



Release Notes for Cisco OFED Release 1.2.5

Release Date: October 9, 2007
Part Number: OL-11537-04

Contents

This document is the release notes for Cisco OFED release 1.2.5 and includes the following sections:

- [Contents, page 1](#)
- [Introduction, page 1](#)
- [System Requirements, page 2](#)
- [New and Changed Information, page 5](#)
- [Caveats, page 8](#)
- [Related Documentation, page 11](#)
- [Service and Support, page 11](#)
- [Obtaining Documentation, Obtaining Support, and Security Guidelines, page 12](#)

Introduction

These release notes describe the features and known issues for the Cisco OpenFabrics Enterprise Distribution (OFED) InfiniBand Host Drivers 1.2.5 for Linux.



Note

Cisco OFED 1.2.5 should only be used with ConnectX HCAs. Cisco OFED 1.2 should be used with other types of HCAs.



System Requirements

This section describes the system requirements for this software release and includes the following topics:

- [Determining the Software Version, page 2](#)
- [Upgrading to a New Software Release, page 2](#)
- [Supported Kernels, page 3](#)
- [Supported MPI Implementations, page 4](#)
- [Supported Compilers for MPI, page 4](#)
- [Supported HCA Firmware, page 5](#)

Cisco supports OFED on all existing Cisco InfiniBand HCAs, Cisco InfiniBand switches, and Cisco Ethernet and Fibre Channel Gateways.

Determining the Software Version

If InfiniBand drivers are already installed on the host, they may be installed in one of several locations.

To determine the version of the Cisco InfiniBand host drivers, log in to the host and enter the following commands at the shell prompt. If the first command produces output, the Cisco Commercial InfiniBand host drivers are installed. If the second or third commands produce a version number, OFED host drivers are installed.

```
host$ rpm -qa | grep topspin
topspin-ib-mpi-rhel4-3.2.0-118
topspin-ib-mod-rhel4-2.6.9-34.ELsmp-3.2.0-118
topspin-ib-rhel4-3.2.0-118
host$ ofed_info | grep OFED
OFED-1.1
host$ grep OFED /usr/local/ofed/BUILD_ID
OFED-1.0
```

Upgrading to a New Software Release

To verify that you are running the latest available release, compare your version against the latest version on the Cisco support website at the following URL:

<http://www.cisco.com/cgi-bin/tablebuild.pl/sfs-linux>.

After registering your product, you should have received a username and password that grant you access to this site.

Switch software and Linux host drivers are released and packaged separately. Cisco OFED 1.2.5 requires that all switches first be upgraded to TopspinOS 2.4.0 or higher. Switches should be upgraded before the InfiniBand hosts.

Use the **ofedinstall** command to install Cisco OFED 1.2.5 on your host. There can only be one set of InfiniBand drivers installed on a host at any time. The **ofedinstall** command will first uninstall any Cisco Commercial InfiniBand host drivers or other OpenFabrics host drivers. If the host has other non-OpenFabrics drivers, those must be uninstalled first before running **ofedinstall**. If you are downgrading OFED, you should also uninstall the newer drivers before running **ofedinstall**.

To enable userspace components such as MPI, **ofedinstall** adds the following entries to `/etc/security/limits.conf`. You may tune the value `unlimited` to a specific amount of RAM if desired.

```
* soft memlock unlimited
* hard memlock unlimited
```

For general information about upgrading to a new software release, see the Installing Host Drivers chapter in the *Cisco OpenFabrics Enterprise Distribution InfiniBand Host Drivers User Guide for Linux*.

Supported Kernels

Cisco OFED 1.2.5 InfiniBand host drivers are supported on the following kernels:

