# Intel® Omni-Path Fabric Software

**Release Notes for 10.3.2**

*September 2017*

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

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

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

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

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

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

Intel, Intel Xeon Phi, Xeon, and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.
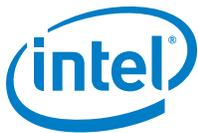
*Other names and brands may be claimed as the property of others.

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

# Contents

## Tables

§

# 1.0    Overview of the Release

## 1.1    Introduction

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

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

## 1.2    Audience

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

## 1.3    Software License Agreement

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

## 1.4    If You Need Help

Technical support for Intel® Omni-Path products is available 24 hours a day, 365 days a year. Please contact Intel Customer Support or visit http://www.intel.com/omnipath/support for additional detail.

## 1.5    New Enhancements and Features in this Release

The 10.3.2 release contains no new features; this release is a maintenance release that contains bug fixes. See Table 2-1 for details.

The following enhancements and features were new in the 10.3.1 release:

- Support for Token ID (TID) RDMA, which is a Verbs protocol extension. See Section 1.16 for details.
- Support for SKX and SKX-F hardware.
- Supports RHEL* 6.7 and CentOS* 6.7.
- Support for active optical cables (AOC) on server platforms using integrated HFI for OPA (commonly known as "-F").

## 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. For details, see the Cable Matrix at: http://www.intel.com/content/www/us/en/high-performance-computing-fabrics/omni-path-cables.html

- MPI applications are provided in a stand-alone package.

- Intel® Xeon® v4 processor (codename Broadwell) support

- Intel® Xeon Phi™ support

- 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.

- Added support for Enterprise Edition for Lustre* software version 3.1.

- 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　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-diagtools-sw, libhfi1verbs

    — Open MPI built for verbs and PSM2 using gcc, and Intel compilers.

    — MVAPICH2 built for verbs and PSM2 using gcc, and Intel compilers.

    — mpitests
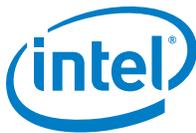
— mpi-selector

— 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 *Building Lustre\* Servers with Intel® Omni-Path Architecture Application Note* for details.

    *Note:* In the Intel® Omni-Path Software package for RHEL\* 7.2, the hfi1 driver and ifs-kernel-updates are supplied as a smaller package.

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 |
|---|---|---|
| HFI1 UEFI Option ROM | HfiPcieGen3_1.3.2.0.0.efi | 1.3.2.0.0 |
| UEFI UNDI | HfiPcieGen3Loader_1.3.2.0.0.rom | 1.3.2.0.0 |
| HFI1 SMBus Microcontroller Firmware (Thermal Monitor) | hfi1_smbus.fw | 10.2.1.0.3 |

## 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\*) 6.7 X86_64 | Update 7 | 2.6.32-573.el6.x86_64 |
| CentOS\* 6.7 X86_64 | Update 7 | 2.6.32-573.el6.x86_64 |
| Red Hat\* Enterprise Linux\* (RHEL\*) 7.2 X86_64 | Update 2 | 3.10.0-327.el7.x86_64 |
| Red Hat\* Enterprise Linux\* (RHEL\*) 7.3 X86_64 | N/A | 3.10.0-514.el7.x86_64 |
| CentOS\* 7.2 X86_64 | N/A | 3.10.0-327.el7.x86_64 |

**Table 1-2.    Operating Systems Supported**

| Operating System | Update/ SP | Kernel Version |
|---|---|---|
| Scientific Linux* 7.2 X86_64 | N/A | 3.10.0-327.el7.x86_64 |
| SUSE* Linux* Enterprise Server (SLES*) 12.1 X86_64 | Service Pack 1 | 3.12.49-11.1-default |
| SUSE* Linux* Enterprise Server (SLES*) 12.2 X86_64 | Service Pack 2 | 4.4.21-69-default |

## 1.10    Parallel File Systems

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

- Intel$^®$ Enterprise Edition Lustre* (IEEL) 3.1
  - RHEL* versions supported by Intel$^®$ Omni-Path Software.
- IBM* General Parallel File System (GPFS) version 4.0.1
  - RHEL* 7.2.

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

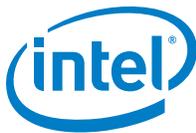## 1.11    Compilers

### 1.11.1    MPI

This release supports the following MPI implementations:

**Table 1-3.    MPI Compilers**

| MPI Implementation | Runs Over | Compiled With |
|---|---|---|
| Open MPI 1.10.4 | 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* 7.2 | gcc (GCC) 4.8.5 20150623 (Red Hat* 4.8.5-4) |
| (GNU) gcc | RHEL* 7.3 | gcc (GCC) 4.8.5 20150623 (Red Hat* 4.8.5-11) |
| (GNU) gcc | SLES* 12 SP 1 | gcc (SUSE* Linux*) version 4.8.5 |
| (GNU) gcc | SLES* 12 SP 2 | gcc (SUSE* Linux*) version 4.8.5 |
| (Intel) icc | RHEL* 7.2 | icc (ICC) 15.0.1 |
| (Intel) icc | RHEL* 7.3 | icc (ICC) 15.0.1 |
| (Intel) icc | SLES* 12 SP 1 | icc (ICC) 15.0.1 |
| (Intel) icc | SLES* 12 SP 2 | icc (ICC) 15.0.1 |

