Cornelis™ Omni-Path Express™ Fabric Software

Release Notes for V10.11.1

Table of Contents

1. Overview of the Release

This document provides a brief overview of the changes introduced into the Cornelis Omni-Path Express Software in 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.

To better represent the Cornelis Networks product line, Cornelis Networks is changing the name of Omni-Path Architecture (OPA) to Omni-Path Express (OPX) and the abbreviation IFS to Omni-Path Express Suite (OPXS). For more information see Section 3.1 "Omni-Path Express Rebranding".

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

The software provided is 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 Cornelis Omni-Path Express products is available 24 hours a day, 365 days a year. For additional details, visit the Cornelis Support page at http://www.cornelisnetworks.com/support, or contact Cornelis Networks Customer Support at support@cornelisnetworks.com.

1.4. Updated Features and Middleware

The following table lists the updated features and middleware supported in this release.

Feature/Middleware	First Introduced
NVIDIA CUDA 11.5	10.11.1
NVIDIA CUDA 11.6 (RHEL 8.5 only)	10.11.1
Open MPI version 4.0.6	10.11.1
libfabric OFI version 1.13.1	10.11.1
Intel OneAPI 2021 update 3	10.11.1

1.5. Product Improvements

The following improvements are included in this release:

- Support for Red Hat Enterprise Linux (RHEL) 7.7, 7.8, 7.9, 8.3, 8.4 and 8.5 (See Operating Systems)

- Support for SUSE Linux Enterprise Server (SLES) 12.4, 12.5, 15.2, and 15.3 (See Operating Systems)

- Support for Rocky Linux 8.4 and 8.5 (See Operating Systems)

- Support for libfabric OFI version 1.13.1

- Support for Intel OneAPI 2021, Update 3

- Support for Open MPI version 4.0.6

- Support for PSM2 NCCL

- Support for NVIDIA CUDA 11.5, and 11.6 (RHEL 8.5 only)

- Enhanced Cable Support with update to HFI1 Firmware version 4.2.5

- Name changes to reflect the Cornelis product line.

1.6. Changed, Deprecated, and Removed Features

- Support for RHEL 8.2 and SLES 15.1 has been removed in the 10.11.1 release.

- As of 10.10.3.1, OFI Verbs (libfabric-verbs) is installed but no longer supported. OFA Verbs remains a supported component of Cornelis Omni-Path Express.

- In releases earlier than 10.10.2, OpenMPI with CUDA support was installed with the non-CUDA version of IFS using INSTALL command (without `-G` argument).

  In 10.10.2 and later, to install the CUDA version of IFS using the INSTALL command, you must use the `-G` argument.

- As of 10.10.1, CongDiscards counters are no 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 Cornelis Omni-Path Express Fabric Software packages:

- **BASIC** for compute, service, and login nodes

  Includes:

  – Software that installs the following packages:

  > **NOTE**
  >
  > These packages will replace or update the packages included in the OS distribution.

- hfi1-firmware, libpsm2 (for RHEL) and libpsm2-2 (for SLES), hfi1-diagtools-sw

- Open MPI and MVAPICH2. See Section 1.12 "MPI Libraries" for details.

- mpitests

- mpi-selector

- Open Fabrics Interface (OFI) libfabric

- **OPXS** (formerly called IFS) for management nodes

  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 OPX 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 Cornelis Omni-Path Express Fabric Software supports the operating systems listed in the following table.

> **NOTE**
>
> RHEL 7.9 is supported through RHEL 7.8 OPA packages.

Table 1. Supported Operating Systems

Operating System	Update/SP	Base Kernel Version
Red Hat Enterprise Linux (RHEL) 7.7 X86_64	Update 7	3.10.0-1062.el7.x86_64
Red Hat Enterprise Linux (RHEL) 7.8 X86_64	Update 8	3.10.0-1127.el7.x86_64
Red Hat Enterprise Linux (RHEL) 7.9 X86_64	Update 9	3.10.0-1160.el7.x86_64
Red Hat Enterprise Linux (RHEL) 8.3 X86_64	Update 3	4.18.0-240.el8.x86_64
Red Hat Enterprise Linux (RHEL) 8.4 X86_64	Update 4	4.18.0-305.el8.x86_64
Red Hat Enterprise Linux (RHEL) 8.5 X86_64	Update 5	4.18.0-348.el8.x86_64
CentOS-7 (1908) X86_64 (corresponds to RHEL 7.7)	1908	3.10.0-1062.el7.x86_64
CentOS-7 (2003) X86_64 (corresponds to RHEL 7.8)	2003	3.10.0-1127.el7.x86_64
CentOS-7 (2009) X86_64 (corresponds to RHEL 7.9)	2009	3.10.0-1160.el7.x86_64
CentOS-8 (2011) X86_64 (corresponds to RHEL 8.3)	2011	4.18.0-240.el8.x86_64

