Socket	opensm-libs
rdma-core	libatomic	sysfsutils	pkgconfig	openssl-devel (1.0.1 or higher)
rdma-core-devel	libgfortran	tcl	python	tcl-devel
	libgomp	tcsh	systemd	valgrind-devel
	libquadmath	zlib	systemd-libs	
	libstdc++			

Table 10. RHEL* 8 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	atlas	libstdc++-devel	bash	autoconf
libibumad	bc	ncurses-compat-libs	irqbalance	automake
libibverbs	coreutils	ncurses-libs	kernel	bison
librdmacm	createrepo	numactl-libs	kernel-modules-extra	elfutils-libelf-devel
opensm-libs	expat	openssl	kmod	expat-devel
perftest	expect	openssl-devel	libgcc	flex
qperf	gcc-gfortran	openssl-libs	perl	gcc-c++.x86_64
rdma-core	glibc	pciutils	perl-Getopt-Long	kernel-abi-whitelists
rdma-core-devel	kernel-devel	redhat-rpm-config	perl-Socket	kernel-rpm-macros
	libatomic	rpm-build	pkgconf	libnl3-devel
	libgfortran	sysfsutils	python2	libpfm
	libgomp	tcl	systemd	libtool
	libquadmath	tcsh		libuuid-devel
	libstdc++	zlib		ncurses-devel
				numactl-devel
				opensm-libs
				openssl-devel (1.1.1 or higher)
				tcl-devel

Table 11. RHEL* 8.1 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	atlas	libstdc++-devel	bash	autoconf
libibumad	bc	ncurses-compat-libs	irqbalance	automake
libibverbs	coreutils	ncurses-libs	kernel	bison
librdmacm	createrepo	numactl-libs	kernel-modules-extra	elfutils-libelf-devel
opensm-libs	expat	openssl	kmod	expat-devel
perftest	expect	openssl-devel	libgcc	flex
qperf	gcc-gfortran	openssl-libs	perl	gcc-c++.x86_64
rdma-core	glibc	pciutils	perl-Getopt-Long	kernel-abi-whitelists
rdma-core-devel	kernel-devel	redhat-rpm-config	perl-Socket	kernel-rpm-macros
	libatomic	rpm-build	pkgconf	libnl3-devel
	libgfortran	sysfsutils	python2	libpfm
	libgomp	tcl	systemd	libtool
	libquadmath	tcsh		libuuid-devel
	libstdc++	zlib		ncurses-devel
				numactl-devel
				opensm-libs
				openssl-devel (1.1.1 or higher)
				tcl-devel

SUSE* Linux* Enterprise Server OS RPMs

The tables below list the RPMs for each supported SLES* release.

NOTE

Some SLES* rpms are available in SLES* Software Development Kit (SDK). Please install the SLES* SDK iso from SUSE* which contains additional, required development packages.

Table 12. SLES* 12.4 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	bc	libnuma1	bash	bison
libibcm1	createrepo	libopenssl1_0_0	glibc	flex
libibmad5	expect	libopenssl-devel	grep	libexpat-devel
libibumad3	gcc-fortran	libquadmath0	irqbalance	libnuma-devel
libibverbs1	kernel-devel	libudev-devel	kmod	libopenssl-devel (1.0.1 or higher)

continued...

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
libpsm_infinipath1	kernel-syms	libz1	libedit0	libuuid-devel
librdmacm1	libatomic1	openssl	libgcc_s1	ncurses-devel
mpi-selector	libexpat1	rpm-build	libstdc++6	opensm-libs3
opensm-devel	libgfortran3	tcl	perl	tcl-devel
opensm-libs3	libgomp1	tcsh	perl-base	valgrind-devel
perftest	libncurses5		pkg-config	
qperf			python-base	
rdma-core			systemd	
rdma-core-devel			udev	
rdma-ndd				

