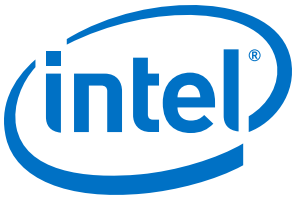


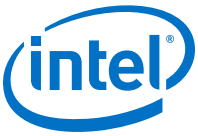
Intel® Omni-Path Fabric Software in Red Hat* Enterprise Linux* 7.6

Release Notes

Rev. 2.0

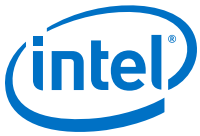
March 2019

Doc. No.: K34562, Rev.: 2.0



Contents

Tables



1.0 Overview of the Release

These Release Notes are intended for Intel® Omni-Path IFS software provided inbox with the OS 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 Document Versions

Intel® Omni-Path publications are available at the following URLs. For documents compatible with this release, refer to Release 10.7.

- Intel® Omni-Path Switches Installation, User, Reference Guides, and Release Notes

  http://www.intel.com/omnipath/SwitchPublications

- Intel® Omni-Path Software Installation, User, Reference Guides, and Release Notes (includes HFI documents)

  http://www.intel.com/omnipath/FabricSoftwarePublications

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

Table 1. Supported Document Versions

Title	Doc. Number	Revision
Intel® Omni-Path Fabric Quick Start Guide	J57479	4.0
Intel® Omni-Path Fabric Setup Guide	J27600	8.0
Intel® Omni-Path Fabric Switches Hardware Installation Guide	H76456	7.0
Intel® Omni-Path Host Fabric Interface Installation Guide	H76466	5.0
Intel® Omni-Path Fabric Software Installation Guide	H76467	9.0
Intel® Omni-Path Fabric Switches GUI User Guide	H76457	9.0
Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide	H76458	9.0
Intel® Omni-Path Fabric Suite FastFabric User Guide (Merged with: *Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide*)	H76469	9.0

continued...

Title	Doc. Number	Revision
Intel® Omni-Path Fabric Suite Fabric Manager User Guide	H76468	9.0
Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide	H76471	9.0
Intel® Omni-Path Fabric Host Software User Guide	H76470	9.0
Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide	H76473	9.0
Intel® Omni-Path Fabric Performance Tuning User Guide	H93143	11.0
Intel® Omni-Path IP and LNet Router Design Guide (Old title: *Intel® Omni-Path IP and Storage Router Design Guide*)	H99668	6.0
Building Containers for Intel® Omni-Path Fabrics using Docker and Singularity* Application Note*	J57474	4.0
Intel® Omni-Path Management API Programmer's Guide	J68876	3.0
Configuring Non-Volatile Memory Express (NVMe*) over Fabrics on Intel® Omni-Path Architecture Application Note*	J78967	1.0
Intel® Omni-Path Fabric Software Release Notes	J95967	1.0
Intel® Omni-Path Fabric Manager GUI Release Notes	J95968	1.0
Intel® Omni-Path Fabric Switches Release Notes (includes managed and externally-managed switches)	J95964	1.0
Intel® Omni-Path Fabric Unified Extensible Firmware Interface (UEFI) Release Notes	J98868	1.0
Intel® Omni-Path Fabric Thermal Management Microchip (TMM) Release Notes	J98871	1.0
Intel® Omni-Path Fabric Firmware Tools Release Notes	J98870	1.0

1.3 Software License Agreement

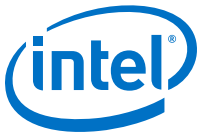
This software is provided under license agreements and may contain third-party software under separate third-party licensing. Please refer to the license files provided with the software for specific details.

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 Packages in This Release

Intel® Omni-Path Software Packages
Packages created by Intel
opa-address-resolution-10.7.0.0.133-1.el7.x86_64
opa-basic-tools-10.7.0.0.133-1.el7.x86_64
opa-fastfabric-10.7.0.0.133-1.el7.x86_64
opa-fm-10.7.0.0.141-1.el7.x86_64
opa-libopamgt-10.7.0.0.133-1.el7.x86_64
continued...



Intel® Omni-Path Software Packages
libfabric-1.6.1-2.el7.x86_64
libpsm2-10.3.58-1.el7.x86_64
Firmware binaries delivered by Intel
8051 firmware version 1.26.1
SBus Master firmware version 0x10130001
PCIe SerDes firmware version 0x4755
Fabric SerDes firmware version 0x1055
Packages used by Intel
rdma-core-17.2-3.el7.x86_64 (libhfi1)
openmpi-1.10.7-2.el7.x86_64
openmpi3-3.0.2-1.el7.x86_64
mpitests-openmpi3-5.4.2-1.el7.x86_64
mpitests-mvapich2-5.4.2-1.el7.x86_64
mvapich2-2.2-psm2-2.2-4.el7.x86_64
mvapich23-psm2-2.3-4.el7.x86_64
mpitests-mvapich2-psm-5.4.2-1.el7.x86_64

To download Intel programmable firmware, refer to the following:

- Unified Extensible Firmware Interface (UEFI)
- Thermal Management Module (TMM)
- Firmware Tools

1.6 Supported Features

- The list of supported hardware is in Table 3 on page 8.
- Product Constraints described in Product Constraints on page 11.
- UEFI, TMM, and Firmware Tools are now standalone rpms.
- Active Optical Cables. For details, see the Cable Matrix at: https://www.intel.com/content/www/us/en/products/network-io/high-performance-fabrics/omni-path-cables.html
  - Support for active optical cables (AOC) on server platforms using integrated HFI for OPA (commonly known as "-F").
  - Support for Power Class 2 active optical cables (AOC). See Product Constraints on page 11 for more information.
- Legacy BIOS Boot Mode Enhancements to support boot over fabric, custom board descriptions, and pre-boot platform configuration data for AOC support.
- Multi-endpoint functionality. See the *Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide* for details.
- Support for OpenFabrics Interfaces (OFI), a framework that includes libraries (including libfabric) and applications used to export fabric communication services to applications.

- Support for NVMe over Fabric Protocol
- 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.
- 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.

1.7 Supported MPI Libraries

The table below lists the different MPI libraries supported by Intel® Omni-Path Fabric Software. Note that the second column indicates whether the MPI library is included in the distribution or not.

Table 2. Supported MPI Libraries

MPI Implementation	Included in Distribution?	Runs Over
Open MPI 2.1.2	Yes	PSM2
MVAPICH2-2.3B	Yes	PSM2

1.8 Intel Hardware

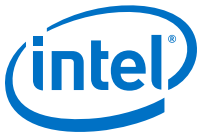
The following table lists the Intel hardware supported in this release.

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

Table 3. 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
Intel® Xeon Phi™ x200 Product Family	Knights Landing CPU-based servers
Intel® Xeon Phi™ 72x5 Processor Family	Knights Mill 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

continued...



Hardware	Description
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

1.9 Intel® OPA Compatibility Matrix

The following component versions are compatible with Intel® Omni-Path software in RHEL* 7.6.

Table 4. Intel® OPA Compatibility Matrix

UEFI	TMM	Managed Switch	Externally-Managed Switch	FM GUI
1.8.1.0.0	10.8.0.0.214	10.8.0.0.186	10.8.0.0.186	10.8.0.0.206
1.7.2.0.0	10.7.0.0.3	10.7.0.0.146	10.7.0.0.144	10.7.0.0.145
1.6.0.0.0	10.4.0.0.146	10.6.1.0.3	10.6.1.0.1	10.6.0.0.136

1.10 Installation Requirements

This section provides instructions and information on installing the software.

1.10.1 Installation Instructions

Perform the steps in this section to install the default Intel® Omni-Path Software configuration.

Assumptions

- You are logged in as root or with root privileges.
- You have a list of IPv4 addresses and netmasks for each IPoIB interface you are going to set up.
- RHEL* packages are available in a yum repository.

References

- Refer to the *Intel® Omni-Path Fabric Software Installation Guide* for related software requirements and next steps.
- Refer to the *Intel® Omni-Path Fabric Switches Hardware Installation Guide* for related firmware requirements.

Procedures

Perform the following steps to install the default Intel® OP Software configuration using RHEL* OS:

Step	Task/Prompt	Action
Install OPA-Basic Software		
1.	At the command prompt, enter the installation command for `opa-basic-tools`.	Type **`yum install -y opa-basic-tools`** and press **Enter**.
2.	At the command prompt, reboot the server.	Type **`reboot`** and press **Enter**.

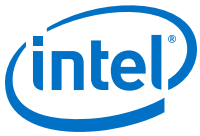
3.	Check your link using `opainfo`.	Type **`opainfo`** and press **Enter**.
		Example output:
		`hfi1_0:1 PortGID: 0xfe80000000000000:001175010163f931`
		`PortState: Active`
		`LinkSpeed Act: 25Gb En: 25Gb`
		`LinkWidth Act: 4 En: 4`
		`LinkWidthDnGrd ActTx: 4 Rx: 4 En: 3,4`
		`LCRC Act: 14-bit En: 14-bit,16-bit, 48-bit Mgmt: True`
		`LID: 0x00000010-0x00000010 SM LID: 0x0000000c SL: 0`
		`QSFP: AOC , 5m FINISAR CORP P/N FCBN425QB1C05 Rev A`
		`Xmit Data: 0 MB Pkts: 251`
		`Recv Data: 0 MB Pkts: 251`
		`Link Quality: 5 (Excellent)`

4.	Install the `rdma-core` rpm.	Type **`yum install -y rdma-core`** and press **Enter**.
5.	On all compute nodes: install the PSM2 library.	Type **`yum install -y libpsm2`** and press **Enter**.
Install Intel® Omni-Path Fabric Suite Components on the Management Node		
6.	Install FastFabric.	Type **`yum install -y opa-fastfabric`** and press **Enter**.
7.	Install the `opa-address-resolution` rpm on all nodes.	Type **`yum install -y opa-address-resolution`** and press **Enter**.
8.	Install Fabric Manager.	Type **`yum install -y opa-fm`** and press **Enter**.
9.	Start the Fabric Manager.	Type **`systemctl start opafm`** and press **Enter**.
Set up IPoIB IPV4 Configuration		