- Red Hat Enterprise Linux 4 (RHEL4)
 - 2.6.9-34.ELsmp (Update 3) for i686
 - 2.6.9-34.ELhugemem (Update 3) for i686
 - 2.6.9-34.ELsmp (Update 3) for x86_64
 - 2.6.9-34.ELlargesmp (Update 3) for x86_64
 - 2.6.9-34.EL (Update 3) for ia64
 - 2.6.9-34.ELlargesmp (Update 3) for ia64
 - 2.6.9-34.EL (Update 3) for ppc64
 - 2.6.9-34.ELlargesmp (Update 3) for ppc64
 - 2.6.9-42.ELsmp (Update 4) for i686
 - 2.6.9-42.ELhugemem (Update 4) for i686
 - 2.6.9-42.ELsmp (Update 4) for x86_64
 - 2.6.9-42.ELlargesmp (Update 4) for x86_64
 - 2.6.9-42.EL (Update 4) for ia64
 - 2.6.9-42.ELlargesmp (Update 4) for ia64
 - 2.6.9-42.EL (Update 4) for ppc64
 - 2.6.9-42.ELlargesmp (Update 4) for ppc64
 - 2.6.9-55.ELsmp (Update 5) for i686
 - 2.6.9-55.ELhugemem (Update 5) for i686
 - 2.6.9-55.ELsmp (Update 5) for x86_64
 - 2.6.9-55.ELlargesmp (Update 5) for x86_64
 - 2.6.9-55.EL (Update 5) for ia64
 - 2.6.9-55.ELlargesmp (Update 5) for ia64
 - 2.6.9-55.EL (Update 5) for ppc64
 - 2.6.9-55.ELlargesmp (Update 5) for ppc64
- SUSE Linux Enterprise Server 10 (SLES10)
 - 2.6.16.21-0.8-smp for i686
 - 2.6.16.21-0.8-smp for x86_64
 - 2.6.16.21-0.8-default for ia64

- 2.6.16.21-0.8-ppc64 for ppc64
- 2.6.16.46-0.12-smp (Service Pack 1) for i686
- 2.6.16.46-0.12-smp (Service Pack 1) for x86_64
- 2.6.16.46-0.12-default (Service Pack 1) for ia64
- 2.6.16.46-0.12-ppc64 (Service Pack 1) for ppc64
- Red Hat Enterprise Linux 5 (RHEL5)
 - 2.6.18-8.el5 for i686
 - 2.6.18-8.el5PAE for i686
 - 2.6.18-8.el5 for x86_64
 - 2.6.18-8.el5 for ia64
 - 2.6.18-8.el5 for ppc64

If you are using a different kernel, you will need to compile OFED from source code. To compile OFED source code, unpack the OFED tarball in `src/` in the Cisco OFED ISO image and follow the instructions in `README.txt`. Sample `.conf` files for `install.sh` are also provided in `src/` in the ISO image. If you compile OFED from source code, we will support you on a best-effort basis, but we recommend that you use the pre-built binary RPMs if possible. More information about OFED is available in the `ofed-docs` package on a host with the drivers installed, <http://www.openfabrics.org>, and the OpenFabrics general mailing list and its mail archives.

Supported MPI Implementations

Cisco OFED 1.2.5 InfiniBand host drivers have been tested with the following MPI implementations:

- Open MPI 1.2.2 (included with Cisco OFED 1.2.5)
- MVAPICH 0.9.9 (included with Cisco OFED 1.2.5)
- MVAPICH2 0.9.8 (included with Cisco OFED 1.2.5)
- Intel MPI 3.0.33
- HP MPI 2.2.5 (uDAPL interface only)

Other MPI implementations support OFED, for more information contact the MPI vendor.

Supported Compilers for MPI

Cisco OFED 1.2.5 InfiniBand host drivers are supported with the following compilers:

- GNU C, C++, Fortran 77, and Fortran 90
- Intel 9.1 and 10.0 C, C++, Fortran 77, and Fortran 90
- PGI (Portland Group) 6.2-5 and 7.0-2 C, C++, Fortran 77, and Fortran 90

Supported HCA Firmware

Cisco OFED 1.2.5 InfiniBand host drivers include the following InfiniBand HCA firmware, which is based on the Mellanox firmware from http://www.mellanox.com/support/firmware_table.php:

HCA Type	Description	Firmware Version
ConnectX	2 port PCIe	2.2.000
Cougar	2 port PCI-X (not low-profile)	3.5.917
Cougar Cub	2 port PCI-X	3.5.917
Lion Cub	2 port PCIe	4.8.917
Lion Mini	2 port PCIe memfree	5.2.917
Cheetah	1 port PCIe memfree	1.2.917
Cheetah DDR	1 port PCIe memfree DDR	1.2.917
IBM BladeCenter PCI-X	uses Cougar firmware	see Cougar
IBM BladeCenter PCIe	based on Lion Cub	4.8.917
Dell PowerEdge 1X55	based on Lion Cub	4.8.917

New and Changed Information

This section describes any new and changed information.

Cisco OFED 1.2.5 is a minor release that introduces support for the ConnectX HCA. The ConnectX HCA is a two-port PCIe HCA that offers improved latency, for example 1.4 microseconds with MVAPICH or MVAPICH2, as measured by the `osu_latency` MPI benchmark.

