



Mellanox OFED for Linux Release Notes

Rev 4.0-2.0.2.0



NOTE:

THIS HARDWARE, SOFTWARE OR TEST SUITE PRODUCT ("PRODUCT(S)") AND ITS RELATED DOCUMENTATION ARE PROVIDED BY MELLANOX TECHNOLOGIES "ASIS" 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 2017. Mellanox Technologies Ltd. All Rights Reserved.

Mellanox®, Mellanox logo, Accelio®, BridgeX®, CloudX logo, CompustorX®, Connect-IB®, ConnectX®, CoolBox®, CORE-Direct®, EZchip®, EZchip logo, EZappliance®, EZdesign®, EZdriver®, EZsystem®, GPUDirect®, InfiniHost®, InfiniBridge®, InfiniScale®, Kotura®, Kotura logo, Mellanox CloudRack®, Mellanox CloudXMellanox®, Mellanox Federal Systems®, Mellanox HostDirect®, Mellanox Multi-Host®, Mellanox Open Ethernet®, Mellanox OpenCloud®, Mellanox OpenCloud Logo®, Mellanox PeerDirect®, Mellanox ScalableHPC®, Mellanox StorageX®, Mellanox TuneX®, Mellanox Connect Accelerate Outperform logo, Mellanox Virtual Modular Switch®, MetroDX®, MetroX®, MLNX-OS®, NP-1c®, NP-2®, NP-3®, Open Ethernet logo, PhyX®, PlatformX®, PSIPHY®, SiPhy®, StoreX®, SwitchX®, Tiler®, Tiler logo, TestX®, TuneX®, The Generation of Open Ethernet logo, UFM®, Unbreakable Link®, Virtual Protocol Interconnect®, Voltaire® and Voltaire logo are registered trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

For the most updated list of Mellanox trademarks, visit <http://www.mellanox.com/page/trademarks>

Table of Contents

Table of Contents	3
Chapter 1 Overview	5
1.1 Content of Mellanox OFED for Linux.	5
1.2 Supported Platforms and Operating Systems	6
1.2.1 Supported Non-Linux Virtual Machines	7
1.3 Hardware and Software Requirements	7
1.4 Supported HCAs Firmware Versions	8
1.5 Compatibility Matrix.	8
1.6 RoCE Modes Matrix	9
Chapter 2 Changes and New Features in Rev 4.0-2.0.2.0	10
2.1 API Changes in MLNX_OFED	10
2.2 Unsupported Functionalities/Features/HCAs	10
Chapter 3 Known Issues	11
Chapter 4 Bug Fixes History	62
Chapter 5 Change Log History	83



Release Update History

Release	Date	Description
Rev 4.0-2.0.2.0	April 30, 2017	Initial release of this version.

1 Overview

These are the release notes of MLNX_OFED for Linux Driver, Rev 4.0-2.0.2.0 which operates across all Mellanox network adapter solutions supporting the following uplinks to servers:

Uplink/HCAs	Uplink Speed
ConnectX®-3/ ConnectX®-3 Pro	<ul style="list-style-type: none"> InfiniBand: SDR, QDR, FDR10, FDR Ethernet: 10GigE, 40GigE and 56GigE^a
ConnectX®-4	<ul style="list-style-type: none"> InfiniBand: SDR, QDR, FDR, FDR10, EDR Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, 56GigE^a, and 100GigE
ConnectX®-4 Lx	<ul style="list-style-type: none"> Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, and 50GigE
ConnectX®-5	<ul style="list-style-type: none"> InfiniBand: SDR, QDR, FDR, FDR10, EDR Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, and 100GigE
ConnectX®-5 Ex	<ul style="list-style-type: none"> InfiniBand: SDR, QDR, FDR, FDR10, EDR Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, and 100GigE
Connect-IB®	<ul style="list-style-type: none"> InfiniBand: SDR, QDR, FDR10, FDR
PCI Express 2.0	2.5 or 5.0 GT/s
PCI Express 3.0	8 GT/s

a. 56 GbE is a Mellanox proprietary link speed and can be achieved while connecting a Mellanox adapter cards to Mellanox SX10XX switch series or connecting a Mellanox adapter card to another Mellanox adapter card.