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

## 1.12    Hardware

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

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

**Table 1-5.    Hardware Supported**

| 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 |
| Next generation Intel® Xeon® Processor (codename Skylake) | Skylake CPU-based servers (pre-production samples) |
| Intel® Xeon Phi™ Processor x200 product family | Knights Landing CPU-based servers |
| Intel® Omni-Path Host Fabric Interface 100HFA016 (x16) | Single Port Host Fabric Interface (HFI) |
| Intel® Omni-Path Host Fabric Interface 100HFA018 (x8) | Single Port Host Fabric Interface (HFI) |
| Intel® Omni-Path Switch 100SWE48Q | Managed 48-port Edge Switch |
| Intel® Omni-Path Switch 100SWE48U | Externally-managed 48-port Edge Switch |
| Intel® Omni-Path Switch 100SWE24Q | Managed 24-port Edge Switch |
| Intel® Omni-Path Switch 100SWE24U | Externally-managed 24-port Edge Switch |
| Intel® Omni-Path Director Class Switch 100SWD24 | Director Class Switch 100 Series, up to 768 ports |
| Intel® Omni-Path Director Class Switch 100SWD06 | Director Class Switch 100 Series, up to 192 ports |

*Note:*    For RHEL* 6.7 and CentOS* 6.7, only the following processors are supported:
- Intel® Xeon® Processor E5-2600 v3 product family
- Intel® Xeon® Processor E5-2600 v4 product family

# 1.13 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
10.3.2 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 Installation Prerequisites for RHEL* 6.7 and CentOS* 6.7

Install the following packages using yum from the RHEL* or CentOS* distributions:

- libibverbs
- librdmacm
- libibcm
- qperf
- perftest
- rdma
- infinipath-psm
- opensm-devel
- expat
- elfutils-libelf-devel
- libstdc++-devel
- gcc-gfortran
- atlas
- c-ares
- tcl
- expect
- tcsh
- sysfsutils
- pciutils
- bc (command line calculator for floating point math)
- rpm-build
- redhat-rpm-config
- kernel-devel

- opensm-libs

### 1.13.2.2 Required Pre-Installation to perform external modules builds on SLES* 12 Systems

*Note:*   This step is required only if the installed distribution kernel has been updated to a distribution security update.

SLES* 12 kernel-development environment is not ready "out-of-box" for external modules build. It has to be prepared prior to the installation. To rebuild SLES* 12 kernel pieces, perform the following steps:

1. Change directory:
   `cd /lib/modules/3.12.28-*****/source`

2. Create the following files:
   `Make cloneconfig`
   `Make oldconfig`
   `Make __headers`

3. Only build as needed.

## 1.14 Product Constraints

None.

## 1.15 Product Limitations

This release has the following product limitations:

- 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.

- PA Failover should **not** be enabled with FMs running on differing software versions. PA Failover is enabled via
  `configuration:<PM>/<ImageUpdateInterval> > 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.

### 1.15.1 RHEL* 6.7 and CentOS* 6.7 Support

- Processor support:
  — Intel® Xeon® Processor E5-2600 v3 product family
  — Intel® Xeon® Processor E5-2600 v4 product family

- File system support:
  — GPFS
  — NFS
  — Lustre*

  *Note:* For Enterprise Edition 3.0 Clients (support RHEL* 6.7) and Enterprise Edition 3.1 Servers (support RHEL* 7.3): You cannot upgrade your Clients beyond version 3.0 until you move to a newer RHEL* version.

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

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

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

## 1.16    Token ID (TID) RDMA Information

Token ID (TID) 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 performance benefits include increased achievable bandwidth with reduced CPU utilization. The TID RDMA protocol accelerates the OpenFabrics Alliance* (OFA) Verbs API with no changes required to API consumers. The acceleration technique is performed by the host driver and the application running over the OFA Verbs API does not need to make any code change.

TID 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.

## 1.17    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
  www.intel.com/omnipath/SwitchPublications
- Intel® Omni-Path Fabric Software Installation, User, and Reference Guides
  www.intel.com/omnipath/FabricSoftwarePublications
- Drivers and Software (including Release Notes)
  www.intel.com/omnipath/downloads

**Table 1-6.  Related Documentation for this Release**

| Document Title |
| --- |
| **Hardware Documents** |
| Intel® Omni-Path Fabric Switches Hardware Installation Guide |
| Intel® Omni-Path Fabric Switches GUI User Guide |
| Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide |
| Intel® Omni-Path Edge Switch Platform Configuration Reference Guide |
| Intel® Omni-Path Fabric Switches Release Notes (includes managed and externally-managed switches) |
| Intel® Omni-Path Host Fabric Interface Installation Guide |
| **Fabric Software Documents** |
| Intel® Omni-Path Fabric Software Installation Guide |
| Intel® Omni-Path Fabric Suite Fabric Manager User Guide |
| Intel® Omni-Path Fabric Suite FastFabric User Guide |
| Intel® Omni-Path Fabric Host Software User Guide |
| Intel® Omni-Path Fabric Suite Fabric Manager GUI Online Help |
| Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide |
| Intel® Omni-Path Fabric Suite FastFabric Command Line Interface Reference Guide |
| Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide |
| Intel® Omni-Path Fabric Performance Tuning User Guide |
| Intel® Omni-Path Host Fabric Interface Platform Configuration Reference Guide |
| Intel® Omni-Path Fabric Software Release Notes |
| Intel® Omni-Path Fabric Manager GUI Release Notes |
| Intel® Omni-Path Storage Router Design Guide |
| Intel® Omni-Path Fabric Staging Guide |
| Building Lustre* Servers with Intel® Omni-Path Architecture Application Note |