Table 13. SLES* 12.5 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	bc	libnuma1	bash	bison
libibmad5	createrepo	libopenssl1_0_0	glibc	flex
libibumad3	expect	libopenssl-devel	grep	libexpat-devel
libibverbs1	gcc-fortran	libquadmath0	irqbalance	libnuma-devel
libpsm_infinipath1	kernel-devel	libudev-devel	kmod	libopenssl-devel (1.0.1 or higher)
librdmacm1	kernel-syms	libz1	libedit0	libuuid-devel
mpi-selector	libatomic1	openssl	libgcc_s1	ncurses-devel
opensm-devel	libexpat1	rpm-build	libstdc++6	opensm-libs3
opensm-libs3	libgfortran3	tcl	perl	tcl-devel
perftest	libgomp1	tcsh	perl-base	valgrind-devel
qperf	libncurses5		pkg-config	
rdma-core			python-base	
rdma-core-devel			systemd	
rdma-ndd			udev	

Table 14. SLES* 15 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	bc	libnuma1	bash	bison
libibmad5	coreutils	libopenssl1_1	glibc	flex
libibumad3	createrepo_c	libopenssl-devel	grep	kernel-devel
libibverbs1	expect	libosmcomp3	irqbalance	libexpat-devel
libnl3-200	gcc-fortran	libquadmath0	kmod	libnuma-devel

continued...

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
libpsm_infinipath1	kernel-devel	libudev-devel	libedit0	libopenssl-devel (1.0.1 or higher)
libquadmath0	kernel-syms	libz1	libgcc_s1	libuuid-devel
librdmacm1	libatomic1	openssl	libncurses5	ncurses-devel
libverbs1	libexpat1	rpm-build	libstdc++6	opensm-libs3
mpi-selector	libgfortran4	tcl	perl	tcl-devel
opensm-devel	libgomp1	tcsh	perl-base	valgrind-devel
opensm-libs3	libncurses6		pkg-config	
perftest			python-base	
qperf			systemd	
rdma-core			udev	
rdma-core-devel				
rdma-ndd				

Table 15. SLES* 15.1 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	bc	libnuma1	bash	bison
libibmad5	coreutils	libopenssl1_1	glibc	flex
libibumad3	createrepo_c	libopenssl-devel	grep	kernel-devel
libibverbs1	expect	libosmcomp3	irqbalance	libexpat-devel
libnl3-200	gcc-fortran	libquadmath0	kmod	libnuma-devel
libpsm_infinipath1	kernel-devel	libudev-devel	libedit0	libopenssl-devel (1.0.1 or higher)
libquadmath0	kernel-syms	libz1	libgcc_s1	libuuid-devel
librdmacm1	libatomic1	openssl	libncurses5	ncurses-devel
libverbs1	libexpat1	rpm-build	libstdc++6	opensm-libs3
mpi-selector	libgfortran4	tcl	perl	tcl-devel
opensm-devel	libgomp1	tcsh	perl-base	valgrind-devel
opensm-libs3	libncurses6		pkg-config	
perftest			python-base	
qperf			systemd	
rdma-core			udev	
rdma-core-devel				
rdma-ndd				

Example

The example below shows the install of a group of OS RPMs on a RHEL* OS server. Not all required OS RPMs are included in this example and some OS RPMs in this example might not be needed.

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

Intel recommends that you build your own list of OS RPMs for installation.

1.17.4 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, service, and login nodes.

The packages in the tgz file are RPMs. Installing individual RPMs is not supported in this 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 Switches Hardware Installation Guide* for related firmware requirements.

1.17.5 Installation Path Changes in Release 10.4 (and later)

If you are upgrading an 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

continued...

