



Connect. Accelerate. Outperform.™

Mellanox OFED for Linux Release Notes

Rev 2.4-1.0.4

NOTE:

THIS HARDWARE, SOFTWARE OR TEST SUITE PRODUCT (“PRODUCT(S)”) AND ITS RELATED DOCUMENTATION ARE PROVIDED BY MELLANOX TECHNOLOGIES “AS-IS” WITH ALL FAULTS OF ANY KIND AND SOLELY FOR THE PURPOSE OF AIDING THE CUSTOMER IN TESTING APPLICATIONS THAT USE THE PRODUCTS IN DESIGNATED SOLUTIONS. THE CUSTOMER’S MANUFACTURING TEST ENVIRONMENT HAS NOT MET THE STANDARDS SET BY MELLANOX TECHNOLOGIES TO FULLY QUALIFY THE PRODUCT(S) AND/OR THE SYSTEM USING IT. THEREFORE, MELLANOX TECHNOLOGIES CANNOT AND DOES NOT GUARANTEE OR WARRANT THAT THE PRODUCTS WILL OPERATE WITH THE HIGHEST QUALITY. ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT ARE DISCLAIMED. IN NO EVENT SHALL MELLANOX BE LIABLE TO CUSTOMER OR ANY THIRD PARTIES FOR ANY DIRECT, INDIRECT, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES OF ANY KIND (INCLUDING, BUT NOT LIMITED TO, PAYMENT FOR PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY FROM THE USE OF THE PRODUCT(S) AND RELATED DOCUMENTATION EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.



Mellanox Technologies
 350 Oakmead Parkway Suite 100
 Sunnyvale, CA 94085
 U.S.A.
www.mellanox.com
 Tel: (408) 970-3400
 Fax: (408) 970-3403

© Copyright 2015. Mellanox Technologies. All Rights Reserved.

Mellanox®, Mellanox logo, BridgeX®, ConnectX®, Connect-IB®, CoolBox®, CORE-Direct®, GPUDirect®, InfiniBridge®, InfiniHost®, InfiniScale®, Kotura®, Kotura logo, Mellanox Connect. Accelerate. Outperform logo, Mellanox Federal Systems®, Mellanox Open Ethernet®, MetroX®, MLNX-OS®, Open Ethernet logo, PhyX®, ScalableHPC®, SwitchX®, TestX®, The Generation of Open Ethernet logo, UFM®, Virtual Protocol Interconnect®, Voltaire® and Voltaire logo are registered trademarks of Mellanox Technologies, Ltd.

CyPU™, ExtendX™, FabricIT™, FPGADirect™, HPC-X™, Mellanox Care™, Mellanox CloudX™, Mellanox NEO™, Mellanox Open Ethernet™, Mellanox PeerDirect™, Mellanox Virtual Modular Switch™, MetroDX™, NVMeDirect™, StPU™, Spectrum™, Switch-IB™, Unbreakable-Link™ are trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

Table of Contents

| | |
|--|-----------|
| Table of Contents | 1 |
| List Of Tables | 3 |
| Release Update History | 5 |
| Chapter 1 Overview | 6 |
| 1.1 Main Features in This Release | 6 |
| 1.2 Content of Mellanox OFED for Linux | 6 |
| 1.3 Supported Platforms and Operating Systems | 7 |
| 1.3.1 Supported Hypervisors | 8 |
| 1.3.2 Supported Non-Linux Virtual Machines | 8 |
| 1.4 Hardware and Software Requirements | 8 |
| 1.5 Supported HCAs Firmware Versions | 9 |
| 1.6 Compatibility | 9 |
| 1.7 RoCE Modes Matrix | 10 |
| Chapter 2 Changes in Rev 2.4-1.0.0 From Rev 2.3-2.0.5 | 11 |
| Chapter 3 Bug Fixes History | 12 |
| Chapter 4 Known Issues | 16 |
| 4.1 IPoIB Known Issues | 16 |
| 4.2 Ethernet Known Issues | 18 |
| 4.3 General Known Issues | 19 |
| 4.4 VGT+ Known Issues | 20 |
| 4.5 eIPoIB Known Issues | 21 |
| 4.6 XRC Known Issues | 21 |
| 4.7 ABI Compatibility Known Issues | 22 |
| 4.8 System Time Known Issues | 22 |
| 4.9 ConnectX®-3 Adapter Cards Family Known Issues | 22 |
| 4.10 Verbs Known Issues | 22 |
| 4.11 Resiliency Known Issues | 23 |
| 4.12 Driver Start Known Issues | 23 |
| 4.13 Performance Tools Known Issues | 24 |
| 4.14 Performance Known Issues | 24 |
| 4.15 Connection Manager (CM) Known Issues | 25 |
| 4.16 SR-IOV Known Issues | 25 |
| 4.17 Port Type Management Known Issues | 27 |
| 4.18 Flow Steering Known Issues | 28 |
| 4.19 Quality of Service Known Issues | 28 |
| 4.20 Installation Known Issues | 28 |
| 4.21 Driver Upload Known Issues | 29 |
| 4.22 UEFI Secure Boot Known Issues | 29 |
| 4.23 Fork Support Known Issues | 29 |
| 4.24 iSCSI over IPoIB Known Issues | 30 |

| | | |
|------------------|--|-----------|
| 4.25 | MLNX_OFED Sources Known Issues | 30 |
| 4.26 | InfiniBand Utilities Known Issues | 30 |
| 4.27 | mlx5 Driver Known Issues | 30 |
| 4.28 | Ethernet Performance Counters Known Issues | 31 |
| 4.29 | Uplinks Known Issues | 31 |
| 4.30 | Tools Known Issues | 31 |
| 4.31 | Resources Limitation Known Issues | 32 |
| 4.32 | RoCE Known Issues | 32 |
| 4.33 | Storage Known Issues | 34 |
| 4.34 | SRP Known Issues | 34 |
| 4.35 | SRP Interop Known Issues | 34 |
| 4.36 | DDN Storage Fusion 10000 Target Known Issues | 34 |
| 4.37 | Oracle Sun ZFS Storage 7420 Known Issues | 34 |
| 4.38 | iSER Initiator Known Issues | 35 |
| 4.39 | iSER Target Known Issues | 35 |
| 4.40 | ZFS Appliance Known Issues | 35 |
| Chapter 5 | Change Log History | 36 |
| Chapter 6 | API Change Log History | 43 |

List Of Tables

| | | |
|-----------|---|----|
| Table 1: | Release Update History | 5 |
| Table 2: | Mellanox OFED for Linux Software Components | 6 |
| Table 3: | Supported Platforms and Operating Systems | 7 |
| Table 4: | Additional Software Packages | 8 |
| Table 5: | Supported HCAs Firmware Versions | 9 |
| Table 6: | MLNX_OFED Rev 2.4-1.0.4 Compatibility Matrix | 9 |
| Table 7: | RoCE Modes Matrix | 10 |
| Table 8: | Changes in v2.4-1.0.0 | 11 |
| Table 9: | Fixed Bugs List | 12 |
| Table 10: | IPoIB Known Issues | 16 |
| Table 11: | Ethernet Known Issues | 18 |
| Table 12: | General Known Issues | 19 |
| Table 13: | VGT+ Known Issues | 20 |
| Table 14: | eIPoIB Known Issues | 21 |
| Table 15: | XRC Known Issues | 21 |
| Table 16: | ABI Compatibility Known Issues | 22 |
| Table 17: | System Time Known Issues | 22 |
| Table 18: | ConnectX®-3 Adapter Cards Family Known Issues | 22 |
| Table 19: | Verbs Known Issues | 22 |
| Table 20: | Resiliency Known Issues | 23 |
| Table 21: | Driver Start Known Issues | 23 |
| Table 22: | Performance Tools Known Issues | 24 |
| Table 23: | Performance Known Issues | 24 |
| Table 24: | Connection Manager (CM) Known Issues | 25 |
| Table 25: | SR-IOV Known Issues | 25 |
| Table 26: | Port Type Management Known Issues | 27 |
| Table 27: | Flow Steering Known Issues | 28 |
| Table 28: | Quality of Service Known Issues | 28 |
| Table 29: | Installation Known Issues | 28 |
| Table 30: | Driver Upload Known Issues | 29 |
| Table 31: | UEFI Secure Boot Known Issues | 29 |
| Table 32: | Fork Support Known Issues | 29 |
| Table 33: | ISCSI over IPoIB Known Issues | 30 |
| Table 34: | MLNX_OFED Sources Known Issues | 30 |
| Table 35: | InfiniBand Utilities Known Issues | 30 |

| | | |
|-----------|--|----|
| Table 36: | mlx5 Driver Known Issues | 30 |
| Table 37: | Ethernet Performance Counters Known Issues | 31 |
| Table 38: | Uplinks Known Issues | 31 |
| Table 39: | Tools Known Issues. | 31 |
| Table 40: | Resources Limitation Known Issues | 32 |
| Table 41: | RoCE Known Issues | 32 |
| Table 42: | Storage Known Issues | 34 |
| Table 43: | SRP Known Issues. | 34 |
| Table 44: | SRP Interop Known Issues | 34 |
| Table 45: | DDN Storage Fusion 10000 Target Known Issues | 34 |
| Table 46: | Oracle Sun ZFS Storage 7420 Known Issues | 34 |
| Table 47: | iSER Initiator Known Issues | 35 |
| Table 48: | iSER Target Known Issues | 35 |
| Table 49: | ZFS Appliance Known Issues | 35 |
| Table 50: | Change Log History. | 36 |
| Table 51: | API Change Log History | 43 |

Release Update History

Table 1 - Release Update History

| Release | Date | Description |
|-----------|---------------|-----------------|
| 2.4-1.0.4 | March 12 2015 | Initial Release |

1 Overview

These are the release notes of Mellanox OFED for Linux Driver, Rev 2.4-1.0.4. Mellanox OFED is a single Virtual Protocol Interconnect (VPI) software stack and operates across all Mellanox network adapter solutions supporting the following uplinks to servers:

- 10, 20, 40 and 56 Gb/s InfiniBand (IB)
- 10, 40 and 56¹ Gb/s Ethernet
- 2.5 or 5.0 GT/s PCI Express 2.0
- 8 GT/s PCI Express 3.0

1.1 Main Features in This Release

MLNX_OFED Rev 2.4-1.0.4 provides the following features:

- Verbs performance optimization
- EQ pool management
- MAC Forwarding DataBase (FDB) status reporting
- Ethtool Settings
- Adaptive Interrupt Moderation Scheme
- RSS support for fragmented IP datagrams
- iSER Target driver compatible with upstream Linux IO Target