Operating System	Update/SP	Base Kernel Version
CentOS-8 (2105) X86_64 (corresponds to RHEL 8.4)	2105	4.18.0-305.el8.x86_64
CentOS-8 (2111) X86_64 (corresponds to RHEL 8.5)	2111	4.18.0-348.el8.x86_64
Rocky Linux 8.4 (corresponds to RHEL 8.4)	Update 4	4.18.0-305.el8.x86_64
Rocky Linux 8.5 (corresponds to RHEL 8.5)	Update 5	4.18.0-348.el8.x86_64
Scientific Linux 7.7 X86_64	Update 7	3.10.0-1062.el7.x86_64
Scientific Linux 7.8 X86_64	Update 8	3.10.0-1127.el7.x86_64
Scientific Linux 7.9 X86_64	Update 9	3.10.0-1160.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) 15.2 X86_64	Service Pack 2	5.3.18-22_default
SUSE Linux Enterprise Server (SLES) 15.3 X86_64	Service Pack 3	5.3.18-57-default

1.10. CUDA Support

In general, PSM2 GPUDirect RDMA with CUDA is supported as shown below:

- CUDA Toolkit 11.6 is supported on RHEL 8.5
- CUDA Toolkit 11.5 is supported on RHEL 7.7, RHEL 7.8, RHEL 7.9, RHEL 8.3, RHEL 8.4, SLES 15 SP2, SLES 15 SP3
- CUDA Toolkit 10.2 is supported on SLES12 SP4, SLES12 SP5

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 Cornelis Omni-Path Express Fabric Software contains the hfi1 driver kernel modules listed in the following table.

> **NOTE**
>
> The version number is shown in bold.

Table 2. Kernel Modules for hfi1 Driver

OS	Non-CUDA/ CUDA	RPM Version
RHEL 7.7	Non-CUDA	kmod-ifs-kernel-updates-3.10.0_1062.el7.x86_64-**2351**.x86_64.rpm
	CUDA	kmod-ifs-kernel-updates-3.10.0_1062.el7.x86_64-**2351cuda**.x86_64.rpm
RHEL 7.8	Non-CUDA	kmod-ifs-kernel-updates-3.10.0_1127.el7.x86_64-**2351**.x86_64.rpm
	CUDA	kmod-ifs-kernel-updates-3.10.0_1127.el7.x86_64-**2351cuda**.x86_64.rpm
RHEL 7.9 (built during installation)	Non-CUDA	kmod-ifs-kernel-updates-3.10.0_1160.el7.x86_64-**2351**.x86_64.rpm
	CUDA	kmod-ifs-kernel-updates-3.10.0_1160.el7.x86_64-**2351cuda**.x86_64.rpm

OS	Non-CUDA/CUDA	RPM Version
RHEL 8.3	Non-CUDA	kmod-ifs-kernel-updates-4.18.0_240.el8.x86_64-**2351**.x86_64.rpm
	CUDA	kmod-ifs-kernel-updates-4.18.0_240.el8.x86_64-**2351cuda**.x86_64.rpm
RHEL 8.4	Non-CUDA	kmod-ifs-kernel-updates-4.18.0_305.el8.x86_64-**2351**.x86_64.rpm
	CUDA	kmod-ifs-kernel-updates-4.18.0_305.el8.x86_64-**2351cuda**.x86_64.rpm
RHEL 8.5	Non-CUDA	kmod-ifs-kernel-updates-4.18.0-348.el8.x86_64-**2351**.x86_64.rpm
	CUDA	kmod-ifs-kernel-updates-4.18.0-348.el8.x86_64-**2351cuda**.x86_64.rpm
SLES 12.4	Non-CUDA	ifs-kernel-updates-kmp-default-4.12.14_94.41_default_k4.12.14_94.41-**2351**.x86_64.rpm
	CUDA	ifs-kernel-updates-kmp-default-4.12.14_94.41_default_k4.12.14_94.41-**2351cuda**.x86_64.rpm
SLES 12.5	Non-CUDA	ifs-kernel-updates-kmp-default-4.12.14_120_default_k4.12.14_120-**2351**.x86_64.rpm
	CUDA	ifs-kernel-updates-kmp-default-4.12.14_120_default_k4.12.14_120-**2351cuda**.x86_64.rpm
SLES 15.2	Non-CUDA	ifs-kernel-updates-kmp-default-5.3.18-57-default_k5.3.18_22-**2351**.x86_64.rpm
	CUDA	ifs-kernel-updates-kmp-default-5.3.18-57-default_k5.3.18_22-**2351cuda**.x86_64.rpm
SLES 15.3	Non-CUDA	ifs-kernel-updates-kmp-default-5.3.18_22.default_k5.3.18_22-**2351**.x86_64.rpm
	CUDA	ifs-kernel-updates-kmp-default-5.3.18_22.default_k5.3.18_22-**2351cuda**.x86_64.rpm

1.12. MPI Libraries

NOTE

Cornelis recommends that you source the Intel MPI mpivars.sh or setvars.sh script when using Intel MPI with OPX and OFI. Refer to the Intel MPI Library documentation (https://software.intel.com/content/www/us/en/develop/tools/oneapi/components/mpi-library.html)

If you are not using Intel MPI, do not run this script.

1.12.1. Supported MPI Libraries

The table below lists the different MPI libraries supported by Cornelis Omni-Path Express 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 Cornelis Omni-Path Express Fabric Software package.