# 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.    Issues resolved in this release**

| ID | Description | Resolved in Release |
|----|-------------|---------------------|
| 137950 | OFED DELTA 10.3.1 package fails to build on RHEL* 7.2 kernel version: kernel-devel-3.10.0-327.46.1.el7.x86_64 | 10.3.2 |
| 138475 | IntelOPA-IFS.RHEL72-x86_64.10.3.1.0.22 fails to build on RHEL* 7.2 kernel version: 3.10.0-327.49.2.el7.x86_64 | 10.3.2 |
| 139031 | IntelOPA-IFS.SLES12-x86_64.10.3.1.0.22 fails to build on SLES* 12 SP1 kernel version: 3.12.74-60.64.40 | 10.3.2 |

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

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

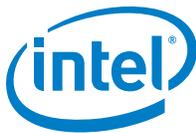| ID | Description | Resolved in Release |
|----|-------------|---------------------|
| 132219 | Server platforms running IFS 10.3.0 release (or Intel® OPA software delivered in certain Linux* OS distributions) and using integrated HFI for OPA (commonly known as "-F") may not support Active Optical Cables (AOC) after boot up. | 10.3.1 |
| 133377 | irqbalance settings are not being honored correctly after a reboot. | 10.3 |
| 133707 | Updating to the RHEL* 7.2 kernel for the CVE-2016-0728 update in OSes prior to 7.2 causes the Intel® Omni-Path installation to fail. | 10.3 |
| 134111 | On some older HFI and HFI-like cards, running hfi1_eprom -V -c to inquire the version of the AOC configuration file on the card may return an invalid version of "etnIRFWI". | 10.3 |
| 134124 | HFI port stuck in INIT state due to SM failure to set pkeys. | 10.3 |
| 134135 134429 | When running communication-intensive workloads with 10KB MTU, it is possible to encounter node and/or job failures. | 10.3 |
| 134283 | When downgrading on a SLES* 12.X system from Intel® OPA version 10.2.X to a previous version, the following install errors occur:<br>`ERROR - Failed to install`<br>and<br>`error: Failed dependencies: libibmad5 is needed by opa-basic-tools...` | 10.3 |
| 134772 | opatmmtool will fail if provided with a filename (full path) that is longer than 63 characters. | 10.3 |
| 134866 | `hostverify.sh` cannot properly detect if SRP is enabled on target node. | 10.3.1 |

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

| ID | Description | Resolved in Release |
|---|---|---|
| 134956 | ib0 fails to become ready on warm reboots. | 10.3.1 |
| 135000 | Fabric Manager configuration files that specify IncludeGroup fields with undefined or nonexistent device groups could cause Fabric Manager failure. | 10.3 |
| 135649 | The XPPSL kernel changes conflict with items in the SLES* 12 SP1 kernel RPM. This causes the recompile of the SLES* 12 SP1 compat-rdma package to have an error. | 10.3.1 |
| 135729 135870 | KNL-F/SKL-F ports are offline in pre-boot setting when connected with AOC. | 10.3.1 |
| 135812 | FM may crash and restart in the event of a failure during topology assignments. This may result in mismatched port physical states on a link. While unlikely, this event may occur when there are integrity issues on a link. | 10.3.1 |
| 135958 | Spurious segmentation faults with greater than 2MB PSM2 transfers on Intel® Xeon Phi™ platforms. | 10.3.1 |
| 136027 | IFS `hostverify.sh` script does not provide reliable results for pstates_on and governor tests on RHEL* 7.3 and SLES* 12 SP2. | 10.3.1 |
| 136028 | Two versions of the UEFI firmware are contained in the hfi-uefi RPM in the 10.3.0 IFS and BASIC packages. The files are functionally identical except the unsigned files (`HfiPcieGen3Loader_<version number>.unsigned.rom` and `HfiPcieGen3_<version number>.unsigned.efi`) are not signed for secure boot. | 10.3.1 |
| 136152 | Server platforms using integrated HFI for OPA (commonly known as "-F") require BIOS that provides UEFI version 1.3.1.0.0 and a configuration data file for pre-boot support of Active Optical Cables (AOC). Some servers may not have these files available in BIOS and will therefore not support AOC in pre-boot. | 10.3.1 |
| 136215 | For RHEL6.7, the `opaconfig` command will not change the autostart settings for OPA service. | 10.3.1 |
| 136318 | SM crashes showing segfault errors in logs and high CPU usage. These crashes were caused by a mismatch of pahistory file versions. | 10.3 |
| 136621 | PCIe Fatal Errors during reboot cycles on server platforms using integrated HFI for OPA (commonly known as "-F"). | 10.3.1 |
| 136628 | A bug in the Linux* kernel (CVE-2016-5195, also called Dirty COW) requires you to update the kernel for your operating system. | 10.3.1 |
| 136723 | Upgrading your OPA installation from version 10.2 to 10.3 may not install the correct host driver. | 10.3.1 |