1.1 Content of Mellanox OFED for Linux

Mellanox OFED for Linux software contains the following 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 initiator, iSER initiator and target, NVMeoF host and target
OpenFabrics utilities	<ul style="list-style-type: none"> OpenSM: IB Subnet Manager with Mellanox proprietary Adaptive Routing Diagnostic tools Performance tests SSA (SLES12): libopensmssa plugin for OpenSM, ibssa, ibacm
MPI	<ul style="list-style-type: none"> Open MPI stack 1.6.5 and later supporting the InfiniBand interface MPI benchmark tests (OSU benchmarks, Intel MPI benchmarks, Presta)

Components	Description
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 v3.x (MPI/PGAS collective operations acceleration library over InfiniBand) • KNEM, Linux kernel module enabling high-performance intra-node MPI/PGAS communication for large messages
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	
HCAs	<ul style="list-style-type: none"> • ConnectX-4 EN driver Rev 4.0-2.0.0.1 • ConnectX-3 EN driver Rev 4.0-2.0.0.1
Documentation	

1.2 Supported Platforms and Operating Systems

The following are the supported OSs in MLNX_OFED Rev 4.0-2.0.2.0:

Table 1 - Supported Platforms and Operating Systems

Operating System	Platform
Ubuntu 17.04	x86_64



32 bit platforms are no longer supported in MLNX_OFED.



For RPM based distributions, 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 Manual for instructions.



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

1.2.1 Supported Non-Linux Virtual Machines

The following are the supported Non-Linux (InfiniBand only) Virtual Machines in MLNX_OFED Rev 4.0-2.0.2.0:

- Windows Server 2012 R2

1.3 Hardware and Software Requirements

The following are the hardware and software requirements of MLNX_OFED Rev 4.0-2.0.2.0.

- 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:

Operating System	Required Packages Installation Command
Ubuntu/Debian	<code>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 pkg-config libnuma logrotate ethtool</code>

1.4 Supported HCAs Firmware Versions

MLNX_OFED Rev 4.0-2.0.2.0 supports the following Mellanox network adapter cards firmware versions:

Table 2 - Supported HCAs Firmware Versions

HCA	Recommended Firmware Rev.	Additional Firmware Rev. Supported
ConnectX®-3	2.40.7000	2.36.5150
ConnectX®-3 Pro	2.40.7000	2.36.5150
ConnectX®-4	12.18.2000	12.18.1000
ConnectX®-4 Lx	14.18.2000	14.18.1000
ConnectX®-5	16.19.1200	16.18.1000
ConnectX®-5 Ex	16.19.1200	16.18.1000
Connect-IB®	10.16.1020	N/A

For the official firmware versions, please see:

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

1.5 Compatibility Matrix

MLNX_OFED Rev 4.0-2.0.2.0 is compatible with the following:

Table 3 - Compatibility Matrix

Mellanox Product	Description/Version
MLNX-OS®	MSX6036 w/w MLNX-OS® version 3.4.3202 ^a
Grid Director™	4036 w/w Grid Director™ version 3.9.1-985
Unified Fabric Manager (UFM®)	v5.8
MXM	v3.5
HPC-X UPC	v2.22
FCA ^b	v3.x
OpenMPI	v1.10

- MLNX_OFED Rev 4.0-2.0.2.0 was tested with this switch. However, additional switches may be supported as well.
- FCA v3.x (HCOLL) is now the default FCA version used in HPC-X, starting from HPC-X v1.8. This FCA version replaces v2.x.