NOTE

When using newer releases of Open MPI with PSM2 MTL or OFI MTL there are new performance limitations to be aware of. The 5.0.x builds of Open MPI have removed support for sending "one sided communications" via the Open MPI MTL (Message Transport Layer) and replaced it with using the OFI BTL (Basic Transport Layer) instead.

While this yields some performance improvements in Open MPI it also consumes twice the number of hardware resources on Omni-Path (and other RDMA) hardware. This can have the effect of reducing the total number of ranks-per-node that Open MPI can support on an Omni-Path fabric.

For this reason Cornelis recommends using the 4.0.6 or 4.0.7 releases of Open MPI on Omni-Path fabrics. If a 5.0.x version of Open MPI is required for a particular application, and is using one-sided communications, the user will need to restrict the maximum number of ranks-per-node their application uses to about 1/2 the number of CPU cores available on the nodes. (For example, if a particular fabric features 72 cores per node only 36 will be usable.)

Alternately, the system administrator can update the HFI1 module settings to increase the number of contexts reserved for users by increasing the `num_user_contexts` parameter. This can be found in the `/etc/modprobe.d/hfi1.conf`, for example, but may have other performance penalties. See the *Cornelis Omni-Path Express Host Fabric Interface Installation Guide* for more information.

For 4.1.x versions of Open MPI, the ability to use the MTL for one-sided communications is still present but may no longer be selected by default. For this reason, when running with Open MPI versions 4.1.x we recommend using the settings:

`--mca mtl ofi --mca osc pt2pt --mca btl self,vader --mca pml ^ucx`

(`--mca pml ^ucx` can be omitted if ucx support was not compiled into Open MPI.)

NOTE

As of 10.10.0, the Open MPI build in IFS, renamed OPXS with this 10.11.1.0 release, no longer includes native verbs support (openib BTL).

Table 3. Supported MPI Libraries

MPI Implementation	Included in Basic Package?	Runs Over	Compiled With
Open MPI 4.0.6, 4.0.7, 4.1.1,4.1.2[1]	Yes	PSM2, OFI	GCC
Open MPI 4.0.6-cuda[1]	Yes	PSM2	GCC
MVAPICH2-2.3B	Yes	PSM2	GCC
Intel OneAPI 2021 Update 3, 4 [2]	No	OFI	N/A
NOTE:			
1. Open MPI 4.1.0 is not supported at this time.			
2. PSM2 is compiled using the Intel Compiler (ICC).			

1.12.2. 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 4. Compiler Versions and Distributions

Compiler	OS Distribution	Compiler Version
(GNU) gcc	RHEL 7.7	gcc (GCC) 4.8.5 20150623 (Red Hat 4.8.5-39)
(GNU) gcc	RHEL 7.8	gcc (GCC) 4.8.5 20150623 (Red Hat 4.8.5-39)
(GNU) gcc	RHEL 7.9	gcc (GCC) 4.8.5 20150623 (Red Hat 4.8.5-44)
(GNU) gcc	RHEL 8.3	gcc (GCC) 8.3.1 20191121 (Red Hat 8.3.1-5)
(GNU) gcc	RHEL 8.4	gcc (GCC) 8.4.1 20200928 (Red Hat 8.4.1-1)
(GNU) gcc	RHEL 8.5	gcc (GCC) 8.5.0 20210514 (Red Hat 8.5.0-3)
(GNU) gcc	SLES 12 SP4	gcc version 4.8.5 (SUSE Linux)
(GNU) gcc	SLES 12 SP5	gcc version 4.8.5 (SUSE Linux)
(GNU) gcc	SLES 15 SP2	gcc version 7.5.0 (SUSE Linux)
(GNU) gcc	SLES 15 SP3	gcc version 7.5.0 (SUSE Linux)

NOTE

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

1.13. Supported Hardware

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

NOTE

The PSM2 implementation has a limit of four (4) HFIs per server.

Table 5. Supported Hardware

Hardware	Description
Cornelis Omni-Path Express Host Fabric Interface 100HFA016 (x16)	Single Port Host Fabric Interface (HFI)
Cornelis Omni-Path Express Host Fabric Interface 100HFA018 (x8)	Single Port Host Fabric Interface (HFI)
2nd Gen AMD EPYC Processor	Rome Zen 2 microarchitecture
3rd Gen AMD EPYC Processor	Milan Zen 3 microarchitecture
Intel Xeon Processor E5-2600 v3 product family	Haswell CPU-based servers
Intel Xeon Processor E5-2600 v4 product family	Broadwell CPU-based servers
Intel Xeon Scalable Processor	Skylake CPU-based servers
2nd Generation Intel Xeon Scalable Processor	Cascade Lake CPU-based servers
3rd Generation Intel Xeon Scalable Processor (i3, i5, i7)	Ice Lake CPU-based servers

1.14. Switch Firmware

The following firmware is supported for Cornelis Omni-Path Express switches:

- Cornelis Omni-Path Express Switch Firmware 10.8.x revision (managed and externally-managed switches)

- Cornelis Omni-Path Express Switch Firmware 10.7.x revision (managed and externally-managed switches)

Refer to the *Cornelis Omni-Path Express Fabric Switches Release Notes* for more information.