## 2.3 Open Issues

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

**Table 2-3. Open Issues (Sheet 1 of 7)**

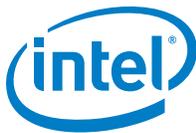| ID | Description | Workaround |
|---|---|---|
| 129563 | Memory allocation errors with Mvapich2-2.1/Verbs. | When running mvapich2 jobs with a large number of ranks (for example, between 36 and 72 ranks), you must set the following parameters in `/etc/security/limits.conf`:<br>* hard memlock unlimited<br>* soft memlock unlimited<br>Also, you must increase the lkey_table_size:LKEY table size in bits (2^n, 1 <= n <= 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*. |
| 130336 | `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. |
| 131017 | Verbs ib_send_bw, ib_read_bw, and ib_write_bw are not working with the -R option to use the RDMA CM API to create QPs and exch data. | The following combinations of client and server DO NOT allow RDMA CM connections:<br>Client       Server<br>SLES* 12.1    SLES* 12.1 or          (intermittent) RHEL* 7.2<br>Using SLES* clients results in "Unexpected CM event bl blka 0" errors.<br>Additionally, there are long (5-10 sec) initial delays when using these combinations:<br>Client       Server<br>SLES* 12.1    RHEL* 7.2<br>SLES* 12.0    RHEL* 7.2 |
| 131745 | When running OpenMPI 1.10.0 on SLES* 12 with large number of ranks per node (over 40), it may happen that the ORTE daemon (orted) "hangs" during the finalization of job. | Stopping and resuming the "hung" orted process allows the job to finish normally.<br>To find the hung process, run the ps and find a node with several job zombie processes.<br>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 | Kernel crash caused by the ib_srpt module. | Install this kernel patch:<br>https://git.kernel.org/cgit/linux/kernel/git/torvalds/linux.git/commit/?id=51093254bf879bc9ce96590400a87897c7498463 |
| 133596 | The install script does not check for packages that are necessary prerequisites for installation. | Refer to the *Intel® Omni-Path Fabric Software Installation Guide*, in the OS RPMs Installation Prerequisites section for the complete list. |
| 133604 | Bonding driver incorrectly shows hardware address of IPoIB interfaces. | Use the `opainfo` command to retrieve the PortGUID and `ip addr show ib0` to get the correct 20-byte hardware address of OPA network interface. |
| 133633 | OpenMPI and Mvapich2 compiles fail to link properly when using the Intel compilers. | No workaround available. |
| 134268 | The Option ROM image (e.g. containing a UEFI driver) may not be executed if the BIOS configures the HFI Expansion ROM BAR with an address that is not 16MB aligned. | Use a BIOS that configures the HFI Expansion ROM BAR with an address that is 16MB aligned.<br>Note that the main memory BAR of the device is 64MB in size and therefore requires 64MB alignment. A BIOS implementation that places the Expansion ROM BAR immediately after the main memory BAR automatically provides this workaround. Many BIOS implementations have this property and automatically meet the workaround criteria. |

**Table 2-3.    Open Issues (Sheet 2 of 7)**

| ID | Description | Workaround |
|---|---|---|
| 134353 | Very infrequently, when a link goes down, the logical link state can remain stuck in the 'Init' state. | The device containing the affected port must be rebooted in order to resolve the issue.  Ports in this state have a logical link state of 'Init' but do NOT have a physical port state of 'LinkUp'. |
| 134471 | The HFI UEFI driver cannot boot via PXE using Grub 2. | Use Elilo instead. |
| 134493 | When using Mvapich2 with Intel® Omni-Path PSM2, users will notice unexpected behavior when seeding the built-in random number generator with functions like srand or srandom before MPI_Init is called. MPI_Init re-seeds the random number generator with its own value and does not restore the seed set by the user application. This causes different MPI ranks to generate different sequences of random numbers even though they started with the same seed value. | Seed the random number generator after MPI_Init is called or use the reentrant random number generator functions such as drand48_r. |
| 134494 | Open MPI uses srand() family functions at MPI_Init() time. Therefore, if the user sets srand() before calling MPI_Init(), the values will be altered. | a) Fixed in Open MPI 2.0.1. <br> b) Call srand() functions family after calling MPI_Init(). |
| 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, working version of the EFI shell from the embedded shell. |
| 134821 | The UEFI network stack is initialized with a default network address before the driver receives a MAD packet containing an updated and actual subnet prefix. Therefore, in ARP and IP UEFI drivers the old (default) HW address is still used, causing problems with packet receiving and transmitting. | Use default subnet prefix, 0xfe80000000000000, when configuring a subnet for PXE boot over OPA. |
| 134904 | Legacy PXE boot using iPXE while the HFI UEFI driver is loaded causes a hang. | Configure PXE operation to boot using UEFI boot mode. |
| 135040 | You can't currently specify portions of an Intel® DCS chassis that is not populated and is not expected to be populated. If CoreFull is 1, all the internal links for that chassis are generated when run against `opaxlattopology`. If CoreFull is 0, none of the links are generated. | Copy internal configuration of desired Intel® DCS switch into the main topology tab of the spreadsheet. Then delete all lines corresponding to leafs or spines that are not present in the configuration. |
| 135068 | When PXE booting using older versions of Grub 2 over Ethernet while the HFI UEFI driver is loaded, some servers will crash with an RSOD (Red Screen of Death). | Upgrade to the latest version of Grub 2. |
| 135084 | During extensive power cycle testing, HFI adapter might fail to appear in PCI config space. | Reboot or power cycle the platform. |
| 135180 | OpenMPI/PSM2 timeouts during MPI stress tests on Haswell and Intel® Xeon® mixed fabrics. | Load the hfi1 module with the following parameter: <br> `sudo modprobe hfi1 rcvhdrcnt=16352` |
| 135259 | On rare occasions, the HFI links do not come up after a reboot. | Reboot or bounce the link. |
| 135326 | Calling `opasmaquery` fails when called from a non-SM node to a node which has not booted to the OS. | Use the SM node when calling `opasmaquery` in this way. |