1.2 Content of Mellanox OFED for Linux

Mellanox OFED for Linux software contains the following components:

Table 2 - Mellanox OFED for Linux Software Components

| Components | Description |
|---------------------------|---|
| OpenFabrics core and ULPs | <ul style="list-style-type: none"> • InfiniBand and Ethernet HCA drivers (mlx4, mlx5) • core • Upper Layer Protocols: IPoIB, SRP, iSER and iSER Initiator and Target |
| OpenFabrics utilities | <ul style="list-style-type: none"> • OpenSM: IB Subnet Manager with Mellanox proprietary Adaptive Routing • Diagnostic tools • Performance tests |
| MPI | <ul style="list-style-type: none"> • OSU MPI (mvapich2-2.0) stack supporting the InfiniBand interface • Open MPI stack 1.6.5 and later supporting the InfiniBand interface • MPI benchmark tests (OSU benchmarks, Intel MPI benchmarks, Presta) |
| PGAS | <ul style="list-style-type: none"> • HPC-X OpenSHMEM v2.2 supporting InfiniBand, MXM and FCA • HPC-X UPC v2.2 supporting InfiniBand, MXM and FCA |
| HPC Acceleration packages | <ul style="list-style-type: none"> • Mellanox MXM v3.0 (p2p transport library acceleration over InfiniBand) • Mellanox FCA v2.5 (MPI/PGAS collective operations acceleration library over InfiniBand) • KNEM, Linux kernel module enabling high-performance intra-node MPI/PGAS communication for large messages |

1. 56 GbE is a Mellanox propriety link speed and can be achieved while connected to Mellanox SX10XX switch series

Table 2 - Mellanox OFED for Linux Software Components

| Components | Description |
|---|---|
| Extra packages | <ul style="list-style-type: none"> • ibutils2 • ibdump • MFT |
| Sources of all software modules (under conditions mentioned in the modules' LICENSE files) except for MFT, OpenSM plugins, ibutils2, and ibdump | |
| Documentation | |

1.3 Supported Platforms and Operating Systems

The following are the supported OSs in MLNX_OFED Rev 2.4-1.0.4:

Table 3 - Supported Platforms and Operating Systems

| Operating System | Platform |
|-----------------------------|----------------|
| RHEL/CentOS 6.2 | x86_64/PPC64 |
| RHEL/CentOS 6.3 | x86_64/PPC64 |
| RHEL/CentOS 6.4 | x86_64/PPC64 |
| RHEL/CentOS 6.5 | x86_64/PPC64 |
| RHEL/CentOS 6.6 | x86_64/PPC64 |
| RHEL/CentOS 7.0 | x86_64/PPC64 |
| RHEL/CentOS 7.1 | x86_64/PPC64 |
| SLES11 SP2 | x86_64/PPC64 |
| SLES11 SP3 | x86_64/PPC64 |
| SLES12 | x86_64 |
| OEL 6.4 | x86_64 |
| OEL 6.5 | x86_64 |
| OEL 6.6 | x86_64 |
| Citrix XenServer Host 6.2 | i686 |
| Fedora 19 | x86_64 |
| Fedora 21 | x86_64 |
| Ubuntu 12.04.4 | x86_64 |
| Ubuntu 14.04 | x86_64/PPC64le |
| Ubuntu 14.10 | x86_64/PPC64le |
| Debian 6.0.9 | x86_64 |
| Debian 7.5 | x86_64 |
| Debian 7.6 | x86_64 |
| kernel 3.10.28 ^a | |
| kernel 3.11.10 ^a | |
| kernel 3.14.3 ^a | |
| kernel 3.15 ^a | |
| kernel 3.16 ^a | |

Table 3 - Supported Platforms and Operating Systems

| Operating System | Platform |
|--------------------------|----------|
| kernel 3.17 ^a | |

a. This kernel is supported only when using the Operating Systems stated in the table above.



If you wish to install OFED on a different kernel, you need to create a new ISO image, using `mlnx_add_kernel_support.sh` script. See the MLNX_OFED User Guide for instructions.



Upgrading MLNX_OFED on your cluster requires upgrading all of its nodes to the newest version as well.

1.3.1 Supported Hypervisors

The following are the supported hypervisors in MLNX_OFED Rev 2.4-1.0.4:

- KVM: RedHat 6.4, 6.5, 7.0, Ubuntu 14.04, Sles11SP3
- Xen4.2
- XenServer6.2

1.3.2 Supported Non-Linux Virtual Machines

The following are the supported Non-Linux (InfiniBand only) Virtual Machines in MLNX_OFED Rev 2.4-1.0.4:

- Windows Server 2012 R2
- Windows Server 20012
- Windows Server 2008 R2

1.4 Hardware and Software Requirements

The following are the hardware and software requirements of MLNX_OFED Rev 2.4-1.0.4.

- Linux operating system
- Administrator privileges on your machine(s)
- Disk Space: 1GB

For the OFED Distribution to compile on your machine, some software packages of your operating system (OS) distribution are required.

To install the additional packages, run the following commands per OS:

Table 4 - Additional Software Packages

| Operating System | Required Packages Installation Command |
|---------------------|---|
| RHEL/OEL/ Fedora | <code>yum install perl pciutils python gcc-gfortran libxml2-python tclsh libnl.i686 libnl expat glib2 tcl libstdc++ bc tk gtk2 atk cairo numactl</code> |

Table 4 - Additional Software Packages

| Operating System | Required Packages Installation Command |
|------------------|---|
| XenServer | yum install perl pciutils python libxml2-python libnl expat glib2 tcl bc libstdc++ tk |
| SLES 11 SP2 | zypper install perl pciutils python libnl-32bit libxml2-python tclsh libnl libstdc++46 expat glib2 tcl bc tklibcurl4 gtk2 atk cairo |
| SLES 11 SP3 | zypper install perl pciutils python libnl-32bit libxml2-python tclsh libstdc++43 libnl expat glib2 tcl bc tk libcurl4 gtk2 atk cairo |
| Ubuntu/Debian | apt-get install perl dpkg autotools-dev autoconf libtool automake1.10 automake m4 dkms debhelper tcl tcl8.4 chrpath swig graphviz tcl-dev tcl8.4-dev tk-dev tk8.4-dev bison flex dpatch zlib1g-dev curl libcurl4-gnutls-dev python-libxml2 libvirt-bin libvirt0 libnl-dev libglib2.0-dev libgfortran3 automake m4 |

1.5 Supported HCAs Firmware Versions

MLNX_OFED Rev 2.4-1.0.4 supports the following Mellanox network adapter cards firmware versions:

Table 5 - Supported HCAs Firmware Versions

| HCA | Recommended Firmware Rev. | Additional Firmware Rev. Supported |
|-----------------|---------------------------|------------------------------------|
| Connect-IB® | Rev 10.10.5020 | Rev 10.10.4064 |
| ConnectX®-3 Pro | Rev 2.33.5100 | Rev 2.33.5000 |
| ConnectX®-3 | Rev 2.33.5100 | Rev 2.33.5000 |
| ConnectX®-2 | Rev 2.9.1000 | Rev 2.9.1000 |

For official firmware versions please see:

http://www.mellanox.com/content/pages.php?pg=firmware_download

1.6 Compatibility

MLNX_OFED Rev 2.4-1.0.4 is compatible with the following:

Table 6 - MLNX_OFED Rev 2.4-1.0.4 Compatibility Matrix

| Mellanox Product | Description/Version |
|-------------------------------|--|
| MLNX-OS® | MSX6036 w/w MLNX-OS® version 3.3.4304 ^a |
| Grid Director™ | 4036 w/w Grid Director™ version 3.9.1-985 |
| FabricIT™ EFM | IS5035 w/w FabricIT EFM version 1.1.3000 |
| FabricIT™ BXM | MBX5020 w/w FabricIT BXM version 2.1.2000 |
| Unified Fabric Manager (UFM®) | v4.8 |
| MXM | v3.2 |
| HPC-X UPC | v2.18.0 |
| HPC-X OpenSHMEM | v1.8.3 |

Table 6 - MLNX_OFED Rev 2.4-1.0.4 Compatibility Matrix

| Mellanox Product | Description/Version |
|------------------|---------------------|
| FCA | v2.5 and v3.1 |
| HPC-X MPI | v1.8.3 |
| MVAPICH | v2.0 |

- a. MLNX_OFED v2.3-1.0.1 was tested with this switch however, additional switches might be supported as well.

1.7 RoCE Modes Matrix

The following is RoCE modes matrix:

Table 7 - RoCE Modes Matrix

| Software Stack / Inbox Distribution | RoCE IP Based (Layer 2) Supported as of Version | RoCEv2 (Layer 3) Supported as of Version |
|-------------------------------------|---|--|
| MLNX_OFED | 2.1-x.x.x | 2.3-x.x.x |
| Kernel.org | 3.14 | |
| RHEL | 6.6; 7.0 | |
| SLES | 12 | |
| Ubuntu | 14.04 | |

2 Changes in Rev 2.4-1.0.0 From Rev 2.3-2.0.5

Table 8 - Changes in v2.4-1.0.0

| Category | Description |
|-----------|---|
| Bug Fixes | See “Bug Fixes History” on page 12. |

3 Bug Fixes History

Table 9 lists the bugs fixed in this release.