1.15. Document Versions

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

Table 6. Supported Document Versions

Title	Doc. Number	Revision
Cornelis Omni-Path Express Fabric Quick Start Guide	J57479	9.0
Cornelis Omni-Path Express Fabric Setup Guide	J27600	14.0
Cornelis Omni-Path Express Fabric Switches Hardware Installation Guide	H76456	15.0
Cornelis Omni-Path Express Host Fabric Interface Installation Guide	H76466	9.0
Cornelis Omni-Path Express Fabric Software Installation Guide	H76467	21.0
Cornelis Omni-Path Express Fabric Switches GUI User Guide	H76457	15.0
Cornelis Omni-Path Express Fabric Switches Command Line Interface Reference Guide	H76458	15.0
Cornelis Omni-Path Express Fabric Suite FastFabric User Guide	H76469	20.0
Cornelis Omni-Path Express Fabric Suite Fabric Manager User Guide	H76468	18.0
Cornelis™ Omni-Path Express Fabric Suite Fabric Manager GUI User Guide	H76471	17.0
Cornelis Omni-Path Express Fabric Host Software User Guide	H76470	18.0
Cornelis Omni-Path Express Gateway Installation and Setup Guide	A00034	1.0
Cornelis Performance Scaled Messaging 2 (PSM2) Programmer's Guide	H76473	17.0
Cornelis Omni-Path Express Fabric Performance Tuning User Guide	H93143	23.0
Cornelis Omni-Path Express IP and LNet Router Design Guide	H99668	13.0
Building Containers for Cornelis Omni-Path Express Fabrics using Docker and Singularity Application Note	J57474	10.0
Cornelis Omni-Path Express Management API Programmer's Guide	J68876	9.0
Cornelis Omni-Path Express Fabric Software Release Notes	A00061	3.0
Cornelis Omni-Path Express Fabric Manager GUI Software Release Notes	K59649	2.0
Cornelis Omni-Path Express Fabric Switches Release Notes (includes managed and externally-managed switches)	A00044	1.0
Cornelis Omni-Path Express Fabric Unified Extensible Firmware Interface (UEFI) Release Notes	A00015	1.0
Cornelis Omni-Path Express Fabric Thermal Management Microchip (TMM) Release Notes	A00068	1.0
Cornelis Omni-Path Express Fabric Firmware Tools Release Notes	A00069	1.0

1.16. Installation Requirements

This section provides installation requirements for this release.

1.16.1. Best Practices

Note the following Cornelis Omni-Path Express recommendations:

- Cornelis recommends that users update to the latest versions of Cornelis Omni-Path Express 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.

- To improve security, Cornelis recommends updating the default HTTPS certificate. Refer to the *Cornelis Omni-Path Express Fabric Switches GUI User Guide*, "Updating the Certificate" for details.

- To improve security, Cornelis recommends configuring the `MgmtAllowed` setting and consider limiting access to port configuration changes by limiting access to Userspace Management Datagrams (UMADs). Refer to the *Cornelis Omni-Path Express Fabric Software Installation Guide*, "About User Queries Settings" for more information.

1.16.2. Software and Firmware Requirements

Operating Systems lists the operating systems supported by this release. For the required packages, refer to OS RPMs Installation Prerequisites.

1.16.3. OS RPMs Installation Prerequisites

Ensure that the following requirements are met before installing the software.

- Along with normal OS installation options, additional OS RPMs must be installed before you can install the Cornelis Omni-Path Express 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.

1.16.3.1. RHEL OS RPMs

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

> **NOTE**
> - 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 7. RHEL 7.7 , 7.8, and 7.9 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

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
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 8. RHEL 8.3, 8.4 and 8.5 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

1.16.3.2. SLES 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 9. 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)
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 10. 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 11. SLES 15.2 and 15.3 Distribution RPMs