Old Location (Release 10.3 and earlier)	New Location (Release 10.4 and later)
`/etc/sysconfig/opaff.xml`	`/etc/opa/opaff.xml` Changed to `/etc/opa/opamgt_tls.xml` in 10.7.
`/etc/sysconfig/opamon.conf`	`/etc/opa/opamon.conf`
`/etc/sysconfig/ports`	`/etc/opa/ports`
`/etc/sysconfig/switches`	`/etc/opa/switches`

1.18 Product Constraints

- Power class 2 AOC are supported. You must use 10.5 (or newer) host software and 1.5 (or newer) UEFI for proper operation. Integrated HFI (-F) requires a specific BIOS level to support power class 2 AOC; contact your BIOS vendor for more information.
- The PM congestion weight for `XmitWaitPct` is set to 0 by default which causes the counter to be ignored. Setting a value other than 0 may lead to overreporting of congestion.
- Special considerations are required if using *Intel® Parallel Studio 2019, Update 2*.
  - The libpsm2 library provided in the Intel® OPA IFS package has been compiled with *Intel® Parallel Studio 2019, Update 5*. Note that it does not compile correctly with the Intel® Parallel Studio 2019, Update 2.

    Please use the *Intel® Parallel Studio 2019 - Update 5* compiler if you need to recompile the PSM2 library.
  - When linking OpenMPI against the IFS-provided version of libfabric and Intel® MPI Library 2019, Update 2 is installed, the following environment variable setting must be made to ensure proper linking with the IFS libfabric:

    ```
    source <IntelMPI installdir>/intel64/bin/mpivars.sh -
    ofi_internal=0
    ```

1.19 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:
  - One ASIC per Intel® Omni-Path Edge Switch 100 Series
  - Two ASICs per Intel® Omni-Path Director Class Switch 100 Series Leaf module
  - Two 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/opa-fm/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.

1.20 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 changes.

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

It is not necessary to enable Accelerated RDMA on all nodes in a cluster. The performance effects of enabling Accelerated RDMA on only a subset of nodes has not been characterized. It is highly recommended that you test to see if enabling this feature works best for your applications.

2.0 Issues

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

2.1 Resolved Issues

2.1.1 Issues Resolved in this Release

The following table lists issues that are resolved in this release.

Table 16. Issues Resolved in this Release

ID	Description	Resolved in Release
133604	Bonding driver shows incorrect hardware address of IPoIB interfaces. As of 10.10.1, this issue is no longer valid.	10.10.1
135028	NVMe over Fabric Protocol is only supported on Intel® OPA with Linux* kernel 4.5 and later versions. As of 10.10.1, this issue is no longer valid.	10.10.1
139368	Some applications compiled with older compilers may use a personality bit that signifies that READ should imply EXECUTE permissions. To improve system security, the hfi1 driver does not allow execute permissions on PSM memory maps. Therefore, applications that use READ implies EXECUTE will fail to run. As of 10.10.1, this issue is no longer valid.	10.10.1
145771	Due to changes in the SLES* 15 kernel, a user space application cannot access a hardware resource if that resource is being used by a kernel driver. The result is that the hfi1_eprom cannot access the EEPROM on an HFI when hfi1 kernel driver is using the device.	10.10.1
STL-43764	An FM running on SLES* 12.3 on certain older CPUs may experience instability due to a CPU hardware errata.	10.10.1
140310 (STL-46700)	On RHEL* 7.5, if an ipofabric interface name is non- standard, the ifcfg file is not read correctly. For example, if you are using the ipofabric interface name opa_ib0, and the connected mode and MTU size is specified in the ifcfg-opa_ib0 file, when you bring up the interface with ifup, the settings do not take effect. *Note:* RHEL* 7.5 is no longer supported in this release.	10.10
141005 (STL-47901)	In 10.8, the output of the module parameter `num_user_contexts` has been changed. The value of the module parameter `num_user_context` used at driver init time determines the number of receive contexts reserved for PSM. The default value of -1 caused the driver init sequence to determine how many CPUs are available, and assigned the context count to that value. *It would then update the module parameter `num_user_contexts` with that value.* This incorrect behavior has been fixed. `num_user_contexts` is no longer updated.	10.10

continued...