Table 9 - Fixed Bugs List

| # | Issue | Description | Discovered in Release | Fixed in Release |
|-----|---------------------|---|-----------------------|------------------|
| 1. | Security | CVE-2014-8159 Fix: Prevented integer overflow in IB-core module during memory registration. | 2.0-2.0.5 | 2.4-1.0.4 |
| 1. | mlx5_ib | Fixed the return value of max inline received size in the created QP. | 2.3-2.0.1 | 2.4-1.0.0 |
| 2. | | Resolved soft lock on massive amount of user memory registrations | 2.3-2.0.1 | 2.4-1.0.0 |
| 3. | InfiniBand Counters | Occasionally, port_rcv_data and port_xmit_data counters may not function properly. | 2.3-1.0.1 | 2.4-1.0.0 |
| 4. | mlx4_en | LRO fixes and improvements for jumbo MTU. | 2.3-2.0.1 | 2.4-1.0.0 |
| 5. | | Fixed a crash occurred when changing the number of rings (ethtool set-channels) when interface connected to netconsole. | 2.2-1.0.1 | 2.4-1.0.0 |
| 6. | | Fixed ping issues with IP fragmented datagrams in MTUs 1600-1700. | 2.2-1.0.1 | 2.4-1.0.0 |
| 7. | | The default priority to TC mapping assigns all priorities to TC0. This configuration achieves fairness in transmission between priorities but may cause undesirable PFC behavior where pause request for priority "n" affects all other priorities. | 2.3-1.0.1 | 2.4-1.0.0 |
| 8. | mlx5_ib | Fixed an issue related to large memory regions registration. The problem mainly occurred on PPC systems due to the large page size, and on non PPC systems with large pages (contig pages). | 2.3-2.0.1 | 2.3-2.0.5 |
| 9. | | Fixed an issue in verbs API: fallback to glibc on contiguous memory allocation failure | 2.3-2.0.1 | 2.3-2.0.5 |
| 10. | IPoIB | Fixed a memory corruption issue in multi-core system due to intensive IPoIB transmit operation. | 2.3-2.0.1 | 2.3-2.0.5 |
| 11. | IB MAD | Fixed an issue to prevent process starvation due to MAD packet storm. | 2.3-2.0.1 | 2.3-2.0.5 |

Table 9 - Fixed Bugs List

| # | Issue | Description | Discovered in Release | Fixed in Release |
|-----|-----------|--|-----------------------|------------------|
| 12. | IPoIB | Fixed an issue which prevented the spread of events among the closet NUMA CPU when only a single RX queue existed in the system. | 2.3-1.0.1 | 2.3-2.0.0 |
| 13. | | Returned the CQ to its original state (armed) to prevent traffic from stopping | 2.3-1.0.1 | 2.3-2.0.0 |
| 14. | | Fixed a TX timeout issue in CM mode, which occurred under heavy stress combined with ifup/ ifdown operation on the IPoIB interface. | 2.1-1.0.0 | 2.3-2.0.0 |
| 15. | mlx4_core | Fixed "sleeping while atomic" error occurred when the driver ran many firmware commands simultaneously. | 2.3-1.0.1 | 2.3-2.0.0 |
| 16. | mlx4_ib | Fixed an issue related to spreading of completion queues among multiple MSI-X vectors to allow better utilization of multiple cores. | 2.1-1.0.0 | 2.3-2.0.0 |
| 17. | | Fixed an issue that caused an application to fail when attaching Shared Memory. | 2.3-1.0.1 | 2.3-2.0.0 |
| 18. | mlx4_en | Fixed dmesg warnings: "NOHZ: local_soft-irq_pending 08". | 2.3-1.0.1 | 2.3-2.0.0 |
| 19. | | Fixed erratic report of hardware clock which caused bad report of PTP hardware Time Stamping. | 2.1-1.0.0 | 2.3-2.0.0 |
| 20. | mlx5_core | Fixed race when async events arrived during driver load. | 2.3-1.0.1 | 2.3-2.0.0 |
| 21. | | Fixed race in mlx5_eq_int when events arrived before eq->dev was set. | 2.3-1.0.1 | 2.3-2.0.0 |
| 22. | | Enabled all pending interrupt handlers completion before freeing EQ memory. | 2.3-1.0.1 | 2.3-2.0.0 |
| 23. | mlnx.conf | Defined mlnx.conf as a configuration file in mlnx-ofa_kernel RPM | 2.1-1.0.0 | 2.3-2.0.0 |
| 24. | SR-IOV | Fixed counter index allocation for VFs which enables Ethernet port statistics. | 2.3-1.0.1 | 2.3-2.0.0 |
| 25. | iSER | Fixed iSER DIX sporadic false DIF errors caused in large transfers when block merges were enabled. | 2.3-1.0.1 | 2.3-2.0.0 |
| 26. | RoCE v2 | RoCE v2 was non-functional on big Endian machines. | 2.3-1.0.1 | 2.3-2.0.0 |
| 27. | Verbs | Fixed registration memory failure when fork was enabled and contiguous pages or ODP were used. | 2.3-1.0.1 | 2.3-2.0.0 |

Table 9 - Fixed Bugs List

| # | Issue | Description | Discovered in Release | Fixed in Release |
|-----|-----------|---|-----------------------|------------------|
| 28. | IPoIB | Changing the GUID of a specific SR-IOV guest after the driver has been started, causes the ping to fail. Hence, no traffic can go over that InfiniBand interface. | 2.1-1.0.0 | 2.3-1.0.1 |
| 29. | XRC | XRC over ROCE in SR-IOV mode is not functional | 2.0-3.1.0 | 2.2-1.0.1 |
| 30. | mlx4_en | Fixed wrong calculation of packet true-size reporting in LRO flow. | 2.1-1.0.0 | 2.2-1.0.1 |
| 31. | | Fixed kernel panic on Debian-6.0.7 which occurred when the number of TX channels was set above the default value. | 2.1-1.0.0 | 2.2-1.0.1 |
| 32. | | Fixed a crash incidence which occurred when enabling Ethernet Time-stamping and running VLAN traffic. | 2.0-2.0.5 | 2.2-1.0.1 |
| 33. | IB Core | Fixed the QP attribute mask upon smac resolving | 2.1-1.0.0 | 2.1-1.0.6 |
| 34. | mlx5_ib | Fixed a send WQE overhead issue | 2.1-1.0.0 | 2.1-1.0.6 |
| 35. | | Fixed a NULL pointer de-reference on the debug print | 2.1-1.0.0 | 2.1-1.0.6 |
| 36. | | Fixed arguments to kzalloc | 2.1-1.0.0 | 2.1-1.0.6 |
| 37. | mlx4_core | Fixed the locks around completion handler | 2.1-1.0.0 | 2.1-1.0.6 |
| 38. | mlx4_core | Restored port types as they were when recovering from an internal error. | 2.0-2.0.5 | 2.1-1.0.0 |
| 39. | | Added an N/A port type to support port_type_array module param in an HCA with a single port | 2.0-2.0.5 | 2.1-1.0.0 |
| 40. | SR-IOV | Fixed memory leak in SR-IOV flow. | 2.0-2.0.5 | 2.0-3.0.0 |
| 41. | | Fixed communication channel being stuck | 2.0-2.0.5 | 2.0-3.0.0 |
| 42. | mlx4_en | Fixed ALB bonding mode failure when enslaving Mellanox interfaces | 2.0-3.0.0 | 2.1-1.0.0 |
| 43. | | Fixed leak of mapped memory | 2.0-3.0.0 | 2.1-1.0.0 |
| 44. | | Fixed TX timeout in Ethernet driver. | 2.0-2.0.5 | 2.0-3.0.0 |
| 45. | | Fixed ethtool stats report for Virtual Functions. | 2.0-2.0.5 | 2.0-3.0.0 |
| 46. | | Fixed an issue of VLAN traffic over Virtual Machine in paravirtualized mode. | 2.0-2.0.5 | 2.0-3.0.0 |
| 47. | | Fixed ethtool operation crash while interface down. | 2.0-2.0.5 | 2.0-3.0.0 |

Table 9 - Fixed Bugs List

| # | Issue | Description | Discovered in Release | Fixed in Release |
|-----|-------|--|-----------------------|------------------|
| 48. | IPoIB | Fixed memory leak in Connected mode. | 2.0-2.0.5 | 2.0-3.0.0 |
| 49. | | Fixed an issue causing IPoIB to avoid pkey value 0 for child interfaces. | 2.0-2.0.5 | 2.0-3.0.0 |

4 Known Issues

The following is a list of general limitations and known issues of the various components of this Mellanox OFED for Linux release.

4.1 IPoIB Known Issues

Table 10 - IPoIB Known Issues

| Index | Description | Workaround |
|-------|---|--|
| 1. | When user increases receive/send a buffer, it might consume all the memory when few child's interfaces are created. | - |
| 2. | The size of send queue in Connect-IB® cards cannot exceed 1K. | - |
| 3. | In 32 bit devices, the maximum number of child interfaces that can be created is 16. Creating more that, might cause out-of-memory issues. | - |
| 4. | In RHEL7.0, the Network-Manager can detect when the carrier of one of the IPoIB interfaces is OFF and can decide to disable its IP address. | Set "ignore-carrier" for the corresponding device in NetworkManager.conf. For further information, please refer to " <i>man NetworkManager.conf</i> " |
| 5. | IPoIB interface does not function properly if a third party application changes the PKey table. We recommend modifying PKey tables via OpenSM. | - |
| 6. | Fallback to the primary slave of an IPoIB bond does not work with ARP monitoring. (https://bugs.openfabrics.org/show_bug.cgi?id=1990) | - |
| 7. | Out-of memory issue might occur due to overload of interfaces created. | To calculate the allowed memory per each IPoIB interface check the following: <ul style="list-style-type: none"> • Num-rings = min(num-cores-on-that-device, 16) • Ring-size = 512 (by default, it is module parameter) • UD memory: 2 * num-rings * ring-size * 8K • CM memory: ring-size * 64k • Total memory = UD mem + CM mem |
| 8. | Connect-IB does not reach the bidirectional line rate | Optimize the IPoIB performance in Connect-IB: <pre>cat /sys/class/net/<interface>/device/local_cpus > /sys/class/net/<interface>/queues/rx-0/rps_cpus</pre> |
| 9. | If the CONNECTED_MODE=no parameter is set to "no" or missing from the ifcfg file for Connect-IB® IPoIB interface then the "service network restart" will hang. | Set the CONNECTED_MODE=yes parameter in the ifcfg file for Connect-IB® interface. |
| 10. | Joining a multicast group in the SM using the RDMA_CM API requires IPoIB to first join the broadcast group. | - |

Table 10 - IPoIB Known Issues (Continued)

| Index | Description | Workaround |
|-------|--|---|
| 11. | <p>Whenever the IOMMU parameter is enabled in the kernel it can decrease the number of child interfaces on the device according to resource limitation. The driver will stuck after unknown amount of child interfaces creation.</p> <p>For further information, please see: https://access.redhat.com/site/articles/66747 http://support.citrix.com/article/CTX136517 http://www.novell.com/support/kb/doc.php?id=7012337 https://bugzilla.redhat.com/show_bug.cgi?id=1044595</p> | <p>To avoid such issue:</p> <ul style="list-style-type: none"> • Decrease the amount of the RX receive buffers (module parameter, the default is 512) • Decrease the number of RX rings (sys/fs or ethtool in new kernels) • Avoid using IOMMU if not required <p>For KVM users: Run: <pre>echo 1 > /sys/module/kvm/parameters/allow_unsafe_assigned_interrupts</pre></p> <p>To make this change persist across reboots, add the following to the <code>/etc/modprobe.d/kvm.conf</code> file (or create this file, if it does not exist): <pre>options kvm allow_unsafe_assigned_interrupts=1 kernel parameters</pre></p> |
| 12. | <p>System might crash in <code>skb_checksum_help()</code> while performing TCP retransmit involving packets with 64k packet size.</p> <p>A similar out to the below will be printed: kernel BUG at net/core/dev.c:1707! invalid opcode: 0000 [#1] SMP RIP: 0010:[<ffffffff81448988>] skb_checksum_help+0x148/0x160 Call Trace: <IRQ> [<ffffffff81448d83>] dev_hard_start_xmit+0x3e3/0x530 [<ffffffff8144c805>] dev_queue_xmit+0x205/0x550 [<ffffffff8145247d>] neigh_connected_output+0xbd/0x1 </p> | Use UD mode in IPoIB |
| 13. | When InfiniBand ports are removed from the host (e.g when changing port type from IB to Eth or removing a card from the PCI bus) the remaining IPoIB interface might be renamed. | <p>To avoid it and have persistent IPoIB network devices names for ConnectX ports, add to the <code>/etc/udev/rules.d/70-persistent-net.rules</code> file:</p> <pre>SUBSYSTEM=="net", ACTION=="add", DRIVERS=="?*", ATTR{address}=="*<Port GID>", NAME="ibN"</pre> <p>Where N is the IPoIB required interface index</p> |
| 14. | After releasing a bond interface that contains IPoIB slaves, a call trace might be printed into the <code>dmesg</code> . | - |