IB/OPX 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				

1.16.3.3. Example

The example below shows the installation of a group of OS RPMs on a SLES 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
```

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

1.16.4. Installation Instructions

There are two Cornelis Omni-Path Express Fabric Software packages:

- `CornelisOPX-OPXS.<distro>-x86_64.<version>.tgz` for the management node.

- `CornelisOPX-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.

> **❗ IMPORTANT**
>
> If you want to install the CUDA versions of the OPXS software using the INSTALL command, you must use the `-G` option.

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

1.16.5. Installation Path Changes in Release 10.4 (and later)

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

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

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

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

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

1.17. 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.

- Some newer distros include an updated version of ipoib which has an incompatibility with FMs prior to this release. This could result in the inability to establish ipoib connections. Ensure that OS versions are consistent across the fabric.

1.18. 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 Cornelis Omni-Path Express Edge Switch 100 Series

  – Two ASICs per Cornelis Omni-Path Express Chassis 100 Series Leaf module

  – Two ASICs per Cornelis Omni-Path Express Chassis 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.

1.19. Accelerated RDMA Information

Accelerated RDMA is a Verbs protocol extension to improve the performance of RDMA write and RDMA read operations on Cornelis Omni-Path Express 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 Accelerated (or 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=0x4c09a09cbba` to the `/etc/modprobe.d/hfi1.conf` file. Instructions on how to do this are in the *Cornelis Omni-Path Express Fabric Performance Tuning User Guide*, "Setting HFI1 Driver Parameters" section.

2. Issues

This section lists the resolved and open issues in the Cornelis Omni-Path Express Fabric 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 12. Issues Resolved in this Release

ID	Description	Resolved in Release
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.	10.11.1
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.	10.11.1
STL-60129	When attempting to launch large scale jobs, you may see messages such as: ```Received eager message(s) ptype=0x1 opcode=0xcc from an unknown process (err=49)``` PSM2 connections are timing out and OFI is not honoring the error returned by PSM2 and later attempting to send messages over the failed connection. OFI overrides the PSM2_CONNECT_TIMEOUT variable to 5 seconds (`.conn_timeout = 5` in `psmx2_init.c`), whereas the PSM2 code specifies 0 (no timeout).	10.11.1
STL-61241 STL-61299	On some platforms, the hfi1 device may not show up in BIOS/UEFI boot menus and may not be available as a PXE boot device. This is caused by the platform not loading the UEFI driver for the hfi1 adapter.	10.11.1
STL-61605	OMPI 4.1.1 fails to build with PSM2 support when PSM2 is installed in a non-default location.	10.11.1
STL-62722	When freeing dma memory associated with eager buffers, the test for "is allocated" incorrectly checks the dma address. The dma allocation is then leaked.	10.11.1
STL-62926	The OPXS installation fails on newer distro kernels due to the prerequisite package kernel-abi-whitelists package being replaced by kernel-abi-stablelists.	10.11.1
STL-62933	A kernel panic occurred when the AIP send queue size was increased from its default value.	10.11.1
STL-63090	Under rare circumstances when a job is killed while a rank is being brought up a panic could occur.	10.11.1

2.1.2. Issues Resolved in Prior Releases

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

Table 13. Issues Resolved in Prior Releases

ID	Description	Resolved in Release
STL-61143	Certain BIOS versions or settings may cause the BIOS to report a numa_node for hfi1 as `NO_NUMA_NODE`, resulting in a probe time panic.	10.11.0.1
STL-61504 STL-61602	The addition of an upstream patch (633d61021298 "RDMA/ipoib: Remove racy Subnet Manager sendonly join checks") to 3.10.0-1160.25.1 caused issues with both the INSTALL RPM build and with the FM handling of a new join type.	10.11.0.1
STL-61604	Certain RedHat Z-stream kernel releases fail to build due to newly incorporated changes. RedHat backported a patch that the IFS driver needs to handle.	10.11.0.1
STL-61689	A build issue occurs with SLES security kernel update and IFS driver due to redefined components in the common_compat.h file.	10.11.0.1
STL-60901	Child process can access the parent memory management through the hfi dev file handle.	10.11.0
STL-61152	atomic_fetch_add_unless() was added to the Kernel API in 4.18.0-193.28.1. OPA IFS driver cannot be compiled and installed on RHEL 8.2 and its equivalent CentOS-8 (2004).	10.11.0
STL-59934	Customers experience errors similar to `hfi1_0: Send Context 8(151) Error: WriteOverflow` in dmesg or console logs.	10.11.0
STL-59971	The topology thread responsible for handling the sweep functionality of the SM occasionally hangs. This prevents the SM from responding to changes within the fabric.	10.10.3.2
STL-55704	IPoIB and core/mcast overreact to SM LID change by unnecessarily flushing connections, path records, and multicast membership.	10.10.3.1
STL-60776	While running a verbs upper-layer protocol (ULP), incorrect reporting of queue pair (QP) errors associated with receiving messages may cause a QP to enter the error state. The observed symptom may be reported as an error by the ULP or the system may hang if the ULP does not report the error associated with the QP state transition to error. Other symptoms are possible and ULP-dependent.	10.10.3.1