1.6 RoCE Modes Matrix

The following is RoCE modes matrix:

Table 4 - RoCE Modes Matrix

Software Stack / Inbox Distribution	RoCEv1 (IP Based GIDs) Supported as of Version		RoCEv2 Supported as of Version		RoCEv1 & RoCEv2 (RoCE per GID) Supported as of Version
	ConnectX-3/ ConnectX-3 Pro	ConnectX-4/ ConnectX-4 Lx/ ConnectX-5/ ConnectX-5 Ex ^a	ConnectX-3 Pro	ConnectX-4/ ConnectX-4 Lx/ ConnectX-5/ ConnectX-5 Ex ^a	ConnectX-3 Pro/ConnectX-4/ ConnectX-4 Lx/ConnectX-5/ ConnectX-5 Ex ^a
MLNX_OFED	2.1-x.x.x	3.0-x.x.x	2.3-x.x.x	3.0-x.x.x	3.0-x.x.x
Kernel.org	3.14	4.4	4.4	4.4	4.4
RHEL	6.6, 7.0	-	-	-	-
SLES	12	-	-	-	-
Ubuntu	14.04.4, 16.04, 15.10	-	-	-	-

a. Note that support for ConnectX-5 and ConnectX-5 Ex adapter cards in MLNX_OFED starts from v4.0.

2 Changes and New Features in Rev 4.0-2.0.2.0

The following are the changes and/or new features that have been added to this version of MLNX_OFED.

Table 5 - Changes and New Features in Rev 4.0-2.0.2.0

HCA's	Feature/Change	Description
All	Operating System	Added support for Ubuntu v17.04.

For additional information on the new features, please refer to the MLNX_OFED User Manual.

2.1 API Changes in MLNX_OFED



Note that the following APIs will be deprecated and replaced with the new APIs as of MLNX_OFED version 4.0, as listed in the table below.

Table 6 - API Changes

Feature	Type	Current API	New API
Rereg MR	Verb	ibv_exp_rereg_mr	ibv_rereg_mr
Memory Window	Verb	ibv_exp_bind_mw	ibv_bind_mw
	Structure	ibv_exp_send_wr -> bind_mw	ibv_send_wr -> bind_mw
		IBV_EXP_WR_SEND_WITH_INV	IBV_WR_SEND_WITH_INV
		IBV_EXP_WR_LOCAL_INV	IBV_WR_LOCAL_INV
	Capability	IBV_EXP_WR_BIND_MW	IBV_WR_BIND_MW
		IBV_EXP_DEVICE_MEM_WINDOW	IBV_DEVICE_MEM_WINDOW
Completion	IBV_EXP_WC_WITH_INV	IBV_WC_WITH_INV	

2.2 Unsupported Functionalities/Features/HCA's

The following are the unsupported functionalities/features/HCA's in MLNX_OFED:

- ConnectX®-2 Adapter Card
- Ethernet IP over InfiniBand (eIPoIB)
- Relational Database Service (RDS)
- Ethernet over InfiniBand (EoIB) - mlx4_vnic
- Ethernet IP over InfiniBand (EIPoIB)
- mthca InfiniBand driver

3 Known Issues

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

Table 7 - Known Issues

Internal Ref	Issue
1006032	<p>Description: Currently, iSER initiator supports T10-PI offload mechanism only in the following OSs:</p> <ul style="list-style-type: none"> • Ubuntu 16.10, 16.04 • RedHat 7.3, 7.2 • SLES 12.02
	<p>Workaround: N/A</p>
	<p>Keywords: iSER initiator</p>
1007830	<p>Description: When working on Xenserver hypervisor with SR-IOV enabled on it, make sure the following instructions are applied:</p> <ol style="list-style-type: none"> 1. Right after enabling SR-IOV, unbind all driver instances of the virtual functions from their PCI slots. 2. It is not allowed to unbind PF driver instance while having active VFs.
	<p>Workaround: N/A</p>
	<p>Keywords: SR-IOV</p>
1006768	<p>Description: An rdma_cm connection between a client and a server that are on the same host is not possible when working over VLAN interfaces.</p>
	<p>Workaround: N/A</p>
	<p>Keywords: RDMA-CM</p>
1008583	<p>Description: A soft lockup in the CQ polling flow might occur when running very high stress on the GSI QP (RDMA-CM applications). This is a transient situation and the driver recovers from it after a while.</p>
	<p>Workaround: N/A</p>
	<p>Keywords: RDMA-CM</p>
1007356	<p>Description: Creating a PKEY interface using “ip link” is not supported.</p>
	<p>Workaround: Use sysfs to create a PKEY interface.</p>
	<p>Keywords: IPoIB, PKEY</p>