**Table 2-3.     Open Issues (Sheet 3 of 7)**

| ID | Description | Workaround |
|---|---|---|
| 135355 | Due to changes in where the IFS packages are installed, customers using the FastFabric tools and upgrading to 10.3 from an earlier release must find each occurrence of `/opt/opa` in the `opafastfabric.conf` file and replace the string with `/usr/lib/opa`. | This includes:<br>`export FF_MPI_APPS_DIR=${FF_MPI_APPS_DIR:-/opt/opa/src/mpi_apps}`<br>`export FF_SHMEM_APPS_DIR=${FF_SHMEM_APPS_DIR:-/opt/opa/src/shmem_apps}`<br>`export FF_PRODUCT=${FF_PRODUCT:-IntelOPA-Basic.`cat`<br>`/opt/opa/tools/osid_wrapper`}`<br><br>These lines should be changed to:<br>`export FF_MPI_APPS_DIR=${FF_MPI_APPS_DIR:-/usr/lib/opa/src/mpi_apps}`<br>`export FF_SHMEM_APPS_DIR=${FF_SHMEM_APPS_DIR:-/usr/lib/opa/src/shmem_apps}`<br>`export FF_PRODUCT=${FF_PRODUCT:-IntelOPA-Basic.`cat`<br>`/usr/lib/opa/tools/osid_wrapper`}` |
| 135390 | For certain older HFI adapters, the card may not be flashed with the AOC configuration file, or may be flashed with an older version of the AOC configuration file. With release 10.3, these adapters will fail to link up with AOCs, and these messages may be seen in dmesg:<br>`[   26.903186] hfi1 0000:d5:00.0: hfi1_0: parse_platform_config:Bad config file`<br>`[   26.903186] hfi1 0000:d5:00.0: hfi1_0: parse_platform_config:File claims to be larger than read size`<br>In addition to:<br>`[   27.351555] hfi1 0000:d5:00.0: hfi1_0: tune_serdes: Unknown port type` | For HFI and HFI-equivalent customers, the resolution for this issue is to flash the AOC configuration file provided in `/lib/firmware/updates/hfi1_platform.dat` to the card using the following command line:<br>`# hfi1_eprom -w -c /lib/firmware/updates/hfi1_platform.dat -o /opt/opa/bios_images/HfiPcieGen3Loader_<version>.rom -b /opt/opa/bios_images/HfiPcieGen3_<version>.efi -d /sys/bus/pci/devices/<PCI address>/resource0` |
| 135545 | A change has been made to several SA record attributes which causes incompatibilities between the Fabric tool suite and the SA. | You must update both the SA and the Fabric tool suite at the same time to avoid this incompatibility. |
| 135648 | MPI applications are installed under the `/usr/lib` directory structure, which may be set up to be read-only overall. This causes resulting FastFabric operations to fail since mpi_apps contain source code and run scripts for sample MPI applications, test programs and benchmarks. | If you want to build them via the "Rebuild MPI Library and Tools" option in the `opafastfabric` menu, you must first perform the following steps to copy the files out of `/usr/lib` that may be a read-only directory:<br>1) Create a mirror directory named `mpi_apps` in a writable area, such as under $HOME.<br>`  mkdir $HOME/mpi_apps`<br>2) Copy `/usr/lib/opa/src/mpi_apps/*` to the mirror directory.<br>`  cp -r /usr/lib/opa/src/mpi_apps/* $HOME/mpi_apps`<br>3) Edit `/etc/sysconfig/opa/opafastfabric.conf` and change the setting of `FF_MPI_APPS_DIR` to the new mirror directory.<br>`  export FF_MPI_APPS_DIR=${FF_MPI_APPS_DIR:-$HOME/mpi_apps}`<br>You can now execute the "Rebuild MPI Library and Tools" option safely. |