STL-60897	Links/ports inconsistency are being generated due to a race-condition caused by multiple threads accessing data structures not properly protected with mutex semaphore.	10.10.3.1
STL-47125	In rare cases, a program crash may be experienced during process cleanup when running with the OpenMPI PSM2 MTL (Matching Transport Layer). This issue no longer exists with Open MPI 3.1.4 and later.	10.10.3
STL-48480	When running workloads with 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. As of 10.10.3, OFI Verbs is no longer supported.	10.10.3
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. As of 10.10.3, OFI Verbs is no longer supported.	10.10.3
STL-60729	Eight bytes of memory will leak with every load of the hfi1 driver. This is associated with affinity processing.	10.10.3
STL-60134	PSM2-CUDA cache false-positives can occur when different processes on the same node share different regions of GPU memory that have the same CUDA virtual addresses. This results in unnecessary cache evictions and sub-optimal performance.	10.10.2.2
STL-60330 STL-60473	Issues with the maintenance of the waiting list that support SDMA-independent progress for PSM and AIP.	10.10.2.2
STL-60488	PSM2-CUDA cache does not handle CUDA IPC open function returning an ALREADY_MAPPED error. PSM2-CUDA cache does not properly update linked-list pointers, resulting in aborts or sub-optimal caching behavior.	10.10.2.2
STL-60542 STL-60629	PSM2 fails to open a port on a system with multiple HFIs if an inactive HFI appears before an active HFI in the list (e.g., hfi1_0 is inactive but hfi1_1 is active).	10.10.2.2

ID	Description	Resolved in Release
STL-60533	A compatibility issue introduced with SLES batch update, sles12.4-4.12.14-95.51, causes installation to fail.	10.10.2.1
STL-49210	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.	10.10.2
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.	10.10.2
STL-58443	When Accelerated RDMA (also known as TID RDMA) is enabled with 256 nodes or greater, it is possible that a storage write may not complete.	10.10.2
STL-59923	The version output of `modinfo hfi1` shows 10.9-0 for IFS 10.10 releases.	10.10.2
STL-59944	Installing IFS on SLE HPC 15.x fails because there is no file `/etc/SuSE-release`.	10.10.2
STL-60295	OpenMPI with CUDA support is installed when a customer installs the non-CUDA IFS using the command INSTALL, that is, runs INSTALL without the argument `-G`. This issue is benign and can be ignored.	10.10.2
135830 (STL-46193)	On systems, failure observed during software upgrade when rebuilding the boot image. Error message contains: Rebuilding boot image with "/usr/bin/dracut -f" As of 10.10.2, this information has been moved into the *Cornelis Omni-Path Express Fabric Software Installation Guide* and will remain there until RHEL 7.x is discontinued.	10.10.2
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.	10.10.2
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	over Fabric Protocol is only supported on OPX 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

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, OPX 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 OPX'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

2.2. Open Issues

The following table lists the open issues for this release.

Table 14. Open Issues

ID	Description	Workaround
134819 (STL-49733)	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.
136822 (STL-49728)	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 (STL-49732)	The Subsystem Vendor and Subsystem Device ID in the PCI configuration space of Cornelis Omni-Path Express 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 Cornelis Omni-Path Express Fabric 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-48921	On some Distributions, 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-49313	On OS versions (RHEL 7.7 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`).

ID	Description	Workaround
| 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:<br><br>```
systemctl stop irqbalance
modprobe -r ib_ipoib
modprobe -r hfi1
modprobe hfi1
modprobe ib_ipoib
systemctl start irqbalance
```<br><br>**NOTE:** Additional steps (like stopping Lustre) may be required on more complicated fabrics. |
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 *Cornelis Omni-Path Express Fabric Performance Tuning User Guide*, 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 Cornelis Networks Customer Support if you need further guidance.
STL-60932	On 3rd Generation Intel Xeon Scalable Processors, IPoFabric small message latency may measure higher than expected depending on the tunings being used.	Cornelis recommends that you use the Tuned latency-performance profile to significantly reduce IPoFabric small message latency. Tuned is a utility (for example, tuned-2.10.0-15.el8.noarch) that allows for dynamic and adaptive tuning of CPU behavior. Note that using Tuned latency-performance profile may have a negative performance impact to other bandwidth/message rate measurements.
| STL-61176 | Running perftest, the following text may be displayed:<br><br>```
Couldn't post send: qp 0 scnt=189
Failed to complete run_iter_bw function successfully
```<br><br>**NOTE:** This failure is only seen on perftest and does not seem to impact any other upper layer protocols (ULPs). However, it is worse on RHEL 8.3. | Disable TID RDMA. |
STL-61215	In rare instances, the host may boot and the HFI device is not visible on the PCI bus, and is therefore not usable.	Reboot/power cycle the host.
STL-61244	Servers with 2nd Gen AMD EPYC Processor may experience an unexpected server reboot when running under load with a high number of processes per node.	Contact Cornelis Networks Customer Support for more information.