Table 7 - Known Issues

Internal Ref	Issue
1009930	<p>Description: Reassembly of packets larger than 64k may fail when ipfrag threshold is low. This issue is present only on RHEL 6.3, 6.4, 6.5, and Ubuntu 12.04. This packet drop can be seen from netstat tool, indicated by “packet reassembles failed” counter.</p>
	<p>Workaround: Change sysctl ipfrag threshold. For instructions on how to perform this change, refer to the following link: https://access.redhat.com/solutions/891273</p>
	<p>Keywords: IPoIB, packet fragmentation</p>
1000197	<p>Description: Displaying multicast groups using sysfs may not show all the entries on Fedora 23 OS.</p>
	<p>Workaround: N/A</p>
	<p>Keywords: IPoIB</p>
1010148	<p>Description: Upgrading from MLNX_OFED v3.x to v4.x using yum and apt-get repositories fails.</p>
	<p>Workaround: Remove MLNX_OFED v3.x using the <code>ofed_uninstall.sh</code> script, and only then install MLNX_OFED v4.x as usual.</p>
	<p>Keywords: Installation</p>
1005786	<p>Description: When using ConnectX-5 adapter cards, the following error might be printed to dmesg, indicating temporary lack of DMA pages: <pre> mlx5_core ... give_pages:289:(pid x): Y pages alloc time exceeded the max permitted duration mlx5_core ... page_notify_fail:263:(pid x): Page allocation failure notification on func_id(z) sent to fw mlx5_core ... pages_work_handler:471:(pid x): give fail -12" </pre> </p>
	<p>Example: This might happen when trying to open more than 64 VFs per port.</p>
	<p>Workaround: N/A</p>
1008066	<p>Description: Performing some operations on the user end during reboot might cause call trace/panic, due to bugs found in the Linux kernel. For example: Running <code>get_vf_stats</code> (via <code>iptool</code>) during reboot.</p>
	<p>Workaround: N/A</p>
	<p>Keywords: <code>mlx5_core</code>, reboot</p>

Table 7 - Known Issues

Internal Ref	Issue
1009488	<p>Description: Mounting MLNX_OFED to a path that contains special characters, such as parenthesis or spaces is not supported. For example, when mounting MLNX_OFED to “/media/CDROM(vcd)”, installation will fail and the following error message will be displayed:</p> <pre># cd /media/CDROM(vcd\)/ # ./mlnxofedinstall sh: 1: Syntax error: "(" unexpected</pre> <p>Workaround: N/A</p> <p>Keywords: Installation</p>
982144	<p>Description: When offload traffic sniffer is on, the bandwidth could decrease up to 50%.</p> <p>Workaround: N/A</p> <p>Keywords: Offload Traffic Sniffer</p>
981045	<p>Description: On kernels below v4.2, when removing a bonding module with devices different from ARPHRD_ETHER, a call trace may be received.</p> <p>Workaround: Remove the bond in the following order: Remove the slaves, delete the bond, and only then remove the bonding module.</p> <p>Keywords: Bonding</p>
980066	<p>Description: Soft RoCE does not support Extended Reliable Connection (XRC).</p> <p>Workaround: N/A</p> <p>Keywords: Soft RoCE, XRC</p>
982534	<p>Description: In ConnectX-3, when using a server with page size of 64K, the UAR BAR will become too small. This may cause one of the following issues:</p> <ol style="list-style-type: none"> 1. mlx4_core driver does not load. 2. The mlx4_core driver does load, but calls to <code>ibv_open_device</code> may return ENOMEM errors. <p>Workaround:</p> <ol style="list-style-type: none"> 1. Add the following parameter in the firmware's ini file under [HCA] section: <code>log2_uar_bar_megabytes = 7</code> 2. Re-burn the firmware with the new ini file. <p>Keywords: PPC</p>