**Table 2-3.    Open Issues (Sheet 4 of 7)**

| ID | Description | Workaround |
|---|---|---|
| 135711 | After generating the `opafm.xml` file from the `config_generate` script, the FE is not enabled. | None. |
| 135873 | hostverify.sh fails with RHEL* 6.7 due to the Intel P-State driver not being the default cpufreq driver. | Item 1 in the Documentation Errata section of this document describes how to set the Intel P-State driver as the default. |
| 135929 | Intel® Omni-Path Boot nodes occasionally dropped from fabric when switching master SM from one node to another. | Reboot PXE client node. |
| 135951 | When creating host verify punchlist, the following error message is displayed:<br>`unable to parse filter -s Invalid slot number` | None. |
| 135963 | Cannot install IFS software on RHEL* 7.3 using the command: `./INSTALL -vv -a` | Use the `-v` option instead. |
| 135975 | After performing an OPA software configuration update, some unmanaged switches do not update the settings for LinkWidth and LinkWidthDnGrade enables. | A reboot is required for configuration changes made to an externally managed switch to become active. |
| 136049 | The expected width of a card is not showing up correctly in opaverifyhosts. | For a cluster with mixed server or HFI configurations, the correct edited hostverify.sh script should be pushed to each group of servers.<br>If using the TUI:<br>• Create a /etc/sysconfig/opa/myhosts file for each type of server configuration. For example: computehosts, storagehosts, mgmthosts, etc.<br>• Pick the desired hosts file in menu item 0 of the "Host Verification/Admin" menu, then run the "Perform Single Host Verification" function.<br>• Edit the sample hostverify.sh script, putting in the proper settings for the server config (HFI PCIe bus, server memory size, expected single node HPL performance for server, etc).<br>• When prompted, run the hostverify function on the given subset.<br>• Repeat for each of the hosts files. |
| 136137 | The hfi1_eprom tool man page contains incorrect information in the -d device option. | The following text is correct in the man page:<br>-d device<br>Specify the device file to use. The default method which uses the resource0 file will attempt to autodetect the first device. This can be changed by pointing to the correct file found in /sys/bus/pci/devices/XXXXX/resource0<br>The incorrect text shown below will be removed in a future release:<br>For the alternative driver method the device file should be /dev/hfi1_N, where N represents the desired hfi. Note for cards with multiple HFIs, any device can be used as there is only one EPROM per card. |
| 136160 | On some Intel® Xeon Phi™ with integrated Intel® Omni-Path fabric platforms, the second integrated HFI is discovered first and is subsequently identified as the first HFI device. As a result, when issuing Intel® Omni-Path commands, the second HFI appears first in the results.<br>In Linux* and various Intel® Omni-Path tools, the HFI reporting order may be the opposite of the order appearing on the Intel® Xeon Phi™ with integrated Intel® Omni-Path fabric cable/faceplate. | You can identify the second integrated HFI by inspecting the Node GUID or Port GUID/Port GID reported by `opainfo` or other commands such as `hfi1_control -i`. Note that bit 39 of the PortGUID, the most significant bit, is set for the second HFI, and is clear for the first HFI.<br>Keep in mind that when issuing various Intel® Omni-Path CLI commands targeted at a specific HFI using the `-h` option, `-h 1` correlates to the device that is listed as `hfi1_0`. As a result, the issued command affects the second HFI instance in cases where the second HFI port instance appears first. |
| 136432 | Certain perftest tools such as ib_write_bw do not work on RHEL* 7.3 when using the RDMA CM with UD QPs. | Roll back the perftest package to the level found in RHEL* 7.2, which is perftest-2.4. Then install this package on RHEL* 7.3. |

**Table 2-3.    Open Issues (Sheet 5 of 7)**

| ID | Description | Workaround |
|---|---|---|
| 136436 | On SLES* 12.2, node_desc is not populated with the host name when system is booted up. | Install and run the rdma-ndd daemon on each node.<br>1.  Unpack IFS:<br>`# tar xzf IntelOPA-IFS.SLES122-x86_64.10.3.0.0.81.tgz`<br>`# ls`<br>`IntelOPA-IFS.SLES122-x86_64.10.3.0.0.81`<br>`IntelOPA-IFS.SLES122-x86_64.10.3.0.0.81.tgz`<br>2.  Uninstall infiniband-diags and libibnetdisc5 libraries. (SLES* splits out the libibnetdisc library but it is included in the IFS infiniband-diags version.)<br>`# rpm -e infiniband-diags`<br>`# rpm -e libibnetdisc5`<br>3.  Install the older version of infiniband-diags from the IFS package.<br>`# cd IntelOPA-OFED_DELTA.SLES122-x86_64.10.3.0.0.82/`<br>`# rpm -Uvh ./infiniband-diags-1.6.7-2.x86_64.rpm`<br>4.  Enable rdma-ndd:<br>`# systemctl daemon-reload`<br>`# systemctl status rdma-ndd`<br>`  rdma-ndd.service - RDMA Node Description Daemon`<br>`   Loaded: loaded (/usr/lib/systemd/system/rdma-ndd.service; disabled; vendor preset: disabled)`<br>`   Active: inactive (dead)`<br>`# systemctl enable rdma-ndd`<br>`Created symlink from /etc/systemd/system/multi-user.target.wants/rdma-ndd.service to /usr/lib/systemd/system/rdma-ndd.service.`<br>5.  Start rdma-ndd and check the status:<br>`# systemctl start rdma-ndd`<br>`# systemctl status rdma-ndd`<br>6.  Test that it is working:<br>`# cat /sys/class/infiniband/hfi1_0/node_desc`<br>`phs1fnive08u26 hfi1_0`<br>`# hostname foo`<br>`# cat /sys/class/infiniband/hfi1_0/node_desc`<br>`foo hfi1_0`<br>`# hostname phs1fnive08u26`<br>`# cat /sys/class/infiniband/hfi1_0/node_desc`<br>`phs1fnive08u26 hfi1_0`<br>`# reboot`<br>`...`<br>`# cat /sys/class/infiniband/hfi1_0/node_desc`<br>`phs1fnive08u26 hfi1_0` |
| 136437 | When using RHEL* 7.2, the default generic PXE boot image does not work due to missing driver and firmware files. | None. |
| 136500 | RDMA perftests can hang on start on a client side when RDMA CM (-R option) is used. | None. |