ID	Description	Resolved in Release
STL-57790	A known issue with OpenMPI 2.1.2 and one-sided operations (such as MPI_Accumulate) when running under RHEL 8 results in an MPI internal error.	10.10
STL-57800 STL-58303	INSTALL enables ARPTABLE_TUNING by default. With this flag, Intel® OPA modifies `/etc/sysctl.conf` for arp table size adjustment. If you modify this file after ARPTABLE_TUNING is enabled, your changes are not permanent. Rebooting the machine will restore the file to Intel® OPA's version.	10.10
STL-58338	Under RHEL 8, installing mpitests_openmpi_gcc_hfi with yum fails because no package provides `libmpi.so.20()`(64-bit). This shared library is shipped in openmpi_gcc_hfi; however, the openmpi_gcc_hfi rpm fails to claim that it provides `libmpi.so.20`.	10.10

2.1.2 Issues Resolved in Prior Releases

The following table lists issues that were resolved in prior releases.

Table 17. Issues Resolved in Prior Releases

ID	Description	Resolved in Release
STL-58309	It is possible for a memory region not to be freed due to the lack of rcu synchronization for the memory region's refcount. This has been observed when executing the `mmvdisk` while running GPFS version RAID version 5.0.2-3.0.1.	10.9.3.1
STL-46789	SLES 12.2 shipped with an old version ibacm that does not include the unit file. Since IFS INSTALL relies on systemd to turn on/off services, enabling ibacm with IFS INSTALL fails. *Note:* SLES 12.2 is no longer supported.	10.9.3
STL-56887	When installing OPA with CUDA, `openmpi_gcc_cuda_hfi` is required. When creating a repository for OPA without CUDA, opacreaterepo script recommends `openmpi_gcc_cuda_hfi` by mistake.	10.9.3
STL-47578	Due to a SLES* 15 kernel setting, hfi1_eprom cannot work while the HFI driver is loaded. The tool and driver are mutually exclusive.	10.9.2
STL-55620	When Accelerated IPoFabric (AIP) is enabled, less cores are available for HFI driver sdma interrupts. The exact number depends on the number of CPU cores in the system. This could negatively impact send performance.	10.9.2
STL-56600	Using `./INSTALL -n` to do a fresh installation will not set the irqbalance policy. This can cause a performance issue.	10.9.2
STL-47546	When an ISL goes down in the middle of an FM sweep (due to a disruption in the fabric such as a reboot), the SA copy of topology becomes invalid when the Fattree routing algorithm is used. SA queries that use this topology (e.g., path record query) fail. *Note:* A path record query failure can be seen in FM log as "INVALID TOPOLOGY" messages. The issue will resolve after the FM's next successful sweep.	10.9.1
STL-46211 STL-47907	For SLES OSes, when IPoIB is enabled and working, the IFS INSTALL erroneously displays IPoIB as "disabled". When the IFS is upgraded, this display error causes the INSTALL to disable IPoIB in order to keep the same configuration.	10.9.1
STL-48065	In some boot scenarios the BIOS can order the device paths such that it passes the device handle of a device implementing the EFI_NETWORK_INTERFACE_IDENTIFIER_PROTOCOL, which is not a child of the HFI device path. In these cases, the HFI UEFI driver can attempt to obtain access to internal structures based on the physical memory location of the controller name passed. This can result in an unexpected behavior in the BIOS. This is addressed by explicitly confirming that the controller handle passed by the BIOS matches the HFI device controller for child device handles.	10.9.0.1

continued...