Table 7 - Known Issues

Internal Ref	Issue
981362	Description: On several OSs, setting a number of TC is not supported via the tc tool.
	Workaround: Set the number of TC via the <code>/sys/class/net/<interface>/qos/tc_num</code> sysfs file.
	Keywords: Ethernet, TC
980257	Description: An issue in InfiniBand bond interfaces may cause memory corruption in Ubuntu v14.04 and v14.10 OSs. The memory corruption happens when attempting to reload the driver while the bond is up with InfiniBand slaves.
	Workaround: Delete the bond before restarting the driver.
	Keywords: Bonding, IPoIB
980034	Description: Soft RoCE counters located under <code>/sys/class/infiniband/<rxn-inf>/ports/1/counters/</code> directory are not supported.
	Workaround: N/A
	Keywords: Soft RoCE
979907	Description: Only the following two experimental verbs are supported for Soft RoCE: <ul style="list-style-type: none"> • <code>ibv_exp_query_device</code> • <code>ibv_exp_poll_cq</code>.
	Workaround: N/A
	Keywords: Soft RoCE
979457	Description: When setting <code>IOMMU=ON</code> , a severe performance degradation may occur due to a bug in IOMMU.
	Workaround: Make sure the following patches are found in your kernel: <ul style="list-style-type: none"> • <code>iommu/vt-d: Fix PASID table allocation</code> • <code>iommu/vt-d: Fix IOMMU lookup for SR-IOV Virtual Functions</code> Note: These patches are already available in Ubuntu 16.04.02 and 17.04 OSs.
	Keywords: Performance, IOMMU
977852	Description: <code>rdma_cm</code> running over IB ports does not support UD QPs on ConnectX-3 adapter cards.
	Workaround: N/A
	Keywords: SR-IOV, RDMA CM

Table 7 - Known Issues

Internal Ref	Issue
955113	Description: In RoCE LAG over ConnectX-4 adapter cards, the script <code>ibdev2netdev</code> may show a wrong port state for the bonded device. This means that although the IB device/port <code>mlx5_bond_0/1</code> is up (as seen in <code>ibstat</code>), <code>ibdev2netdev</code> may report that it is down.
	Workaround: N/A
	Keywords: RoCE, LAG, bonding
942161	Description: On some kernels, there might be an issue in csum calculations of tunneled packets when the driver sets <code>CHECKSUM_COMPLETE</code> for the packet. This might print csum error messages to the <code>dmesg</code> log file.
	Workaround: Make sure your kernel version includes this fix .
	Keywords: Ethernet, checksum, tunneling
931574	Description: When using a kernel with Generic Receive Offload (GRO) support, UDP performance results will reveal degradation in comparison to the UDP performance results in <code>MLNX_OFED v1.5.x</code> .
	Workaround: Turn off the GRO feature to get better UDP performance. Run: <code>#ethtool -K <interface> gro off</code>
	Keywords: GRO, UDP, performance
920707	Description: In SLES12 SP2, you may get a memory low warning at the netlink layer when configuring a large number of VFs.
	Workaround: N/A
	Keywords: SR-IOV, SLES12
969467	Description: On SLES PPC64, the removal of packages with names starting with <code>kernel-mft-mlnx</code> might fail with such an error: "Error: package <code>kernel-mft-mlnx-kmp-default</code> seems to contain modules for multiple kernel versions"
	Workaround: Use the following command to remove the <code>kernel-mft</code> packages: <code>rpm -e --noscripts \$(rpm -qa grep kernel-mft-mlnx)</code>
	Keywords: Installation
918880	Description: The driver version shown in <code>modinfo</code> and <code>ethtool</code> outputs is 3.4-1.0.6 instead of 3.4-2.0.0.
	Workaround: N/A
	Keywords: Installation