Changes from Release 1.2 to 1.2.5

This section describes the new features and resolved caveats since the Cisco OFED 1.2 release.



Note

The ID number from the OpenFabrics Defect Tracking System, if applicable. The current status of all issues is available online at the following URL: <https://bugs.openfabrics.org>.

General

- Support has been added for the ConnectX HCA.

Changes from Release 1.1 to 1.2

This section describes the new features and resolved caveats since the Cisco OFED 1.1 release.

General

- Support has been added for RHEL4 U5, RHEL5, and SLES10 SP1.
- Cisco OFED is now packaged in a single ISO image plus smaller ISO images for RHEL4, SLES10, and RHEL5, which allow for smaller downloads for these distributions.
- OFED now installs into /usr instead of /usr/local/ofed.
- The OFED uninstall program has been renamed from uninstall.sh to ofed_uninstall.sh.
- Upgraded HCA firmware is included.
- The hca_self_test utility now works with HP BladeSystem HP-supplied HCA firmware.
- 265
OFED now installs and uninstalls cleanly when multiple kernel-ib packages are installed.
- 428
32-bit libraries are now included along with 64-bit libraries on 64-bit Operating Systems. These 32-bit libraries have not been fully tested and are not supported by Cisco.

IPoIB

- 509
IPoIB CM (Connected Mode) has been added and is now the default. IPoIB CM creates one IB Reliable Connection between pairs of hosts that communicate using IPoIB. IPoIB CM has higher throughput than traditional IPoIB (which uses IB Unreliable Datagram connections). IPoIB CM and non-CM are interoperable, IPoIB CM will transparently fall back to using non-CM communication when necessary. Note that IP multicast still uses non-CM communication.
- 549
IPoIB High Availability has been added through the IPoIB bonding driver. It is based on the Linux Ethernet Bonding Driver and was adopted to work with IPoIB. The ib-bonding package contains the bonding driver and a utility called ib-bond to manage and control the driver operation. The IPoIB bonding drivers supports active/passive failover, pkey interfaces, multiple slave interfaces and multiple HCAs, and configuration at boot time in openib.conf.
- 506
A problem has been fixed where the maximum IPoIB multicast throughput was less than 1.5 Gbps. Note that this fix requires new HCA firmware, which is included with Cisco OFED 1.2.
- 261
IPoIB pkey interfaces can now be configured to start at boot time via ifcfg files.
- IPv6 has been tested and is now documented and supported for IPoIB.

MPI

- Open MPI has been upgraded from 1.1.2 to 1.2.2. More details on Open MPI 1.2.2 are available at <http://www.open-mpi.org>.
- Open MPI is now the recommended MPI implementation for Cisco customers.
- MVAICH has been upgraded from 0.9.7 to 0.9.9. More details on MVAICH 0.9.9 is available at <http://mvapich.cse.ohio-state.edu>.
- MVAICH2 0.9.8 has been added. More details on MVAICH2 0.9.8 is available at <http://mvapich.cse.ohio-state.edu>.
- A new MPI Selector feature (**mpi-selector-menu** and **mpi-selector** commands) has been added to set system or user default on which MPI implementation to use.
- MPI applications may now use the `system()` and `fork()` functions on RHEL4, SLES10, and RHEL5 if they first set the environment variable `IBV_FORK_SAFE`.
- Support for PGI (Portland Group) compilers (C, C++, Fortran 77, and Fortran 90) has been added.
- 188
Open MPI performance on Pallas/Intel MPI collective benchmarks is now improved.
- 187
Open MPI now supports MPI-2 Remote Memory Access.
- 136
Open MPI latency has been improved on ia64.

RDS

- RDS (Reliable Datagram Sockets) has been added back into OFED 1.2. RDS is supported at this time only for use with Oracle RAC.

SDP

- 175
A new `sdpnetstat` program has been added that is based on `netstat` and supports SDP.
- 108
SDP throughput with messages smaller than 8 KB and larger than 64 KB has been improved.
- The default behavior for `libsdp` has changed. Server applications that use `libsdp` will now by default listen on both SDP and TCP. Client applications that use `libsdp` will now by default try SDP first and then fallback to TCP if SDP is not available.