## Table 2-3.    Open Issues (Sheet 6 of 7)

| ID | Description | Workaround |
|---|---|---|
| 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.)<br><br>In Release 10.3.1 a fix was included to address the situation that occurs in fabrics of >1000 nodes when numerous links bounce (or hosts are rebooted) at once. | The following workarounds are recommended:<br>• When rebooting nodes on a production cluster, perform reboots in batches of 300 nodes or less.<br>• During cluster deployment, carefully follow the procedures in the *Intel® Omni-Path Fabric Staging Guide* and use FastFabric to check signal integrity and placement of all cables. Correct or disable any problematic links before starting production use of the cluster.<br>• When replacing or expanding a production cluster, repeat the procedures in the *Intel® Omni-Path Fabric Staging Guide* to verify the new hardware. Correct or disable any problematic links before resuming production use of the cluster.<br>• Use the PM, FM logs, FM GUI, FastFabric, and other tools to monitor signal integrity and link stability. Correct or disable any problematic links when discovered. |
| 136733 | Slow memory deregistration has been observed. | None. |
| 136822 | AOC support is not available on integrated HFI platforms (-F platforms) if the Intel UEFI driver is not executed during boot. Some BIOS will not execute the UEFI driver in Legacy BIOS boot mode. Also, some BIOS configuration settings or other system settings will bypass execution of the HFI driver. | Avoid the use of Legacy BIOS boot mode if your platform does not execute the HFI driver in that mode.<br>Avoid BIOS settings or other configuration settings that do not execute the HFI driver during boot. |
| 136901 | Occasionally, pre-boot nodes are dropped by the Fabric Manager during fabric sweeps, where the system containing the dropped pre-boot node has more than one HFI on a single socket. | Bounce the link of the dropped pre-boot port. |
| 136902 | A snapshot file with a multicast group with rate 10g will not be read properly.<br>The following error is returned:<br>opafabricanalysis: Port 0:0 Error: Unable to analyze fabric snapshot. See /var/usr/lib/opa/analysis/latest/fabric.0:0.links.stderr<br>opafabricanalysis: Possible fabric errors or changes found | On all nodes running opafm, run:<br>`systemctl stop opafm`<br>On all switches running ESM, run:<br>`smControl stop`<br>For all nodes/servers running ibacm:<br>1. Create a file /etc/rdma/ibacm_opts.cfg with one line:<br>`min_rate 40`<br>2. Restart ibacm<br><br>In the nodes running the host FM, restart opafm or start opafm with the command:<br>`systemctl start opafm`<br>In the switches running ESM, run:<br>`smControl start` |
| 136945 | When using the TID RDMA feature, certain Mvapich over Verbs tests may cause error messages. | None. |
| 136971 | When using the TID RDMA feature, certain Verbs Multi-PPN tests may cause error messages. | None. |
| 136985 | opahfirev has output errors when the HFI driver is not installed. | None. |
| 136995 | The opahfirev tool output uses the term "HWRev" to indicate the revision of the silicon on the card. | None. |
| 137015 | The state and configuration of ipoib interfaces are controlled by the NetworkManager service. The NetworkManager in RHEL* 7.2 mistakenly assumes the ipoib interface is type 'ethernet' and fails to initialize it, due to a mismatch against its actual type which is 'infiniband'. | Add "TYPE=InfiniBand" in the ifcfg-ib* files to ensure correct NetworkManager configuration.<br>In order for ifcfg-ib* changes to take effect, you must reboot the host or run the following command as root:<br>`nmcli connection reload` |

## Table 2-3. Open Issues (Sheet 7 of 7)

| ID | Description | Workaround |
|---|---|---|
| 137054 | Pinging an Intel® OPA UEFI permanent IP address from a DHCP server fails on subsequent reboots unless the corresponding network interface has first been initialized in the UEFI network stack. | Before pinging a UEFI permanent IP address, first initialize the corresponding network interface in the UEFI network stack. |
| 137096 | The IFS package does not install all the RPMs that it contains. In particular, infiniband-diags and libibmad are not automatically installed.<br>The absence of infiniband-diags may result in failure of node descriptions to be populated, such that all hosts have the same `hfi1_0` description. | Manually install the infiniband-diags and libibmad packages. |
| 137106 | When running SLES* 12.2 with typical OS drivers installed and connected with copper QSFP, the state does not change from "Offline" to "Physical Linkup (Init)" as expected. | Add a `platform.dat` file in `/lib/firmware/updates`, then restart. |
| 137108 | When using the TID RDMA feature, virtual machines, and other cases where the IOMMU is enabled, do not operate correctly. This can lead to stability issues, and possibly data corruption, because the address used to receive data into will be incorrect. | None. |
| 137142 | When using the TID RDMA feature, certain MPI benchmark tests may cause Kernel panic. | None. |
| 137212 | The RHEL* 6.7 base version of the perftest package includes a ib_send_lat utility that may cause a segmentation fault when run with the -z option. | Run the utility without using the -z option. |
| 137221 | Querying for switch info with opasmaquery while using the -g option will print incorrect IPv4 addresses. | Do not use the -g option. |

# 3.0    Documentation Errata

This section describes issues in the user documentation that are relevant to this release. The documentation updates will be made in a future release.

**1.    Include hostverify changes in performance guide**

Issue:          136684

Document:       *Intel® Omni-Path Fabric Performance Tuning User Guide*

New Text:       A new section will be added to the document as follows.