ID	Description	Resolved in Release
132207	Kernel crash caused by the ib_srpt module.	10.9.0
139743 143031 143115	Under a very heavy load through the IPoIB interface, the kernel warning `NETDEV WATCHDOG: ib0 (hfi1): transmit queue 0 timed out`, followed by the messages `queue stopped 1, tx_head xxx, tx_tail xxx` and `transmit timeout: latency xxxx msecs` may be seen. As of V10.9.0, this issue is no longer valid.	10.9.0
143296	When irqbalance uses the argument `--hintpolicy=exact`, it applies the policy of setting the hardware interrupts to CPU core mappings according to device drivers preferences. For the HFI1 driver, it is strongly recommended to preserve interrupt locality for low latency and high bandwidth by having a dedicated CPU core per interrupt.	10.9.0
141740	Third-Party Issue: Due to the race condition during boot up ipoib driver can miss PKey change event that will leave ipoib interface in disconnected state.	10.9.0
144996	Running workloads with more than 78 ranks with the Open MPI OFI MTL over OFI Verbs;OFI_RXM provider may result in a hang with message sizes larger than 65 KB.	10.9.0
145415	ESM hangs after the first sweep in a configuration that includes all `<VirtualFabric>` sections that are enabled and with `QOS =1`, where `<BaseSL>` is defined in all with values other than `0`. Symptoms of this hang are: • SA or PA queries to ESM time out (e.g., opareport fails with FTIMEOUT) • SM does not resweep after SweepInterval, multicast membership change, or fabric change • smControl shutdown is issued but ESM never reaches stopped state	10.9.0
145474	OFI Verbs mpi_stress may cause verbs/MSG provider completion queue overrun that results in dropped completions. They show up as sequence errors in the test.	10.9.0
145585	SLES* 15 `./INSTALL` script does not properly run `dracut -f`.	10.9.0
145623	For systems running on SLES* 15, there is a known issue with irqbalance.	10.9.0

145855	If the Admin VF is not running on VL0, the HSM may get into a state where it is unable to talk to the fabric. The sweep will log the following errors:
	`opamgt ERROR: [<pid>] omgt_send_mad2: send failed; Invalid argument, agent id 2`
	`MClass 0x81 method 0x1 attrId 0x11 attrM 0x0`
	`WARN [topology]: SM: sm_send_stl_request_impl: Error Sending to Path:[ 1 ]`
	`Lid:[0xffffffff] [Can't find node in topology!]. AID:[NODEINFO]`
	`TID:[0x0000000000000031] Status:[OK (0x00000000)]`
	`WARN [topology]: SM: topology_main: TT: too many errors during sweep, will`
	`re-sweep in a few seconds rc: 108: unrecoverable error`
		10.9.0

146271	When using `libpsm2` with `PSM2_CUDA=1`, workloads may assert() during MPI_Recv or MPI_ collective operations where the source buffer is a GPU buffer and located on the same node, i.e., intranode transfer. The destination buffer locality has no influence. This assert() will only affect transfers that are not aligned to the start of the source buffer.	10.9.0
146456	In a fabric with only one Edge switch using the fat tree routing algorithm, a port can get stuck in the `Init (LinkUp)` state after the port is bounced.	10.9.0
STL-46606 STL-47956 STL-48661	Bouncing a link or rebooting a device under certain fabric conditions may cause a switch in the fabric to be removed from the Fabric Manager's internal view of the topology leading to fabric disruptions and instability.	10.9.0
STL-46790	In cases where GSI services are active and the FM is receiving capability change traps (common after node reboots), FM responsiveness may be impacted. This could result in data traffic disruption or unexpected FM failovers. GSI traffic would include the PM, SA, and DBSync (FM failover).	10.9.0

2.2 Open Issues

The following table lists the open issues for this release.