SRP

- 443
SRP High Availability has been added through the `srp_daemon` program and Device Mapper Multipath, which is included with both RHEL and SLES. Device Mapper Multipath supports both active/active and active/passive high availability, depending on the capability of the storage device.
- The command `cisco_srp_add_targets` has been removed, and has been replaced with `srp_daemon -e -o -n`.
- 474
The `srp_daemon` command is now supported with the Cisco Fibre Channel Gateway.

uDAPL

- 350
uDAPL test programs `dapltest` and `dtest` are now included.

Caveats

This section describes temporary limitations of this release. These restrictions will be resolved in a future release of this product.

General

- The `tvflash` command does not yet support the Mellanox ConnectX HCA. The `mstflint` command must be used to burn ConnectX HCA firmware.
- ConnectX HCAs require up-to-date switch firmware, for example TopspinOS 2.9.0 build 170 or higher, in order to operate at DDR. With older switch firmware, ConnectX HCA will operate at SDR.
- Cisco does not support the OpenSM InfiniBand subnet manager.
- 270
`tvflash` does not work with HCA recovery jumper.
- 183
The IB node description will sometimes not have the hostname in it.
- 292
`tvflash` does not work on SLES10 ia64.

IPoIB

- 541
IPoIB High Availability may experience slow failover time. This problem does not occur with the Cisco Commercial Linux IB drivers.
- IPoIB performance varies greatly by type of motherboard. The best performance will be achieved with the latest Intel and AMD processors and PCIe HCAs.

- IP multicast traffic with messages larger than 2048 bytes will cause a few benign kernel syslog message like “ib0: packet len 65520 (> 2048) too long to send, dropping” to appear. These messages indicate the path MTU is being changed, only for IP multicast traffic, not UDP or TCP traffic.
- 266
IPoIB IP multicast is not supported on RHEL4 U4. U4 IP multicast can only receive from U4 senders, and U4 senders sent traffic to all nodes, not just nodes that have joined the multicast address. U4 multicast traffic can cause a Cisco InfiniBand switch to reboot. This problem is fixed in RHEL4 U5 and RHEL5.
- The RHEL4 U3 and U4 **tcpdump** command does not work with OFED IPoIB. This problem has been fixed in RHEL4 U5.
- RHEL4 and SLES10 **ifconfig -a** commands do not display the IPoIB interface address correctly. The workaround is to use the **ip link show ib0** command instead.
- 351
The ib1 interface cannot be configured at boot time through `/etc/sysconfig/network/ifcfg-ib1` on SLES10. Novell plans to fix this in SLES10 SP1.
- 269
IPoIB non-CM does not interoperate with the mgmt-ib IPoIB management interface on the Cisco SFS 3001 and Topspin 90 switch. Note that IPoIB non-CM is not the default in OFED 1.2.

MPI

- Open MPI has not yet been optimized for ConnectX. The `osu_latency` benchmark will report at best 1.9 microseconds latency.
- 567
MPI is not supported on RHEL5 ppc64.
- Open MPI performance on simple microbenchmarks (OSU MPI benchmarks, Intel MPI benchmarks, NetPIPE, and so on) is improved by adding `--mca -mpi_leave_pinned 1` to the `mpirun` command line. Note that this parameter is only useful for MPI applications that reuse their communication buffers. More information on tuning Open MPI performance for OFED is available at the following URL:
<http://www.open-mpi.org/faq/?category=openfabrics>
- MPI does not have any built-in High Availability, MPI applications have to handle interface failover.
- 585
Intel MPI and HP MPI can be used with IPoIB bonding interfaces, but the `bond0` uDAPL provider must be used instead of the `ib0` provider.
- 369
HP MPI 2.2.5 does not support IB Verbs API (`mpirun -IBV`) with OFED 1.2. The workaround is to use the uDAPL API (`mpirun -UDAPL`). This will be fixed in the next release of HP MPI.
- 550
To use HP MPI with uDAPL and multiple processes per node, the entries for `ib1`, `ib2`, and `ib3` must be commented out in `/etc/dat.conf`.
- 258
MVAPICH does not support GNU Fortran 90 on ppc64.