STL-61256	On 2nd Gen AMD EPYC Processors, AIP interrupts do not map optimally in all situations, especially when certain BIOS settings (such as NPS) are modified. This can impact IPoFabric performance.	Manually remap the AIP interrupts to give the best performance. Refer to the *Cornelis Omni-Path Express Fabric Performance Tuning User Guide*, "Driver IRQ Affinity Assignments" section for more details. Or, contact Cornelis Networks Customer Support if you need further guidance."
STL-61329	On RHEL 7.x servers with two HFIs installed (dual-rail), unbinding both HFIs from the hfi1 driver may result in a kernel warning trace.	None.

ID	Description	Workaround
STL-61331	When installing on SLES 12.x from a yum/zypper repository, the installation may fail due to missing dependencies.	Do not use `-y` when performing the yum/zypper installation and select the option to ignore the dependencies.
STL-61527	Bidirectional host to host workloads using a GPU IFS installation do not achieve maximum performance.	Use a non-GPU IFS install for host-host workloads, or set `PSM2_CUDA=0`, `PSM2_GPUDIRECT=0` in the job environment for host-only workloads.
STL-61590	Open MPI (OMPI) has changed the support for one-sided operations since OMPI 4.0.5. Because of this change, OMPI is now launching both PSM2 MTL and OFI Byte Transfer Layer (BTL), which consumes more contexts than it actually needs. This can result in runtime warnings/errors on some high PPN jobs.	Build OMPI by adding --with-ofi=no to the configure line. Alternatively, you can run -mca btl self,vader. This string explicitly tells OMPI to use -mca btl self,vader as the BTL instead of OFI and prevents OMPI from consuming more contexts than it actually needs. **Note:** Applications that require BTL to communicate may be adversely affected by this change. These are limited to a subset of MPI applications. Intel MPI benchmarks (IMB) will not be impacted.
STL-62119	Issues running the allreduce IMB benchmark using older versions of Intel MPI (ie: 2020.4) on AMD Milan systems have been observed.	Upgrade to use a more recent version of IMPI and IMB if this issue is seen.
STL-62604 STL-62605	Running certain IMB benchmarks using the the IMB source code packaged with oneAPI 2021.3 (or earlier) can result in benchmark crashes.	Upgrade to use a more recent version of IMB or upgrade to oneAPI 2021.4
STL-62613	The capmask for Accelerated IP (AIP) has changed starting in OPXS 10.11.1.0. This can cause AIP to be disabled if you are using the old cap mask referenced in the performance tuning guide, which is used to enable Accelerated RDMA.	The correct capmask is to use is now 0x4c09a09cbba.

2.2.1. Third Party Open Issues

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

Table 15. 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 PSM2. 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 *Cornelis Omni-Path Express Fabric Performance Tuning User Guide*, HFI1 Driver Module Parameters chapter.
141273 (STL-46935)	The in-distro version of perftests has bugs.	Use the upstream version of perftest from https://github.com/linux-rdma/perftest.

ID	Description	Workaround
STL-47571	When trying to run an MPI/PSM job with more MPI ranks than CPU cores (oversubscribing), the job may fail with the following error message: ```hfi_userinit: assign_context command failed: Device or resource busy PSM2 can't open hfi unit: -1 (err=23)```	Set `PSM2_MULTI_EP=0` (user environment variable) before or during job launch. For details, see the *Cornelis Performance Scaled Messaging 2 (PSM2) Programmer's Guide*.
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-59449	Occasionally, an unload of the hfi1 driver will produce a message similar to the following: ```WARNING: CPU: 15 PID: 119002 at ../kernel/workqueue.c:4091 destroy_workqueue+0x61/0x240()``` This message will be followed by a call trace showing the hfi1 unload. This issue is related to memory reclaim activity just before the unload.	None. **NOTE:** This is a bug in the Linux kernel. Cornelis Networks is working with the Linux community to resolve this issue.
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. `size=$(blockdev --getsize /dev/ram0)` 3. **`dmsetup`** `create mdevice0 --table "0 ${size} linear /dev/ram0 0"`	None.
STL-60633	When you install `gpfs.callhome-ecc-client-5.0.5-0` under SLES using zypper, you may receive an error code if there are no existing credentials.	Ignore the error code. **NOTE:** To avoid this error, you can touch an empty file `/usr/lpp/mmfs/bin/ecc/ECCBase/com.ibm.ws.webservices.thinclient_7.0.0.jar` to force zypper to exit with a success code. For example, run the following command before you install `gpfs.callhome-ecc-client-5.0.5-0`: ```touch /usr/lpp/mmfs/bin/ecc/ECCBase/com.ibm.ws.webservices.thinclient_7.0.0.jar```
STL-61762	osu_mbw_mr can report bandwidth greater than line-rate in OSU micro benchmarks version 4.1 and higher.	None.
LU-14733 (STL-61911)	The lustre Lnet layer can emit the following message: `dk.opal63.llnl.gov.7:00000001:00020000:43.0:162 2598261.714620:0:129525:0: (brw_test.c:415:brw_bulk_ready()) BRW bulk READ failed for RPC from 12345-192.168.128.126@o2ib18: -103` The issue happens when an FRMR invaldation operation is failed in driver. RHEL 8.4 is exposed to the issue due to the release removing the older FMR kernel registration scheme. Lustre is sending an old previously used invalidate.	Use Lustre client release 2.12.8 or Build Lustre from source since the patches are in the Lustre source

3. Related Information

3.1. Omni-Path Express Rebranding

The rebranding changes included in this release make updates to the installation packages for management and compute nodes.

The Intel Fabric Suite or IFS package, intended for fabric management/head nodes is being replaced with Cornelis Omni-Path Express Fabric Suite (OPXS) to align with our rebranding efforts as Cornelis Networks.

The Basic package naming convention is retained for non-management nodes/ compute nodes.

The package names are changed as follows:

Table 16. Package Name Rebranding

Old Package Name	New Package Name
IntelOPA-Basic.<distro>-x86_64.<version>.tgz	CornelisOPX-Basic.<distro>-x86_64.<version>.tgz