Table 10 - IPoIB Known Issues (Continued)

| Index | Description | Workaround |
|-------|--|---|
| 15. | IPoIB interfaces are loaded without an IP address on SLES 12. | <ol style="list-style-type: none"> 1. Open the "/etc/wicked/common.xml" file. 2. Change: <pre>"<use-nanny>>false</use-nanny>" to "<use-nanny>>true</use-nanny>"</pre> 3. Run: <pre># systemctl restart wickedd.ser- vice wicked # ifup all</pre> |
| 16. | In RHEL7.0, running ifdown then ifup on an interface after changing CONNECTED_MODE in its ifcfg file, will cause the interface bring up to fail. | Reload the driver "/etc/init.d/openibd restart" or reboot the system. |
| 17. | Clone interfaces receive a duplicated IPv6 address when a child interface with the same PKey (a.k.a clone interface) is used for all the clones. | - |
| 18. | eth_ipoib module is not loaded after reloading the ib_ipoib module. | To restart the IPoIB driver, run "/etc/init.d/openibd restart". Do not restart it by manually restarting each module. |
| 19. | In Ubuntu and Debian, the default of the recv_queue_size and send_queue_size is 128 according to the io_mmu issue. | - |

4.2 Ethernet Known Issues

Table 11 - Ethernet Known Issues

| Index | Description | Workaround |
|-------|--|--|
| 1. | <p>When creating more than 125 VLANs and SR-IOV mode is enabled, a kernel warning message will be printed indicating that the native VLAN is created but will not work with RoCE traffic.</p> <pre>kernel warning: mlx4_core 0000:07:00.0: vhcr command ALLOC_RES (0xf00) slave:0 in_param 0x7e in_mod=0x107, op_mod=0x1 failed with error:0, status -28</pre> | - |
| 2. | Kernel panic might occur during FIO splice in kernels before 2.6.34-rc4. | Use kernel v2.6.34-rc4 which provides the following solution: baff42a net: Fix oops from tcp_collapse() when using splice() |
| 3. | In PPC systems when QoS is enabled a harmless Kernel DMA mapping error messages might appear in kernel log (IOMMU related issue). | - |
| 4. | Transmit timeout might occur on RH6.3 as a result of lost interrupt (OS issue). In this case, the following message will be shown in dmesg: do_IRQ: 0.203 No irq handler for vector (irq -1) | - |

Table 11 - Ethernet Known Issues (Continued)

| Index | Description | Workaround |
|-------|---|--|
| 5. | Mixing ETS and strict QoS policies for TCs in 40GbE ports may cause inaccurate results in bandwidth division among TCs. | - |
| 6. | Creating a VLAN with user priority ≥ 4 on ConnectX®-2 HCA is not supported. | - |
| 7. | Affinity hints are not supported in Xen Hypervisor (an irqblancer issue). This causes a non-optimal IRQ affinity. | To overcome this issues, run: <code>set_irq_affinity.sh eth<x></code> |
| 8. | Reboot might hang in SR-IOV when using the "probe_vf" parameter with many Virtual Functions. The following message is logged in the kernel log: "waiting for eth to become free. Usage count =1" | - |
| 9. | In ConnectX®-2, RoCE UD QP does not include VLAN tags in the Ethernet header | |
| 10. | VXLAN may not be functional when configured over Linux bridge in RH7.0 or Ubuntu14.04. The issue is within the bridge modules in those kernels. In Vanilla kernels above 3.16 issues is fixed. | - |
| 11. | In RH6.4, ping may not work over VLANs that are configured over Linux bridge when the bridge has a mlx4_en interface attached to it. | - |
| 12. | The interfaces LRO needs to be set to "OFF" manually when there is a bond configured on Mellanox interfaces with a Bridge over that bond. | Run: <code>ethtool -K ethX lro off</code> |

4.3 General Known Issues

Table 12 - General Known Issues

| Index | Description | Workaround |
|-------|--|---|
| 1. | On ConnectX-2/ConnectX-3 Ethernet adapter cards, there is a mismatch between the GUID value returned by firmware management tools and that returned by fabric/driver utilities that read the GUID via device firmware (e.g., using ibstat). Mlxburn/flint return 0xffff as GUID while the utilities return a value derived from the MAC address. For all driver/firmware/software purposes, the latter value should be used. | N/A. Please use the GUID value returned by the fabric/driver utilities (not 0xffff). |

4.4 VGT+ Known Issues

Table 13 - VGT+ Known Issues

| Index | Description | Workaround |
|-------|--|------------|
| 1. | Before adding a VLAN on the VM, the parent interface should be brought up. Otherwise VLAN creation will fail and the following message will be presented: "Fail to register network rule." | - |
| 2. | Bringing down and up the parent interface with VLANs configured over it, may result in traffic over VLANs being lost. | |
| 3. | On some of the OSes the callback <code>ndo_vlan_rx_add/kill_vid</code> returns void, therefore <code>ethX.Y</code> is created. However, if VLAN Y is not listed in the set of the allowed VLANs, no traffic will pass. | |
| 4. | When un-tagged traffic is not allowed, the below message will appear on Dom0 after driver restart on DomU: <code>mlx4_core 0000:16:00.0: vhcr command ALLOC_RES (0xf00) slave:1 in_param 0x0 in_mod=0x207, op_mod=0x1 failed with error:0, status -1</code> | - |

4.5 eIPoIB Known Issues

Table 14 - eIPoIB Known Issues

| Index | Description | Workaround |
|-------|--|---|
| 1. | On rare occasions, upon driver restart the following message is shown in the dmesg: 'cannot create duplicate filename '/class/net/eth_ipoib_interfaces' | - |
| 2. | No indication is received when eIPoIB is non functional. | Run 'ps -ef grep ipoibd' to verify its functionality. |
| 3. | eIPoIB requires libvirtd, python | - |
| 4. | eIPoIB supports only active-backup mode for bonding. | - |
| 5. | eIPoIB supports only VLAN Switch Tagging (VST) mode on guests. | - |
| 6. | IPv6 is currently not supported in eIPoIB | - |
| 7. | eIPoIB cannot run when Flow Steering is enabled | - |

4.6 XRC Known Issues

Table 15 - XRC Known Issues

| Index | Description | Workaround |
|-------|---|----------------------|
| 1. | Legacy API is deprecated, thus when recompiling applications over MLNX_OFED v2.0-3.x.x, warnings such as the below are displayed. rdma.c:1699: warning: 'ibv_open_xrc_domain' is deprecated (declared at /usr/include/infiniband/ofa_verbs.h:72) rdma.c:1706: warning: 'ibv_create_xrc_srq' is deprecated (declared at /usr/include/infiniband/ofa_verbs.h:89) These warnings can be safely ignored. | - |
| 2. | XRC is not functional in heterogeneous clusters containing non Mellanox HCAs. | - |
| 3. | XRC options do not work when using qperf tool. | Use perftest instead |
| 4. | Out-of memory issue might occur due to overload of XRC receive QP with non zero receive queue size created. XRC QPs do not have receive queues. | - |

4.7 ABI Compatibility Known Issues

Table 16 - ABI Compatibility Known Issues

| Index | Description | Workaround |
|-------|---|--|
| 1. | MLNX_OFED v2.3-1.0.1 is not ABI compatible with previous MLNX_OFED/OFED versions. | Recompile the application over the new MLNX_OFED version |

4.8 System Time Known Issues

Table 17 - System Time Known Issues

| Index | Description | Workaround |
|-------|--|------------|
| 1. | Loading the driver using the openibd script when no InfiniBand vendor module is selected (for example <code>mlx4_ib</code>), may cause the execution of the <code>/sbin/start_udev</code> script. In RedHat 6.x and OEL6.x this may change the local system time. | - |

4.9 ConnectX®-3 Adapter Cards Family Known Issues

Table 18 - ConnectX®-3 Adapter Cards Family Known Issues

| Index | Description | Workaround |
|-------|--|---|
| 1. | Using RDMA READ with a higher value than 30 SGEs in the WR might lead to "local length error". | Do not set the value of SGEs higher than 30 when RDMA READ is used. |

4.10 Verbs Known Issues

Table 19 - Verbs Known Issues

| Index | Description | Workaround |
|-------|---|---|
| 1. | Using <code>libnl1_1_3~26</code> or earlier, requires <code>ibv_create_ah</code> protection by a lock for multi-threaded applications. | - |
| 2. | In MLNX_OFED v2.4-1.0.0, if several CQEs are received on a CQ, they will be coalesced and a user-space event will be triggered only once. | When getting an event, poll the CQ until it is empty. |

4.11 Resiliency Known Issues

Table 20 - Resiliency Known Issues

| Index | Description | Workaround |
|-------|---|--|
| 1. | Reset Flow can run on XenServer 6 only after the active user space applications running verbs are terminated. | - |
| 2. | SR-IOV non persistent configuration (such as VGT, VST, Host assigned GUIDs, and QP0-enabled VFs) may be lost upon Reset Flow. | Reset Admin configuration post Reset Flow |
| 3. | Upon Reset Flow or after running restart driver, Ethernet VLANs are lost. | Reset the VLANs using the <code>ifup</code> command. |
| 4. | Restarting the driver or running <code>connectx_port_config</code> when Reset Flow is running might result in a kernel panic | - |
| 5. | Networking configuration (e.g. VLANs, IPv6) should be statically defined in order to have them set after Reset Flow as of after restart driver. | - |
| 6. | After recovering from an EEH event, <code>mlx5_core/</code> <code>mlx4_core</code> unload may fail due to a bug in some kernel versions. The bug is fixed in Kernel 3.15 | - |