SDP

- 137
OFED SDP does not interoperate with Cisco Commercial Linux InfiniBand Host Driver SDP.
- SDP performance varies greatly by type of motherboard. The best performance will be achieved with the latest Intel and AMD processors and PCIe HCAs.
- SDP does not have any built-in High Availability, SDP applications have to handle interface failover.
- 38
OFED SDP does not support Linux AIO.
- 25
There is no `#define` for `AF_INET_SDP`, code that uses SDP without `libsdp` must manually add `#define AF_INET_SDP 27` at the beginning of a program.
- 294
SDP does not support using the `connect()` function with `AF_INET_SDP`. The workaround is to use `AF_INET` for this function instead. `AF_INET_SDP` must be used with the `socket()` and `bind()` functions.
- SDP does not support IPv6, although `libsdp` does support IPv4 addresses encapsulated in IPv6 addresses.

SRP

- 577
SRP High Availability may experience slow failover time. This problem does not occur with the Cisco Commercial Linux IB drivers.
- 624
The SRP `ib_srp` kernel module will often fail to unload manually with `rmmod` or `modprobe -r`, even if SRP is not in use. The workaround is to reboot the system.
- 262
SRP filesystems cannot be mounted from `/etc/fstab`. The workaround is to manually mount these filesystems in `/etc/rc.local`.

uDAPL

- uDAPL is supported at this time only for use with Intel MPI and HP MPI.
- uDAPL does not have any built-in High Availability, uDAPL applications have to handle interface failover.
- 48
uDAPL is not available on `ppc64`.

Related Documentation

The following list describes the documentation available with TopspinOS 2.9.0, which is available in electronic form and printed form upon request.



Note

Documentation is included on the TopspinOS 2.9.0 Server Switch CD Image.

You can download the latest documentation updates on the Cisco support site at http://www.cisco.com/en/US/partner/products/ps6418/tsd_products_support_category_home.html.

- *Cisco InfiniBand Hardware Installation and Cabling Guide*
- *Cisco SFS 7000P Switch Installation and Configuration Note*
- *Cisco SFS 7000P Hardware Installation Guide*
- *Cisco SFS 7008P Switch Installation and Configuration Note*
- *Cisco SFS 7008P Hardware Installation Guide*
- *Release Notes for TopspinOS Release 2.9.0*
- *Cisco SFS Product Family Chassis Manager User Guide*
- *Cisco SFS Product Family Element Manager User Guide*
- *Cisco SFS Product Family Command Reference*

Service and Support

For additional support, you must first register your product at <http://www.cisco.com>. After registering, you may contact your supplier for support, or Cisco directly.

Refer to the “[Obtaining Documentation, Obtaining Support, and Security Guidelines](#)” section on [page 12](#) in this document.

When you call Cisco Technical Support or use the Cisco Technical Support website at <http://www.cisco.com>, be prepared to provide the following information to support personnel:

General Information

- Technical Support registration number, if applicable
- Error messages received
- Detailed description of the problem and specific questions
- Description of any troubleshooting steps already performed and results

Server Configuration

- Type of server, chip set, CPU, amount of RAM, and number of nodes
- Attached storage devices (output from `cat /proc/scsi/scsi`)
- InfiniBand configuration (output from `hca_self_test`)

Chassis Configuration

- Chassis model
- Output from the **show running-status all** command

Chassis Serial Number

The chassis serial number and corresponding bar code are provided on the serial number label, as shown in this example:

Model: TS360



SN UST323XXXXXXXXXX

This chassis serial number can be found on the bottom of the chassis or the outside of the chassis box packaging. It can also be found in the output of the **show backplane** command.

Obtaining Documentation, Obtaining Support, and Security Guidelines

For information on obtaining documentation, obtaining support, providing documentation feedback, security guidelines, and also recommended aliases and general Cisco documents, see the monthly *What's New* in Cisco Product Documentation, which also lists all new and revised Cisco technical documentation, at:

<http://www.cisco.com/en/US/docs/general/whatsnew/whatsnew.html>

This document is to be used in conjunction with the documents listed in the “[Related Documentation](#)” section.

CCVP, the Cisco logo, and the Cisco Square Bridge logo are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn is a service mark of Cisco Systems, Inc.; and Access Registrar, Aironet, BPX, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, IP/TV, iQ Expertise, the iQ logo, iQ Net Readiness Scorecard, iQuick Study, LightStream, Linksys, MeetingPlace, MGX, Networking Academy, Network Registrar, PIX, ProConnect, ScriptShare, SMARTnet, StackWise, The Fastest Way to Increase Your Internet Quotient, and TransPath are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0709R)

Any Internet Protocol (IP) addresses used in this document are not intended to be actual addresses. Any examples, command display output, and figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses in illustrative content is unintentional and coincidental.

© 2007 Cisco Systems, Inc. All rights reserved.