Table 18. Open Issues

ID	Description	Workaround
134819	In KNL-F EFI shell, the command `ifconfig -l` does not correctly display the IP address after being assigned via DHCP.	Launch a newer version of the EFI shell from the embedded shell.
135830 (STL-46193)	On Intel® Xeon Phi™ systems, failure observed during software upgrade when rebuilding the boot image. Error message contains: Rebuilding boot image with "/usr/bin/dracut -f"	Install parallel gzip (pigz) and make it the default compression tool prior to upgrading the software.
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.	Avoid the use of Legacy BIOS boot mode if your platform does not execute the HFI driver in that mode. Avoid BIOS settings or other configuration settings that do not execute the HFI driver during boot.
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 `lspci` 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.
139995 (STL-49724)	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.
143174 (STL-47003)	Due to a SLES* 15 kernel setting, hfidiags cannot work while the HFI driver is loaded. The tool and driver are mutually exclusive.	Boot the kernel with `iomem=relaxed`. This turns off the resource exclusivity check.
STL-47125	In rare cases, a program crash may be experienced during process cleanup when running with the OpenMPI PSM2 MTL (Matching Transport Layer).	Restart the workload.
STL-48921	If you attempt to unload or restart the hfi1 driver while the ib_ipoib driver is loaded, it may fail with a message similar to this: `modprobe: FATAL: Module hfi1 is in use`	To avoid this issue, unload the ib_ipoib driver before unloading or restarting the hfi1 driver.
STL-49210	For RHEL* 7.6 and newer and SLES* 15, SLES* 12.4 and newer, when Accelerated IP (AIP) is disabled and IPoFabric is configured in datagram mode, no bulk traffic will occur if the configured ib0 MTU size is larger than what default (non-AIP) IPoFabric can support.	Enable AIP or reduce the locally configured ib0 MTU size to a size that default IPoFabric can support (e.g., 4092 bytes).
STL-49313	On OS versions (RHEL* 7.6 and newer and SLES* 15, SLES* 12.4 and newer) where it is available, configuring IPoIB module parameter `ipoib_enhanced=0` disables AIP on same node.	To run with AIP, ensure enhanced IPoIB is enabled (`ipoib_enhanced=1`).
STL-55704	IPoIB and core/mcast overreact to SM LID change by unnecessarily flushing connections, path records, and multicast membership.	Contact Intel Customer Support for more information.

continued...

ID	Description	Workaround
STL-56414	In SLES* distros, libpsm_infinipath1 is designated to obsolete libpsm2-compat even though the rpms can coexist. For some use cases (such as rebuilding MPI from source), both rpms must be installed. When a user installs IFS using zypper, the process fails when installing libpsm2-compat due to this conflict.	Run the following command under the IFS package folder before the zypper-based IFS installation to force install libpsm2-compat: `rpm -ivh --nodeps repos/OPA_PKGS/RPMS/libpsm2-compat-<version>.x86_64.rpm`
STL-56557	The primary IPoIB network device associated with any RDMA device may fail to join certain multicast groups. This can prevent IPv6 neighbor discovery and possibly other network ULPs from working correctly. Note that the IPv4 broadcast group is not affected as the IPoIB network device handles joining that multicast group directly. This problem does not affect IPoIB child network devices.	Perform one of the following workarounds: • Do not use IPv6 over a parent IPoIB device. • Only load the IPoIB module after the HFI link is ACTIVE with a valid pkey assigned. • Bounce (unload, then load) the IPoIB module anytime a pkey change occurs.

STL-57127	Restarting the irqbalance service while the HFI1 driver is loaded may cause improper or inconsistent distribution of interrupts, which may result in low network transfer performance.	Irqbalance must be stopped before starting the HFI1 module. Perform the following command sequence:
		`systemctl stop irqbalance`
		`modprobe -r ib_ipoib`
		`modprobe -r hfi1`
		`modprobe hfi1`
		`modprobe ib_ipoib`
		`systemctl start irqbalance`