4.12 Driver Start Known Issues

Table 21 - Driver Start Known Issues

| Index | Description | Workaround |
|-------|---|--|
| 1. | "Out-of-memory" issues may rise during drivers load depending on the values of the driver module parameters set (e.g. <code>log_num_cq</code>). | - |
| 2. | When reloading/starting the driver using the <code>/etc/init.d/openibd</code> the following messages are displayed if there is a third party RPM or driver installed: "Module <code>mlx4_core</code> does not belong to <code>MLNX_OFED</code> " or "Module <code>mlx4_core</code> belong to <code><rpm name></code> which is not a part of <code>MLNX_OFED</code> " | Remove the third party RPM/non <code>MLNX_OFED</code> drivers directory, run: " <code>depmod</code> " and then rerun " <code>/etc/init.d/openibd restart</code> " |
| 3. | Occasionally, when trying to repetitively reload the NES hardware driver on SLES11 SP2, a soft lockups occurs that required reboot. | - |

Table 21 - Driver Start Known Issues (Continued)

| Index | Description | Workaround |
|-------|--|--|
| 4. | In ConnectX-2, if the driver load succeeds, the informative message below is presented conveying the below limitations: <ul style="list-style-type: none"> If port type is IB the number of maximum supported VLs is 4 If port type is ETH then the maximum priority for VLAN tagged is 3 <pre>"mlx4_core 0000:0d:00.0: command SET_PORT (0xc) failed: in_param=0x120064000, in_mod=0x2, op_mod=0x0, fw status = 0x40"</pre> | - |
| 5. | “openibd start” unloads kernel modules that were loaded from initrd/initramfs upon boot. This affects only kernel modules which come with MLNX_OFED and are included in initrd/initramfs. | - |
| 6. | If a Lustre storage is used, it must be fully unloaded before restarting the driver or rebooting the machine, otherwise machine might get stuck/panic. | <ol style="list-style-type: none"> Unmount any mounted Lustre storages: # umount<lustre_mount_point> Unload all Lustre modules: # lustre_rmmod |

4.13 Performance Tools Known Issues

Table 22 - Performance Tools Known Issues

| Index | Description | Workaround |
|-------|---|------------|
| 1. | perftest package in MLNX_OFED v2.2-1.0.1 and onwards does not work with older versions of the driver. | - |

4.14 Performance Known Issues

Table 23 - Performance Known Issues

| Index | Description | Workaround |
|-------|---|---|
| 1. | On machines with irqbalancer daemon turned off, the default InfiniBand interrupts will be routed to a single core which may cause overload and software/hardware lockups. | Execute the following script as root: set_irq_affinity.sh <interface or IB device> [2nd interface or IB device] |
| 2. | Out of the box throughput performance in Ubuntu 14.04 is not optimal and may achieve results below the line rate in 40GE link speed. | For additional performance tuning, please refer to Performance Tuning Guide. |
| 3. | UDP receiver throughput may be lower than expected, when running over mlx4_en Ethernet driver. This is caused by the adaptive interrupt moderation routine, which sets high values of interrupt coalescing, causing the driver to process large number of packets in the same interrupt, leading UDP to drop packets due to overflow in its buffers. | Disable adaptive interrupt moderation and set lower values for the interrupt coalescing manually. <pre>ethtool -C <eth>X adaptive-rx off rx-usecs 64 rx-frames 24</pre> Values above may need tuning, depending the system, configuration and link speed. |

Table 23 - Performance Known Issues (Continued)

| Index | Description | Workaround |
|-------|--|------------|
| 4. | Performance degradation might occur when bonding Ethernet interfaces on Centos 6.5 | - |

4.15 Connection Manager (CM) Known Issues

Table 24 - Connection Manager (CM) Known Issues

| Index | Description | Workaround |
|-------|--|-------------------------------------|
| 1. | When 2 different ports have identical GIDs, the CM might send its packets on the wrong port. | All ports must have different GIDs. |

4.16 SR-IOV Known Issues

Table 25 - SR-IOV Known Issues

| Index | Description | Workaround |
|-------|---|--|
| 1. | When using legacy VMs with MLNX_OFED 2.x hypervisor, you may need to set the 'enable_64b_cqe_eqe' parameter to zero on the hypervisor. It should be set in the same way that other module parameters are set for mlx4_core at module load time. For example, add "options mlx4_core enable_64b_cqe_eqe=0" as a line in the file /etc/modprobe.d/mlx4_core.conf. | - |
| 2. | InfiniBand counters are not available in the VM. | - |
| 3. | mlx4_port1_mtu sysfs entry shows a wrong MTU number in the VM. | - |
| 4. | When at least one port is configured as InfiniBand, and the num_vfs is provided but the probe_vf is not, HCA initialization fails. | Use both the num_vfs and the probe_vf in the modprobe line. |
| 5. | When working with a bonding device to enslave the Ethernet devices in active-backup mode and failover MAC policy in a Virtual Machine (VM), establishment of RoCE connections may fail. | Unload the module mlx4_ib and reload it in the VM. |
| 6. | Attaching or detaching a Virtual Function on SLES11 SP3 to a guest Virtual Machine while the mlx4_core driver is loaded in the Virtual Machine may cause a kernel panic in the hypervisor. | Unload the mlx4_core module in the hypervisor before attaching or detaching a function to or from the guest. |
| 7. | When detaching a VF without shutting down the driver from a VM and reattaching it to another VM with the same IP address for the Mellanox NIC, RoCE connections will fail | Shut down the driver in the VM before detaching the VF. |

Table 25 - SR-IOV Known Issues (Continued)

| Index | Description | Workaround |
|-------|---|--|
| 8. | Enabling SR-IOV requires appending the "intel_iommu=on" option to the relevant OS in file /boot/grub/grub.conf. Without that SR-IOV cannot be loaded. | - |
| 9. | On various combinations of Hypervisor/OSes and Guest/OSes, an issue might occur when attaching/detaching VFs to a guest while that guest is up and running. | Attach/detach VFs to/from a VM only while that VM is down. |
| 10. | When working with SR-IOV in Xen-4.2 virtualization platform, only the built-in xen_pciback driver should be loaded. The xen_pciback module in dom0 should not be loaded, as loading them simultaneously may cause interrupts loss and cause the driver to enter the reset flow. | - |
| 11. | The known PCI BDFs for all VFs in kernel command line should be specified by adding xen-pci-back.hide For further information, please refer to http://wiki.xen.org/wiki/Xen_PCI_Passthrough | - |
| 12. | The qemu version (2.0) provided in box with Ubuntu 14.04 does not work properly when more than 2 VMs are run over an Ubuntu 14.04 Hypervisor. | - |
| 13. | SR-IOV UD QPs are forced by the Hypervisor to use the base GID (i.e., the GID that the VF sees in its GID entry at its paravirtualized index 0). This is needed for security, since UD QPs use Address Vectors, and any GID index may be placed in such a vector, including indices not belonging to that VF. | - |
| 14. | Attempting to attach a PF to a VM when SR-IOV is already enabled on that PF may result in a kernel panic. | - |
| 15. | osmtest on the Hypervisor fails when SR-IOV is enabled. However, only the test fails, OpenSM will operate correctly with the host. The failure reason is that if an mcg is already joined by the host, a subsequent join request for that group succeeds automatically (even if the join parameters in the request are not correct). This success does no harm. | - |
| 16. | If a VM does not support PCI hot plug, detaching an mlx4 VF and probing it to the hypervisor may cause the hypervisor to crash. | - |

4.17 Port Type Management Known Issues

Table 26 - Port Type Management Known Issues

| Index | Description | Workaround |
|-------|---|---|
| 1. | OpenSM must be stopped prior to changing the port protocol from InfiniBand to Ethernet. | - |
| 2. | After changing port type using <code>connectx_port_config</code> interface ports' names can be changed. For example. <code>ib1</code> -> <code>ib0</code> if port1 changed to be Ethernet port and port2 left IB. | Use udev rules for persistent naming configuration. For further information, please refer to the User Manual |
| 3. | A working IP connectivity between the RoCE devices is required when creating an address handle or modifying a QP with an address vector. | - |
| 4. | IPv4 multicast over RoCE requires the MGID format to be as follow <code>::ffff:<Multicast IPv4 Address></code> | - |
| 5. | IP routable RoCE does not support Multicast Listener Discovery (MLD) therefore, multicast traffic over IPv6 may not work as expected. | - |
| 6 | DIF: When running IO over FS over DM during unstable ports, block layer BIOs merges may cause false DIF error. | - |

4.18 Flow Steering Known Issues

Table 27 - Flow Steering Known Issues

| Index | Description | Workaround |
|-------|--|---|
| 1. | Flow Steering is disabled by default in firmware version < 2.32.5100. | To enable it, set the parameter below as follow: log_num_mgm_entry_size should set to -1 |
| 2. | IPv4 rule with source IP cannot be created in SLES 11.x OSes. | - |
| 3. | RFS does not support UDP. | - |
| 4. | When working in DMFS:A0 mode and VM/hypervisor is MLNX_OFED 2.3-x.x.x, the second side (hypervisor/VM respectively) should be MLNX_OFED 2.3-x.x.x as well. | - |

4.19 Quality of Service Known Issues

Table 28 - Quality of Service Known Issues

| Index | Description | Workaround |
|-------|---|------------|
| 1. | QoS is not supported in XenServer, Debian 6.0 and 6.2 with uek kernel | - |
| 2. | When QoS features are not supported by the kernel, mlnx_qos tool may crash. | - |

4.20 Installation Known Issues

Table 29 - Installation Known Issues

| Index | Description | Workaround |
|-------|---|---|
| 1. | When upgrading from an earlier Mellanox OFED version, the installation script does not stop the earlier version prior to uninstalling it. | Stop the old OFED stack (/etc/init.d/openibd stop) before upgrading to this new version. |
| 2. | Upgrading from the previous OFED installation to this release, does not unload the kernel module ipoib_helper. | Reboot after installing the driver. |
| 3. | Installation using Yum does not update HCA firmware. | See "Updating Firmware After Installation" in OFED User Manual |
| 4. | "--total-vfs <0-63>" installation parameter is no longer supported | Use '--enable-sriov' installation parameter to burn firmware with SR-IOV support. The number of virtual functions (VFs) will be set to 16. For further information, please refer to the User Manual. |
| 5. | When using bonding on Ubuntu OS, the "ifenslave" package must be installed. | - |
| 6. | On PPC systems, the ib_srp module is not installed by default since it breaks the ibmvscsi module. | If your system does not require the ibmvscsi module, run the mlnxofedinstall script with the "--with-srp" flag. |