IntelOPA-IFS.<distro>-x86_64.<version>.tgz	CornelisOPX-OPXS.<distro>-x86_64.<version>.tgz

The subdirectory structure in the OPXS and Basic packages also replace references to IntelOPA with CornelisOPX. The subdirectory names are changed as follows:

Table 17. Subdirectory Structure Rebranding

IntelOPA-OFA_Delta	CornelisOPX-OFA_DELTA
IntelOPA-Tools	CornelisOPX-Tools[1]
IntelOPA-Tools-FF	CornelisOPX-Tools-FF[2]
IntelOPA-FM	CornelisOPX-FM[2]
Note: 1. Basic only 2. OPXS only	

3.2. Documentation Library

Log in to the Cornelis Customer Center to download the Cornelis Omni-Path Express publications from the Release Library.

Use the tasks listed in this table to find the corresponding Cornelis Omni-Path Express document.

Task	Document Title	Description
Using the OPX documentation set	*Cornelis Omni-Path Express Fabric Quick Start Guide*	A roadmap to Cornelis comprehensive library of publications describing all aspects of the product family. This document outlines the most basic steps for getting your Cornelis Omni-Path Express (OPX) cluster installed and operational.

Task	Document Title	Description
Setting up an OPX cluster	*Cornelis Omni-Path Express Fabric Setup Guide*	Provides a high-level overview of the steps required to stage a customer-based installation of the Cornelis Omni-Path Express Fabric. Procedures and key reference documents, such as Cornelis Omni-Path Express 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	*Cornelis Omni-Path Express Fabric Switches Hardware Installation Guide*	Describes the hardware installation and initial configuration tasks for the Cornelis Omni-Path Express Switches 100 Series. This includes: Cornelis Omni-Path Express Edge Switches 100 Series, 48-port configurable Edge switches, and Cornelis Omni-Path Express Director Class Switches 100 Series.
	Cornelis Omni-Path Express Host Fabric Interface Installation Guide	Contains instructions for installing the HFI in an OPX cluster.
	Cornelis Omni-Path Express Gateway Installation and Setup Guide	Describes the hardware installation and set up tasks for the Cornelis Omni-Path Express Gateways.
Installing host software Installing HFI firmware Installing switch firmware (externally-managed switches)	*Cornelis Omni-Path Express 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)	*Cornelis Omni-Path Express Fabric Switches GUI User Guide*	Describes the graphical user interface (GUI) of the Cornelis Omni-Path Express Fabric Chassis Viewer GUI. This document provides task-oriented procedures for configuring and managing the Cornelis Omni-Path Express Switch family. Help: GUI embedded help files
Managing a switch using the CLI Installing switch firmware (managed switches)	*Cornelis Omni-Path Express Fabric Switches Command Line Interface Reference Guide*	Describes the command line interface (CLI) task information for the Cornelis Omni-Path Express Switch family. Help: -help for each CLI
Managing a fabric using FastFabric	*Cornelis Omni-Path Express 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	*Cornelis Omni-Path Express Fabric Suite Fabric Manager User Guide*	The Fabric Manager uses a well-defined management protocol to communicate with management agents in every Cornelis Omni-Path Express Host Fabric Interface (HFI) and switch. Through these interfaces the Fabric Manager can discover, configure, and monitor the fabric.
	Cornelis Omni-Path Express 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 HFI and IPoIB driver Running MPI applications on OPX	*Cornelis Omni-Path Express 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.

Task	Document Title	Description
Writing and running middleware that uses OPX	*Cornelis Performance Scaled Messaging 2 (PSM2) Programmer's Guide*	Provides a reference for programmers working with the PSM2 Application Programming Interface (API). The Performance Scaled Messaging 2 API (PSM2 API) is a low-level user-level communications interface.
Optimizing system performance	*Cornelis Omni-Path Express Fabric Performance Tuning User Guide*	Describes BIOS settings and parameters that have been shown to ensure best performance, or make performance more consistent, on Cornelis Omni-Path Express 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 OPX	*Cornelis Omni-Path Express 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 Cornelis Omni-Path Express and a legacy InfiniBand fabric.
Building Containers for OPX fabrics	*Building Containers for Cornelis Omni-Path Express 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 Cornelis Omni-Path Express networking technology.
Writing management applications that interface with OPX	*Cornelis Omni-Path Express Management API Programmer's Guide*	Contains a reference for programmers working with the Cornelis Omni-Path Express Architecture Management (OPAMGT) Application Programming Interface (API). The OPAMGT API is a C-API permitting in-band and out-of-band queries of the FM's Subnet Administrator and Performance Administrator.
Using NVM over Fabrics on OPX	*Configuring Non-Volatile Memory Express (NVMe) over Fabrics on Cornelis Omni-Path Express Application Note*	Describes how to implement a simple Cornelis Omni-Path Express Architecture-based point-to-point configuration with one target and one host server.
Learning about new release features, open issues, and resolved issues for a particular release	*Cornelis Omni-Path Express Fabric Software Release Notes*	
	Cornelis Omni-Path Express Fabric Manager GUI Software Release Notes	
	Cornelis Omni-Path Express Fabric Switches Release Notes (includes managed and externally-managed switches)	
	Cornelis Omni-Path Express Fabric Unified Extensible Firmware Interface (UEFI) Release Notes	
	Cornelis Omni-Path Express Fabric Thermal Management Microchip (TMM) Release Notes	
	Cornelis Omni-Path Express Fabric Firmware Tools Release Notes	

3.2.1. How to Search the Cornelis Omni-Path Express 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 Cornelis Omni-Path Express 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.