### 3.2.3   Switching to the Intel P-State Driver to Run Certain FastFabric Tools

Some Intel-provided tools require the use of the Intel P-State driver rather than the acpi_cpufreq driver. For example, the `hostverify.sh` tool fails with RHEL* 6.7 due to the Intel P-State driver not being the default cpufreq driver.

If you are using the acpi_cpufreq driver, perform the following steps to temporarily switch to the Intel P-state driver in order to use the target tool.

**Temporary Switch to Intel P-State Driver**

To temporarily switch to the Intel P-state driver, perform the following steps:

1. Make sure cpupowerutils package is installed.
   ```
   # yum install cpupowerutils
   ```
2. Check if any other cpufreq kernel driver is active.
   ```
   # cpupower frequency-info -d
   ```
3. Unload another cpufreq kernel driver (if any).
   ```
   # rmmod acpi_cpufreq
   ```
4. Load intel_pstate driver.
   ```
   # modprobe intel_pstate
   ```
5. Set cpufreq governor to 'performance'.
   ```
   # cpupower -c all frequency-set -g performance
   ```
6. After using `hostverify.sh` or other tools that needed the Intel P-state set, you may reboot to return to the acpi_cpufreq driver.

**Load Intel P-State Driver at Boot Time**

To load the Intel P-state driver at boot time, perform the following steps:

1. Create a script file `/etc/sysconfig/modules/intel_pstate.modules` and add the below text to it.
   ```
   #!/bin/sh
   /sbin/modprobe intel_pstate >/dev/null 2>&1
   ```
2. Add executable permissions for the file:
   ```
   # chmod +x /etc/sysconfig/modules/intel_pstate.modules
   ```

3. Reboot the system for the changes to take effect.

4. Verify that the Intel P-state driver is loaded.

5. Install the cpupowerutils package, if not already installed:
```
# yum install cpupowerutils
```

6. Set cpufreq governor to 'performance' using the command:
```
# cpupower -c all frequency-set -g performance
```

To re-enable the acpi_cpufreq driver, perform the following:

1. Disable intel_pstate in the kernel command line:
   Edit /etc/default/grub by adding `intel_pstate=disable` to GRUB_CMDLINE_LINUX.
   For example:
```
GRUB_CMDLINE_LINUX=vconsole.keymap=us console=tty0
vconsole.font=latarcyrheb-sun16 crashkernel=256M
console=ttyS0,115200 intel_pstate=disable
```

2. Apply the change using:
```
        if [ -e /boot/efi/EFI/redhat/grub.cfg ]; then
    GRUB_CFG=/boot/efi/EFI/redhat/grub.cfg
else if [ -e /boot/grub2/grub.cfg ]; then
    GRUB_CFG=/boot/grub2/grub.cfg
grub2-mkconfig -o $GRUB_CFG
```

3. Reboot.
   When the system comes back up with `intel_pstate` disabled, the acpi_cpufreq driver is loaded.
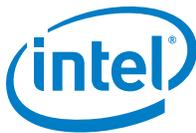
## 2. Update documentation with ESM limitations

Issue: 136895

Document: *Intel® Omni-Path Fabric Suite Fabric Manager User Guide*

New Text: The following section will be updated with new text and a table (see changebars).

### 1.4.3 Choosing Between Host and Embedded FM Deployments

Refer to the *Intel® Omni-Path Fabric Software Release Notes* or embedded firmware for more guidance in choosing between deploying a host or embedded FM solution.

**Table 1. HSM versus ESM Major Capability Differences**

| Capability | HSM | ESM |
|---|---|---|
| Maximum Fabric Node Size | See *Intel® Omni-Path Fabric Software Release Notes* for value. | See *Intel® Omni-Path Fabric Software Release Notes* for value. |
| PA Short Term History | Supported | Not Supported - disabled |
| Dynamic Configuration | Supported | Not Supported - disabled |
| FM Control User Interface | opafmcmd CLI | Chassis CLI |
| FE OOB Security | | |
| File directory location | Configurable | /mmc0:4 |
| Ciphers Support | • ECDHE-ECDSA-AES128-GCMSHA256<br>• ECDHE-ECDSA-AES128-SHA256<br>• DHE-DSS-AES256-SHA | DHE-DSS-AES256-SHA |
| Pre-defined Topology Verification | Supported | Not Supported - disabled |

#### Host-Based Fabric Manager or Embedded Fabric Manager Recommendations

Both fabric managers provide full functionality and the host-based fabric manager can be used in any scenario. The embedded fabric manager is implemented on lightweight internal hardware for smaller fabrics. As a result, the following recommendations exist.

- Director Class Switch configurations - A host-based subnet manager is required. host-based fabric managers run on a minimally configured server. For redundancy, two or more servers can run the host fabric manager simultaneously.

- Managed Edge switches - For fabrics of more than 100 nodes, which typically occur with three or more 48-port switches in a fabric, a host-based fabric manager is suggested.

| | Host-Based Fabric Manager | Embedded Fabric Manager |
|---|---|---|
| HFIs back-to-back | Yes | Not Applicable |
| 48-port Edge switches - less than 100 nodes | Yes | Yes |
| 48-port Edge switches - greater than 100 nodes | Yes | No |
| 192 or 768 Director Class switch | Yes | No |
| Director Class switch + Edge switches | Yes | No |

§ §