4.21 Driver Upload Known Issues

Table 30 - Driver Upload Known Issues

| Index | Description | Workaround |
|-------|--|-----------------------|
| 1. | "openibd stop" can sometime fail with the error: Unloading ib_cm [FAILED] ERROR: Module ib_cm is in use by ib_i- poib | Re-run "openibd stop" |

4.22 UEFI Secure Boot Known Issues

Table 31 - UEFI Secure Boot Known Issues

| Index | Description | Workaround |
|-------|---|---|
| 1. | On RHEL7 and SLES12, the following error is displayed in dmesg if the Mellanox's x.509 Public Key is not added to the system: [4671958.383506] Request for unknown module key 'Mellanox Technologies signing key: 61feb074fc7292f958419386ffdd9d5-ca999e403' err -11 This error can be safely ignored as long as Secure Boot is disabled on the system. | For further information, please refer to the User Manual section "Enrolling Mellanox's x.509 Public Key On your Systems". |
| 2 | Ubuntu12 requires update of user space open-iscsi to v2.0.873 | - |
| 3 | The initiator does not respect interface parameter while logging in. | Configure each interface on a different subnet. |

4.23 Fork Support Known Issues

Table 32 - Fork Support Known Issues

| Index | Description | Workaround |
|-------|--|------------|
| 1. | Fork support from kernel 2.6.12 and above is available provided that applications do not use threads. <code>fork()</code> is supported as long as the parent process does not run before the child exits or calls <code>exec()</code> . The former can be achieved by calling <code>wait(childpid)</code> , and the latter can be achieved by application specific means. The Posix <code>system()</code> call is supported. | - |

4.24 ISCSI over IPoIB Known Issues

Table 33 - ISCSI over IPoIB Known Issues

| Index | Description | Workaround |
|-------|--|------------|
| 1. | When working with ISCSI over IPoIB, LRO must be disabled (even if IPoIB is set to connected mode) due to a bug in older kernels which causes a kernel panic. | - |

4.25 MLNX_OFED Sources Known Issues

Table 34 - MLNX_OFED Sources Known Issues

| Index | Description | Workaround |
|-------|---|------------|
| 1. | MLNX_OFED includes the OFED source RPM packages used as a build platform for kernel code but does not include the sources of Mellanox proprietary packages. | - |

4.26 InfiniBand Utilities Known Issues

Table 35 - InfiniBand Utilities Known Issues

| Index | Description | Workaround |
|-------|---|--|
| 1. | When running the <code>ibdiagnet check nodes_info</code> on the fabric, a warning specifying that the card does not support general info capabilities for all the HCAs in the fabric will be displayed. | Run <code>ibdiagnet --skip nodes_info</code> |
| 2. | <code>ibdump</code> does not work when IPoIB device managed Flow Steering is OFF and at least one of the ports is configured as InfiniBand. | Enable IPoIB Flow Steering and restart the driver. For further information, please refer to MLNX_OFED User Manual section Enable/Disable Flow Steering. |

4.27 mlx5 Driver Known Issues

Table 36 - mlx5 Driver Known Issues

| Index | Description | Workaround |
|-------|---|---|
| 1. | Atomic Operations in Connect-IB® are fully supported on big-endian machines (e.g. PPC). Their support is limited on little-endian machines (e.g. x86) | - |
| 2. | EEH events that arrive while the mlx5 driver is loading may cause the driver to hang. | - |
| 3. | The mlx5 driver can handle up to 5 EEH events per hour. | If more events are received, cold reboot the machine. |

4.28 Ethernet Performance Counters Known Issues

Table 37 - Ethernet Performance Counters Known Issues

| Index | Description | Workaround |
|-------|---|------------|
| 1. | In a system with more than 61 VFs, the 62nd VF and onwards is assigned with the SINKQP counter, and as a result will have no statistics, and loopback prevention functionality for SINK counter. | - |
| 2. | Since each VF tries to allocate 2 more QP counter for its RoCE traffic statistics, in a system with less than 61 VFs, if there is free resources it receives new counter otherwise receives the default counter which is shared with Ethernet. In this case RoCE statistics is not available. | - |
| 3. | In ConnectX®-3, when we enable function-based loopback prevention for Ethernet port by default (i.e., based on the QP counter index), the dropped self-loopback packets increase the IfRxErrorFrames/Octets counters. | - |

4.29 Uplinks Known Issues

Table 38 - Uplinks Known Issues

| Index | Description | Workaround |
|-------|---|--------------------|
| 1. | On rare occasions, ConnectX®-3 Pro adapter card may fail to link up when performing parallel detect to 40GbE. | Restart the driver |

4.30 Tools Known Issues

Table 39 - Tools Known Issues

| Index | Description | Workaround |
|-------|--|--|
| 1. | Running ibdump in InfiniBand mode with firmware older than v2.33.5000, may cause the server to hang due to a firmware issue. | Run ibdump with firmware v2.33.5000 and higher |

4.31 Resources Limitation Known Issues

Table 40 - Resources Limitation Known Issues

| Index | Description | Workaround |
|-------|---|--|
| 1. | The device capabilities reported may not be reached as it depends on the system on which the device is installed and whether the resource is allocated in the kernel or the userspace. | - |
| 2. | mlx4_core can allocate up to 64 MSI-X vectors, an MSI-X vector per CPU. | - |
| 3. | Setting more IP addresses than the available GID entries in the table results in failure and the "update_gid_table error message is displayed: GID table of port 1 is full. Can't add <address>" message. | - |
| 4. | Registering a large amount of Memory Regions (MR) may fail because of DMA mapping issues on RHEL 7.0. | - |
| 5. | Occasionally, a user process might experience some memory shortage and not function properly due to Linux kernel occupation of the system's free memory for its internal cache. | <p>To free memory to allow it to be allocated in a user process, run the <code>drop_caches</code> procedure below.</p> <p>Performing the following steps will cause the kernel to flush and free pages, dentries and inodes caches from memory, causing that memory to become free.</p> <p>Note: As this is a non-destructive operation and dirty objects are not freeable, run <code>`sync'</code> first.</p> <ul style="list-style-type: none"> To free the pagecache: <code>echo 1 > /proc/sys/vm/drop_caches</code> To free dentries and inodes: <code>echo 2 > /proc/sys/vm/drop_caches</code> To free pagecache, dentries and inodes: <code>echo 3 > /proc/sys/vm/drop_caches</code> |

4.32 RoCE Known Issues

Table 41 - RoCE Known Issues

| Index | Description | Workaround |
|-------|--|--------------------|
| 1. | Not configuring the Ethernet devices or independent VMs with a unique IP address in the physical port, may result in RoCE GID table corruption. | Restart the driver |
| 2. | If RDMA_CM is not used for connection management, then the source and destination GIDs used to modify a QP or create AH should be of the same type - IPv4 or IPv6. | - |

Table 41 - RoCE Known Issues (Continued)

| Index | Description | Workaround |
|-------|--|------------|
| 3. | On rare occasions, the driver reports a wrong GID table (read from <code>/sys/class/infiniband/mlx4_*/ports/*/gids/*</code>). This may cause communication problems. | - |
| 4. | MLNX_OFED v2.1-1.0.0 and onwards is not interoperable with older versions of MLNX_OFED. | - |
| 5. | <p>Since the number of GIDs per port is limited to 128, there cannot be more than the allowed IP addresses configured to Ethernet devices that are associated with the port. Allowed number is:</p> <ul style="list-style-type: none"> • "127" for a single function machine • "15" for a hypervisor in a multifunction machine • "$(127 - 15) / n$" for a guest in a multifunction machine (where n is the number of virtual functions) | - |
| 6. | A working IP connectivity between the RoCE devices is required when creating an address handle or modifying a QP with an address vector. | - |
| 7. | IPv4 multicast over RoCE requires the MGID format to be as follow <code>::ffff:<Multicast IPv4 Address></code> | - |
| 8. | IP routable RoCE does not support Multicast Listener Discovery (MLD) therefore, multicast traffic over IPv6 may not work as expected. | - |
| 9. | Using GID index 0 (the default GID) is possible only if the matching IPv6 link local address is configured on the net device of the port. This behavior is possible even though the default GID is configured regardless the presence of the IPv6 address. | - |
| 10. | Using IPv6 link local address (GID0) when VLANs are configured is not supported. | - |

4.33 Storage Known Issues

Table 42 - Storage Known Issues

| Index | Description | Workaround |
|-------|--|--|
| 1. | Older versions of <code>rescan_scsi_bus.sh</code> may not recognize some newly created LUNs. | If encountering such issues, it is recommended to use the '-c' flag. |

4.34 SRP Known Issues

Table 43 - SRP Known Issues

| Index | Description | Workaround |
|-------|--|------------|
| 1. | MLNX_OFED SRP installation breaks the <code>ibmvstgt</code> and <code>ibmvscsi</code> symbol resolution in RHEL7.0 | - |

4.35 SRP Interop Known Issues

Table 44 - SRP Interop Known Issues

| Index | Description | Workaround |
|-------|--|------------|
| 1. | The driver is tested with Storage target vendors recommendations for <code>multipath.conf</code> extensions (ZFS, DDN, TMS, Nimbus, NetApp). | - |

4.36 DDN Storage Fusion 10000 Target Known Issues

Table 45 - DDN Storage Fusion 10000 Target Known Issues

| Index | Description | Workaround |
|-------|--|------------|
| 1. | DDN does not accept non-default <code>P_Key</code> connection establishment. | - |

4.37 Oracle Sun ZFS Storage 7420 Known Issues

Table 46 - Oracle Sun ZFS Storage 7420 Known Issues

| Index | Description | Workaround |
|-------|--|------------|
| 1. | Ungraceful power cycle of an initiator connected with Targets DDN, Nimbus, NetApp may result in temporary "stale connection" messages when initiator reconnects. | - |

4.38 iSER Initiator Known Issues

Table 47 - iSER Initiator Known Issues

| Index | Description | Workaround |
|-------|---|--|
| 1. | On SLES OSs, the <code>ib_iser</code> module does not load on boot. | Add a dummy interface using <code>iscsiadm</code> : <ul style="list-style-type: none"> • <code># iscsiadm -m iface -I ib_iser -o new</code> • <code># iscsiadm -m iface -I ib_iser -o update -n iface.transport_name -v ib_iser</code> |
| 2 | Ubuntu12 requires update of user space <code>open-iscsi</code> to v2.0.873 | - |
| 3 | The initiator does not respect interface parameter while logging in. | Configure each interface on a different subnet. |
| 4 | iSCSID v2.0.873 can enter an endless loop on bind error. | - |
| 5 | iSCSID may hang if target crashes during logout sequence (reproducible with TCP). | - |