STL-57994	The OSU benchmarks built and packaged with the CUDA-enabled MPI packages included in the IFS software distribution are not built with CUDA support.	Download the OSU benchmarks from http://mvapich.cse.ohio-state.edu/benchmarks/. Follow the instructions for enabling CUDA when building.
STL-58332	When upgrading from RHEL 7.5 to a newer version, if a value is set for `ifs_sel_mode` in hfi1.conf HFI driver will not load.	Remove the parameter and value.
STL-58443	When TID RDMA is enabled with 256 nodes or greater, it is possible that a storage write may not complete.	Disable TID RDMA.
STL-59413	Using PSM2 with CUDA enabled on top of a non-CUDA enabled hfi1 driver may result in invalid behavior.	Ensure the PSM2 CUDA library is used only in conjunction with the hfi1 driver with GPUDirect/CUDA support. *Note:* If the gpu-direct enabled driver is present, the string `gpu-direct` will be returned in the Driver version output of `hfi1_control -i`.
STL-59663	Certain conditions result in hfi1 driver interrupts not being distributed correctly, even with irqbalance running. This may result in a large run to run performance variation or low overall performance.	Refer to the*Intel® Omni-Path Fabric Performance Tuning User Guide*, Identifying to Which CPU Core an Interrupt is Bound section for help with determining which cores are running the hfi1 interrupts. If interrupts are arbitrarily distributed or significantly reusing select cpu cores, refer to Manually Changing IRQ Affinity section. Contact Intel Customer Support if you need further guidance.
STL-59923	The version output of `modinfo hfi1` shows 10.9-0 for IFS 10.10 releases.	None. A fix will be available in a later release.

continued...

ID	Description	Workaround
STL-59934	Customers experience errors similar to `hfi1_0: Send Context 8(151) Error: WriteOverflow` in dmesg or console logs.	Contact Intel Customer Support for more information. Setting `pio_threshold=0` may prevent the error.
STL-59944	Installing IFS on SLE HPC 15.x fails because there is no file `/etc/SuSE-release`.	Create symlink from `/etc/SuSE-release` to `/etc/os-release` with the following command: `ln -s /etc/os-release /etc/SuSE-release`
STL-59955	irqbalance cannot start hint script when SElinux is enabled and enforcing.	Disable SELinux. Refer to the Red Hat Bugzilla https://bugzilla.redhat.com/show_bug.cgi?id=1784080

2.2.1 Third Party Open Issues

The following table lists the third party open issues for this release.

Table 19. Third Party Open Issues

ID	Description	Workaround
129563 (STL-47095)	Memory allocation errors with MVAPICH2-2.1/Verbs.	*Note:* To avoid this issue, use MPIs over PSM. If you are using MPIs over verbs, the following workaround is required: • When running MVAPICH2 jobs with a large number of ranks (for example, > 36 ranks but ≤ 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, where $1 \le n \le 23$) from its default of 16 to 17. For instructions on setting module parameters, refer to the *Intel® Omni-Path Fabric Performance Tuning User Guide*, HFI1 Driver Module Parameters chapter.
139924	For SLES*, the ibacmp provided in the OS distribution uses incorrect address information when joining multicast groups. This causes name resolution to fail. The dsap provided in the OS distribution works correctly.	The fix for this issue is available in the library rdma-core-15-2.el7. • Do not use the OS distribution ibacmp. • Install rdma-core-15-2.el7 manually.
141273 (STL-46935)	The in-distro version of perftests has bugs.	Use the upstream version of perftest from https://github.com/linux-rdma/perftest.

continued...