10.	Manually edit or create the `ifcfg-ib`*X* file.	*Note:* Use the OS distribution-supplied instructions for setting up network interfaces.
		Type **`cat /etc/network-scripts/ifcfg-ib0`** and press **Enter**.
		Example output:
		`DEVICE=ib0`
		`BOOTPROTO=static`
		`IPADDR=10.228.200.173`
		`BROADCAST=10.228.203.255`
		`NETWORK=10.228.200.0`
		`NETMASK=255.255.252.0`
		`ONBOOT=yes`
		`CONNECTED_MODE=yes`
		`MTU=65520`

11.	Bring up the ib0 interface.	Type **`ifup ib0`** and press **Enter**.
12.	Perform a test ping.	Type **`ping <remote IPoIB address>`** and press **Enter**.

continued...



Step	Task/Prompt	Action

		For example:
		`ping 10.228.200.161`
		`PING 10.228.200.161 (10.228.200.161) 56(84) bytes of data.`
		`64 bytes from 10.228.200.161: icmp_seq=1 ttl=64 time=0.863 ms`

End Task		

1.11 Product Constraints

- The minimum firmware version for Intel® Omni-Path Host Fabric Interface Silicon 100 Series Switch ASIC is 10.6.
- Power class 2 AOC are supported. You must use RHEL* 7.5 (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.

1.12 Product Limitations

This release has the following product limitations:

- 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.
- The fabtests tests distributed with libfabric 1.6.1 fail for the multi_ep psm test.

  A fix can be found in 1.6.2, located upstream in the following repository:

  https://github.com/ofiwg/libfabric

2.0 Issues

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

2.1 Resolved Issues

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

Table 5. Issues Resolved in this Release

ID	Description	Resolved in Release
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.	RHEL* 7.6
135390	The driver can parse older versions of the platform configuration file.	RHEL* 7.6
136728	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.) A new configuration value is present in the FM configuration that determines how much time will be allotted to timeouts before abandoning a sweep. If you are upgrading from a previous version of the FM and retaining a configuration file that does not include this new parameter, the value will be set too low and cause sweeps to abandon after only a single timeout is witnessed.	RHEL* 7.6
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.	RHEL* 7.6
139834	When using the FastFabric TUI to run "Perform Single Host Verification", the test hangs during operation.	RHEL* 7.6
140229	The opaswitchadmin tool was updated to address a condition that was seen during firmware upgrade of a large number of switches.	RHEL* 7.6
140881	In rare cases when an LNI failure occurs, the link will not come up after manually disabling and re-enabling the link.	RHEL* 7.6
140911	The OFI verbs provider does not support FI_EP_RDM End Point type. This End Point type is needed for Open MPI OFI support. Therefore, Open MPI OFI support will not run over the verbs provider.	RHEL* 7.6

141219	When adaptive routing is disabled, the output for `opasmaquery` for `portgroup` appears as shown below:
	`# opasmaquery -l 1 -o portgroup`
	`PG: 0x0000 Egress:None`
		RHEL* 7.6

141845	Resolved FM process out of memory condition	RHEL* 7.6
141909	Resolved multiple FM synchronization issue that can lead to FM failure.	RHEL* 7.6

2.2 Open Issues

The following table lists the open issues for this release.

Table 6. Open Issues

ID	Description	Workaround
129563	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 $1 \leq n \leq 23$) 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.
135830	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"	Due to the extended processing time of the dracut command on the Intel® Xeon Phi™ platform, Intel recommends the following: • Install and configure Intel® Xeon Phi™ systems separately. • Change the FF_TIMEOUT_MULT value in `opafastfabric.conf` from 2 to 6 for Intel® Xeon Phi™ systems.
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 root, run the execstack tool to clear the executable bit on the binary: `execstack -c <binary>` Alternatively, recompile the binary to not set this personality bit.
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.
142330	MPI applications that leverage the PSM2 library's access to the HFI ASICs Memory Mapped IO and that access the MMIO directly (not via PSM2) can potentially cause an "unsupported opcode" error which some servers handle as a critical error.	Disable upstream error reporting using the AER mask register. • For discrete HFI ASICs, use `setpci -d 8086:24f0 ECAP_AER +8.l=00100000:00100000`

continued...

ID	Description	Workaround
		• For integrated HFIs, use `setpci -d 8086:24f1 ECAP_AER +8.l=00100000:00100000`
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.	Always start the user-space process irqbalance using the argument `--hintpolicy=exact`.
143449	PM will scroll LQI=0 and Integrity Exceeded Threshold logs when an additional VF with QoS enabled and a device group that is not "All". *Note:* This issue does not occur when running against the default opafm.xml configuration file.	Set the `<ProcessVLCounters>` field in the opafm.xml configuration to `0` to stop scrolling of logs related to LQI.