4.39 iSER Target Known Issues

Table 48 - iSER Target Known Issues

| Index | Description | Workaround |
|-------|--|--|
| 1. | Currently only the following OSs are supported: RHEL/ContOS 7.0, SLES12, Ubuntu14.04. | - |
| 2 | Stress login/logout from multiple initiators may cause iSER target to panic. | - |
| 3 | RHEL/CentOS 7.0: Discovery over RDMA is not supported. | - |
| 4 | <code>ib_isert</code> is unavailable on custom kernels after running the <code>mlnx_add_kernel_support.sh</code> script. | 1. Add " <code>isert=y</code> " to the <code>mlnx_add_kernel_support.sh</code> script after " <code>cat << EOF > ofed.conf</code> ". 2. Use the updated script to build MLNX_OFED for the custom kernel. |

4.40 ZFS Appliance Known Issues

Table 49 - ZFS Appliance Known Issues

| Index | Description | Workaround |
|-------|--|------------|
| 1. | Connection establishment occurs twice which may cause iSER to log a stack trace. | - |

5 Change Log History

Table 50 - Change Log History

| Release | Category | Description |
|--|---|--|
| 2.4-1.0.0 | mlx4_en net-device Ethtool | Added support for Ethtool speed control and advertised link mode. |
| | | Added ethtool txvlan control for setting ON/OFF hardware TX VLAN insertion: <code>ethtool -k txvlan [on/off]</code> |
| | | Ethtool report on port parameters improvements. |
| | | Ethernet TX packet rate improvements. |
| | RoCE | RoCE uses now all available EQs and not only the 3 legacy EQs. |
| | InfiniBand | IRQ affinity hints are now set when working in InfiniBand mode. |
| | Virtualization | VXLAN fixes and performance improvements. |
| | libmlx4 & libmlx5 | Improved message rate of short messages. |
| | libmlx5 | Added ConnectX®-4 device (4114) to the list of supported devices (<code>hca_table</code>), |
| | Storage | Added iSER Target driver. |
| | Ethernet net-device | New adaptive interrupt moderation scheme to improve CPU utilization. |
| RSS support of fragmented IP datagram. | | |
| Connect-IB Virtual Function | Added Connect-IB Virtual Function to the list of supported devices. | |
| 2.3-2.0.5 | mlx5_core | <p>Added the following files under <code>/sys/class/infiniband/mlx5_0/mr_cache/</code>:</p> <ul style="list-style-type: none"> <code>rel_timeout</code>: Defines the minimum allowed time between the last MR creation to the first MR released from the cache. When <code>rel_timeout = -1</code>, MRs are not released from the cache <code>rel_imm</code>: Triggers the immediate release of excess MRs from the cache when set to 1. When all excess MRs are released from the cache, <code>rel_imm</code> is reset back to 0. |
| | Bug Fixes | See “Bug Fixes History” on page 12. |
| 2.3-2.0.1 | Bug Fixes | See “Bug Fixes History” on page 12. |

Table 50 - Change Log History

| Release | Category | Description |
|-----------|---------------------------|---|
| 2.3-2.0.0 | Connect-IB | Added Suspend to RAM (S3). |
| | Reset Flow | Added Enhanced Error Handling for PCI (EEH), a recovery strategy for I/O errors that occur on the PCI bus. |
| | Register Contiguous Pages | Added the option to ask for a specific address when the register memory is using contiguous page. |
| | mlx5_core | Moved the mr_cache subtree from debugfs to mlx5_ib while preserving all its semantics. |
| | InfiniBand Utilities | Updated the ibutils package. Added to the ibdiagnet tool the "ibdiagnet2.mlx_cntrs" option to enable reading of Mellanox diagnostic counters. |
| | Bug Fixes | See "Bug Fixes History" on page 12. |
| 2.3-1.0.1 | OpenSM | Added Routing Chains support with Minhop/UPDN/FTree/DOR/Torus-2QoS |
| | | Added double failover elimination. When the Master SM is turned down for some reason, the Standby SM takes ownership over the fabric and remains the Master SM even when the old Master SM is brought up, to avoid any unnecessary re-registrations in the fabric. To enable this feature, set the "master_sm_priority" parameter to be greater than the "sm_priority" parameter in all SMs in the fabric. Once the Standby SM becomes the Master SM, its priority becomes equal to the "master_sm_priority". So that additional SM handover is avoided. Default value of the master_sm_priority is 14. To disable this feature, set the "master_sm_priority" in opensm.conf to 0. |
| | | Added credit-loop free unicast/multicast updn/ftree routing |
| | | Added multithreaded Minhop/UPDN/DOR routing |
| | RoCE | Added IP routable RoCE modes. For further information, please refer to the MLNX_OFED User Manual. |
| | Installation | Added apt-get installation support. |

Table 50 - Change Log History

| Release | Category | Description |
|-----------|----------------------------|--|
| | Ethernet | Added support for arbitrary UDP port for VXLAN. From upstream 3.15-rc1 and onward, it is possible to use arbitrary UDP port for VXLAN. This feature requires firmware version 2.32.5100 or higher. Additionally, the following kernel configuration option <code>CONFIG_MLX4_EN_VXLAN=y</code> must be enabled. |
| | | MLNX_OFED no longer changes the OS sysctl TCP parameters. |
| | | Added Explicit Congestion Notification (ECN) support |
| | | Added Flow Steering: A0 simplified steering support |
| | | Added RoCE v2 support |
| 2.3-1.0.1 | InfiniBand Network | Added Secure host to enable the device to protect itself and the subnet from malicious software. |
| | | Added User-Mode Memory Registration (UMR) to enable the usage of RDMA operations and to scatter the data at the remote side through the definition of appropriate memory keys on the remote side. |
| | | Added On-Demand-Paging (ODP), a technique to alleviate much of the shortcomings of memory registration. |
| | | Added Masked Atomics operation support |
| | | Added Checksum offload for packets without L4 header support |
| | | Added Memory re-registration to allow the user to change attributes of the memory region. |
| | Resiliency | Added Reset Flow for ConnectX®-3 (+SR-IOV) support. |
| | SR-IOV | Added Virtual Guest Tagging (VGT+), an advanced mode of Virtual Guest Tagging (VGT), in which a VF is allowed to tag its own packets as in VGT, but is still subject to an administrative VLAN trunk policy. |
| | Ethtool | Added Cable EEPROM reporting support |
| | | Disable/Enable ethernet RX VLAN tag striping offload via ethtool |
| | | 128 Byte Completion Queue Entry (CQE) |
| | Non-Linux Virtual Machines | Added Windows Virtual Machine over Linux KVM Hypervisor (SR-IOV with InfiniBand only) support |

Table 50 - Change Log History

| Release | Category | Description |
|---------------|---|---|
| Rev 2.2-1.0.1 | mlnxofedinstall | 32-bit libraries are no longer installed by default on 64-bit OS. To install 32-bit libraries use the '--with-32bit' installation parameter. |
| | openibd | Added pre/post start/stop scripts support. For further information, please refer to section “ <i>openibd Script</i> ” in the MLNX_OFED User Manual. |
| | Reset Flow | Reset Flow is not activated by default. It is controlled by the <code>mlx-4_core'internal_err_reset'</code> module parameter. |
| Rev 2.2-1.0.1 | InfiniBand Core | Asymmetric MSI-X vectors allocation for the SR-IOV hypervisor and guest instead of allocating 4 default MSI-X vectors. The maximum number of MSI-X vectors is <code>num_cpu</code> for port ConnectX®-3 has 1024 MSI-X vectors, 28 MSI-X vectors are reserved. <ul style="list-style-type: none"> Physical Function - gets the number of MSI-X vectors according to the <code>pf_msix_table_size</code> (multiple of 4 - 1) INI parameter Virtual Functions – the remaining MSI-X vectors are spread equally between all VFs, according to the <code>num_vfs mlx-4_core</code> module parameter |
| | Ethernet | Ethernet VXLAN support for kernels 3.12.10 or higher |
| | | Power Management Quality of Service: when the traffic is active, the Power Management QoS is enabled by disabling the CPU states for maximum performance. |
| | | Ethernet PTP Hardware Clock support on kernels/OSes that support it |
| | Verbs | Added additional experimental verbs interface. This interface exposes new features which are not integrated yet in to the upstream libibverbs. The Experimental API is an extended API therefore, it is backward compatible, meaning old application are not required to be recompiled to use MLNX-OFED v2.2-1.0.1. |
| Performance | Out of the box performance improvements: <ul style="list-style-type: none"> Use of affinity hints (based on NUMA node of the device) to indicate the IRQ balancer daemon on the optimal IRQ affinity Improvement in buffers allocation schema (based on the hint above) Improvement in the adaptive interrupt moderation algorithm | |

Table 50 - Change Log History

| Release | Category | Description |
|--|--|---|
| Rev 2.1-1.0.6 | IB Core | Added allocation success verification process to <code>ib_alloc_device</code> . |
| | dapl | dapl is recompiled with no FCA support. |
| | openibd | Added the ability to bring up child interfaces even if the parent's <code>ifcfg</code> file is not configured. |
| | libmlx4 | Unmapped the <code>hca_clock_page</code> parameter from <code>mlx4_uninit_context</code> . |
| | scsi_transport_srp | <code>scsi_transport_srp</code> cannot be cleared up when <code>rport</code> reconnecting fails. |
| | mlnxofedinstall | Added support for the following parameters: <ul style="list-style-type: none"> '--umad-dev-na' '--without-<package>' |
| Rev 2.1-1.0.6 | Content Packages Updates | The following packages were updated: <ul style="list-style-type: none"> bupc to v2.2-407 mstflint to v3.5.0-1.1.g76e4acf perftest to v2.0-0.76.gbf9a463 hcoll to v2.0.472-1 Openmpi to v1.6.5-440ad47 dapl to v2.0.40 |
| Rev 2.1-1.0.0 | EoIB | EoIB is supported only in SLES11SP2 and RHEL6.4. |
| | eIPoIB | eIPoIB is currently at GA level. |
| | Connect-IB® | Added the ability to resize CQs. |
| | IPoIB | Reusing DMA mapped SKB buffers: Performance improvements when IOMMU is enabled. |
| | mlnx_en | Added reporting autonegotiation support. |
| | | Added Transmit Packet Steering (XPS) support. |
| | | Added reporting 56Gbit/s link speed support. |
| | | Added Low Latency Socket (LLS) support. |
| Added check for <code>dma_mapping</code> errors. | | |
| eIPoIB | Added non-virtual environment support. | |