ID	Description	Workaround
STL-47571	Since libfabric 1.6 (which is included in IFS 10.8), the psm2 provider maps OFI endpoints directly to HFI contexts instead of multiplexing multiple OFI endpoints to a single HFI context. This relies on the multi-EP feature of the PSM2 library and thus the provider automatically sets PSM2_MULTI_EP=1 if it has not been set. However, enabling the multi-EP feature also disables context sharing. As the result, applications may experience the following runtime error when trying to oversubscribe CPU cores (which is usually the same as available HFI contexts). `hfi_userinit: assign_context command failed: Device or resource busy PSM2 can't open hfi unit: -1 (err=23)` *Note:* Applications that don't use libfabric are not affected.	Set `PSM2_MULTI_EP=0`. *Note:* This only works for applications that open only one OFI endpoint per process.
STL-48480	When running workloads with Intel® MPI or OpenMPI using OFI RXM and Verbs providers, if more than 100 ranks are running, it is possible for the workload to hang during message passing.	If Verbs messaging is not required, use one of the following methods to run with the PSM2 provider: • Intel® MPI: `-genv I_MPI_OFI_PROVIDER 'psm2'` • OpenMPI: `--mca mtl_ofi_provider_include 'psm2'`
STL-56236	When running workloads with OpenMPI using the OFI RXM and Verbs providers, a rare segmentation fault may be encountered during message processing with many ranks. This issue will result in the workload failing and will require the workload to be rerun.	If Verbs messaging is not required, PSM2 messaging can be used to workaround this issue by running: `--mca mtl_ofi_provider_include 'psm2'`
STL-57040	If Intel® MPI 2019, Update 2 and OpenMPI are installed on the same cluster, OpenMPI jobs could link to the Intel® MPI internal libfabric library after running Intel® MPI mpivars.sh.	Before running an OpenMPI job, perform the following to reset the libfabric environment variables that had been previously set by Intel® MPI. `source <installdir>/intel64/bin/mpivars.sh -ofi_internal=0`
STL-59919	Creating a ram disk using `dmsetup` on RHEL 7.7 will cause a kernel panic. The creation sequence is: 1. `modprobe brd rd_nr=1 rd_size=33554432 max_part=1` 2. `blockdev --getsize /dev/ram0` 3. **`dmsetup`** `create mdevice0 --table "0 67108864 linear /dev/ram0 0`	None. Note that RHEL 7.6 does not have this issue.

3.0 Related Information

3.1 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	*Intel® Omni-Path Fabric Quick Start Guide*	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	*Intel® Omni-Path Fabric Setup Guide*	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.
Installing hardware	*Intel® Omni-Path Fabric Switches Hardware Installation Guide*	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.
	Intel® Omni-Path Host Fabric Interface Installation Guide	Contains instructions for installing the HFI in an Intel® OPA cluster.
Installing host software Installing HFI firmware Installing switch firmware (externally-managed switches)	*Intel® Omni-Path Fabric Software Installation Guide*	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)	*Intel® Omni-Path Fabric Switches GUI User Guide*	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

continued...

Task	Document Title	Description
Managing a switch using the CLI Installing switch firmware (managed switches)	*Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide*	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	*Intel® Omni-Path Fabric Suite FastFabric User Guide*	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	*Intel® Omni-Path Fabric Suite Fabric Manager User Guide*	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.
	Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide	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	*Intel® Omni-Path Fabric Host Software User Guide*	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	*Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide*	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	*Intel® Omni-Path Fabric Performance Tuning User Guide*	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	*Intel® Omni-Path IP and LNet Router Design Guide*	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	*Building Containers for Intel® Omni-Path Fabrics using Docker* and Singularity* Application Note*	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	*Intel® Omni-Path Management API Programmer's Guide*	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	*Configuring Non-Volatile Memory Express* (NVMe*) over Fabrics on Intel® Omni-Path Architecture Application Note*	Describes how to implement a simple Intel® Omni-Path Architecture-based point-to-point configuration with one target and one host server.

continued...

Task	Document Title	Description
Learning about new release features, open issues, and resolved issues for a particular release	*Intel® Omni-Path Fabric Software Release Notes*	
	Intel® Omni-Path Fabric Manager GUI Release Notes	
	Intel® Omni-Path Fabric Switches Release Notes (includes managed and externally-managed switches)	
	Intel® Omni-Path Fabric Unified Extensible Firmware Interface (UEFI) Release Notes	
	Intel® Omni-Path Fabric Thermal Management Microchip (TMM) Release Notes	
	Intel® Omni-Path Fabric Firmware Tools Release Notes	

3.1.1 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.