Table 50 - Change Log History

| Release | Category | Description |
|---|-------------------|---|
| Rev 2.0-3.0.0 | Operating Systems | Additional OS support: <ul style="list-style-type: none"> • SLES11SP3 • Fedora16, Fedora17 |
| | Drivers | Added Connect-IB™ support |
| | Installation | Added ability to install MLNX_OFED with SR-IOV support. |
| | | Added Yum installation support |
| | EoIB | EoIB (at beta level) is supported only in SLES11SP2 and RHEL6.4 |
| | mlx4_core | Modified module parameters to associate configuration values with specific PCI devices identified by their bus/device/function value format |
| | mlx4_en | Reusing DMA mapped buffers: major performance improvements when IOMMU is enabled |
| | | Added Port level QoS support |
| | IPoIB | Reduced memory consumption |
| | | Limited the number TX and RX queues to 16 |
| Default IPoIB mode is set to work in Datagram, except for Connect-IB™ adapter card which uses IPoIB with Connected mode as default. | | |
| Rev 2.0-3.0.0 | Storage | iSER (at GA level) |

Table 50 - Change Log History

| Release | Category | Description |
|----------------------------|---|---|
| Rev 2.0-2.0.5 ^a | Virtualization | SR-IOV for both Ethernet and InfiniBand (at Beta level) |
| | Ethernet Network | RoCE over SR-IOV (at Beta level) |
| | | eIPoIB to enable IPoIB in a Para-Virtualized environment (at Alpha level) |
| | | Ethernet Performance Enhancements (NUMA related and others) for 10G and 40G |
| | | Ethernet Time Stamping (at Beta level) |
| | | Flow Steering for Ethernet and InfiniBand. (at Beta level) |
| | | Raw Eth QPs: <ul style="list-style-type: none"> Checksum TX/RX Flow Steering |
| | InfiniBand Network | Contiguous pages: <ul style="list-style-type: none"> Internal memory allocation improvements Register shared memory Control objects (QPs, CQs) |
| | Installation | YUM update support |
| | VMA | OFED_VMA integration to a single branch |
| | Storage | iSER (at Beta level) and SRP |
| | Operating Systems | Errata Kernel upgrade support |
| | API | VERSION query API: library and headers |
| Counters | 64bit wide counters (port xmit/recv data/packets unicast/mcast) | |

a. SR-IOV, Ethernet Time Stamping and Flow Steering are ConnectX®-3 HCA capability.

6 API Change Log History

Table 51 - API Change Log History

| Release | Name | Description |
|---------------|------------|--|
| 2.4-1.0.0 | libibverbs | <p>Added the following verbs interfaces:</p> <ul style="list-style-type: none"> • <code>ibv_create_flow</code> • <code>ibv_destroy_flow</code> • <code>ibv_exp_use_priv_env</code> • <code>ibv_exp_setenv</code> |
| Rev 2.3-1.0.1 | libibverbs | <ul style="list-style-type: none"> • <code>ibv_exp_rereg_mr</code> - Added new API for memory region re-integration (For further information, please refer to MLNX_OFED User Manual) • Added to the experimental API <code>ibv_exp_post_send</code> the following opcodes: <ul style="list-style-type: none"> • <code>IBV_EXP_WR_EXT_MASKED_ATOMIC_CMP_AND_SWP</code> • <code>IBV_EXP_WR_EXT_MASKED_ATOMIC_FETCH_AND_ADD</code> • <code>IBV_EXP_WR_NOP</code> and these completion opcodes: <ul style="list-style-type: none"> • <code>IBV_EXP_WC_MASKED_COMP_SWAP</code> • <code>IBV_EXP_WC_MASKED_FETCH_ADD</code> |
| Rev 2.2-1.0.1 | libibverbs | <p>The following verbs changed to align with upstream libibverbs:</p> <ul style="list-style-type: none"> • <code>ibv_reg_mr</code> - <code>ibv_access_flags</code> changed. • <code>ibv_post_send</code> - opcodes and send flags changed and <code>wr</code> fields removed (<code>task</code>, <code>op</code>, <code>dc</code> and <code>bind_mw</code>) • <code>ibv_query_device</code> - capability flags changed. • <code>ibv_poll_cq</code> - opcodes and <code>wc</code> flags changed. • <code>ibv_modify_qp</code> - mask bits changed • <code>ibv_create_qp_ex</code> - <code>create_flags</code> field removed. <p>The following verbs removed to align with upstream libibverbs:</p> <ul style="list-style-type: none"> • <code>ibv_bind_mw</code> • <code>ibv_post_task</code> • <code>ibv_query_values_ex</code> • <code>ibv_query_device_ex</code> • <code>ibv_poll_cq_ex</code> • <code>ibv_reg_shared_mr_ex</code> • <code>ibv_reg_shared_mr</code> • <code>ibv_modify_cq</code> • <code>ibv_create_cq_ex</code> • <code>ibv_modify_qp_ex</code> |

Table 51 - API Change Log History

| Release | Name | Description |
|---------------|--|---|
| Rev 2.2-1.0.1 | Verbs Experimental API | <p>The following experimental verbs added (replacing the removed extended verbs):</p> <ul style="list-style-type: none"> • <code>ibv_exp_bind_mw</code> • <code>ibv_exp_post_task</code> • <code>ibv_exp_query_values</code> • <code>ibv_exp_query_device</code> • <code>ibv_exp_poll_cq</code> • <code>ibv_exp_reg_shared_mr</code> • <code>ibv_exp_modify_cq</code> • <code>ibv_exp_create_cq</code> • <code>ibv_exp_modify_qp</code> <p>New experimental verbs:</p> <ul style="list-style-type: none"> • <code>ibv_exp_arm_dct</code> • <code>ibv_exp_query_port</code> • <code>ibv_exp_create_flow</code> • <code>ibv_exp_destroy_flow</code> • <code>ibv_exp_post_send</code> • <code>ibv_exp_reg_mr</code> • <code>ibv_exp_get_provider_func</code> |
| Rev 2.1-1.0.0 | Dynamically Connected (DC) | <p>The following verbs were added:</p> <ul style="list-style-type: none"> • <code>struct ibv_dct *ibv_exp_create_dct(struct ibv_context *context, struct ibv_exp_dct_init_attr *attr)</code> • <code>int ibv_exp_destroy_dct(struct ibv_dct *dct)</code> • <code>int ibv_exp_query_dct(struct ibv_dct *dct, struct ibv_exp_dct_attr *attr)</code> |
| | Verbs Extension API: Verbs extension API defines OFA APIs extension scheme to detect ABI compatibility and enable backward and forward compatibility support. | <ul style="list-style-type: none"> • <code>ibv_post_task</code> • <code>ibv_query_values_ex</code> • <code>ibv_query_device_ex</code> • <code>ibv_create_flow</code> • <code>ibv_destroy_flow</code> • <code>ibv_poll_cq_ex</code> • <code>ibv_reg_shared_mr_ex</code> • <code>ibv_open_xrcd</code> • <code>ibv_close_xrcd</code> • <code>ibv_modify_cq</code> • <code>ibv_create_srq_ex</code> • <code>ibv_get_srq_num</code> • <code>ibv_create_qp_ex</code> • <code>ibv_create_cq_ex</code> • <code>ibv_open_qp</code> • <code>ibv_modify_qp_ex</code> |

Table 51 - API Change Log History

| Release | Name | Description |
|---------------|--|--|
| Rev 2.1-1.0.0 | Verbs Experimental API: Verbs experimental API defines MLNX-OFED APIs extension scheme which is similar to the “Verbs extension API”. This extension provides a way to introduce new features before they are integrated into the formal OFA API and to the upstream kernel and libs. | <ul style="list-style-type: none"> • <code>ibv_exp_create_qp</code> • <code>ibv_exp_query_device</code> • <code>ibv_exp_create_dct</code> • <code>ibv_exp_destroy_dct</code> • <code>ibv_exp_query_dct</code> |
| Rev 2.0-3.0.0 | XRC | <p>The following verbs have become deprecated:</p> <ul style="list-style-type: none"> • <code>struct ibv_xrc_domain *ibv_open_xrc_domain</code> • <code>struct ibv_srq *ibv_create_xrc_srq</code> • <code>int ibv_close_xrc_domain</code> • <code>int ibv_create_xrc_rcv_qp</code> • <code>int ibv_modify_xrc_rcv_qp</code> • <code>int ibv_query_xrc_rcv_qp</code> • <code>int ibv_reg_xrc_rcv_qp</code> • <code>int ibv_unreg_xrc_rcv_qp</code> |
| Rev 2.0-2.0.5 | Libibverbs - Extended speeds | <ul style="list-style-type: none"> • Missing the <code>ext_active_speed</code> attribute from the <code>struct ibv_port_attr</code> • Removed function <code>ibv_ext_rate_to_int</code> • Added functions <code>ibv_rate_to_mbps</code> and <code>mbps_to_ibv_rate</code> |
| | Libibverbs - Raw QPs | QP types <code>IBV_QPT_RAW_PACKET</code> and <code>IBV_QPT_RAW_ETH</code> are not supported |
| | Libibverbs - Contiguous pages | <ul style="list-style-type: none"> • Added Contiguous pages support • Added function <code>ibv_reg_shared_mr</code> |
| | Libmverbs | <ul style="list-style-type: none"> • The enumeration <code>IBV_M_WR_CALC</code> was renamed to <code>IBV_M_WR_CALC_SEND</code> • The enumeration <code>IBV_M_WR_WRITE_WITH_IMM</code> was added • In the structure <code>ibv_m_send_wr</code>, the union <code>wr.send</code> was renamed to <code>wr.calc_send</code> and <code>wr.rdma</code> was added • The enumerations <code>IBV_M_WQE_CAP_CALC_RDMA_WRITE_WITH_IMM</code> was added <p>The following enumerations were renamed:</p> <ul style="list-style-type: none"> • From <code>IBV_M_WQE_SQ_ENABLE_CAP</code> to <code>IBV_M_WQE_CAP_SQ_ENABLE</code> • From <code>IBV_M_WQE_RQ_ENABLE_CAP</code> to <code>IBV_M_WQE_CAP_RQ_ENABLE</code> • From <code>IBV_M_WQE_CQE_WAIT_CAP</code> to <code>IBV_M_WQE_CAP_CQE_WAIT</code> • From <code>IBV_M_WQE_CALC_CAP</code> to <code>IBV_M_WQE_CAP_CALC_SEND</code> |