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: When upgrading from an earlier Mellanox OFED version, the installation script does not stop the earlier version prior to uninstalling it.</p> <p>Workaround: Stop the old OFED stack (<code>/etc/init.d/openibd stop</code>) before upgrading to this new version.</p> <p>Keywords: Installation</p>
-	<p>Description: When using bonding on Ubuntu OS, the "ifenslave" package must be installed.</p> <p>Workaround: N/A</p> <p>Keywords: Installation</p>
-	<p>Description: On PPC systems, the <code>ib_srp</code> module is not installed by default since it breaks the <code>ibmvscsi</code> module.</p> <p>Workaround: If your system does not require the <code>ibmvscsi</code> module, run the <code>mlnxofedinstall</code> script with the "--with-srp" flag.</p> <p>Keywords: Installation</p>
679801	<p>Description: Updating MLNX_OFED via Yum (e.g. running "<code>yum update mlnx-ofed-all</code>") can fail with the following error: --> Finished Dependency Resolution Error: Package: <code>mpitests_openmpi__1_8_8-3.2.16-fe5387c.x86_64</code> (installed) Requires: <code>liboshmem.so.3()(64bit)</code> Removing: <code>openmpi-1.8.8-1.x86_64</code> (installed) <code>liboshmem.so.3()(64bit)</code> Updated By: <code>openmpi-1.10.2rc4-1.32008.x86_64</code> (mlnx_ofed) <code>~liboshmem.so.9()(64bit)</code></p> <p>Workaround: Remove the <code>mpitests</code> packages manually: <pre># rpm -e --allmatches \$(rpm -qa grep mpitests_)</pre></p> <p>Keywords: Installation</p>
690799	<p>Description: OpenSM package removal fails with the following error on Ubuntu12.04: Removing opensm ... /sbin/insserv: No such file or directory</p> <p>Workaround:</p> <ol style="list-style-type: none"> 1. Create the missing link by running this command: <pre># ln -s /usr/lib/insserv/insserv /sbin/insserv</pre> 2. Remove the package. <p>Keywords: Installation</p>

Table 7 - Known Issues

Internal Ref	Issue
764204	<p>Description: Weak Updates (KMP) support is broken on RHEL PPC64LE with errata kernels. MLNX_OFED installation will pass, but no links will be created under the weak-updates directory for the new kernel. Therefore, the driver load will fail.</p>
	<p>Workaround:</p> <ul style="list-style-type: none"> As of MLNX_OFED v3.3, use the <code>mlnx_add_kernel_support.sh</code> script, or simply provide the <code>--add-kernel-support</code> flag to <code>mlnxofedinstall</code> script. Update the <code>kmod</code> package using the following link: https://rhn.redhat.com/errata/RHBA-2016-1832.html
	<p>Keywords: Installation</p>
785119	<p>Description: When upgrading ConnectX-4/ConnectX-4 Lx firmware version from v12/14.14.2036 to a newer one (for example:12/14.16.1xxx), power cycle is necessary to enable working in Pass-Through mode. Using <code>mlxfwreset</code> instead of power cycle will print messages similar to the following when Passing-Through the device to Virtual Machine: <pre>"-device vfio-pci,host=04:00.0,id=hostdev0,bus=pci.0,addr=0x7: vfio: Error: Failed to setup INTx fd: No such device 2016-05-22T06:46:39.164786Z qemu-kvm: -device vfio-pci,host=04:00.0,id=hostdev0,bus=pci.0,addr=0x7: Device initialization failed."</pre></p>
	<p>Workaround: N/A</p>
	<p>Keywords: Installation</p>
-	<p>Description: "openibd stop" can sometime fail with the error: <pre>Unloading ib_cm [FAILED] ERROR: Module ib_cm is in use by ib_ipoib</pre></p>
	<p>Workaround: Re-run "openibd stop"</p>
	<p>Keywords: Driver Unload</p>
-	<p>Description: 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>).</p>
	<p>Workaround: N/A</p>
	<p>Keywords: Driver Start</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: 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 mlx4_core does not belong to MLNX_OFED" or "Module mlx4_core belong to <rpm name> which is not a part of MLNX_OFED"</p> <p>Workaround: Remove the third party RPM/non MLNX_OFED drivers directory, run: "depmod" and then rerun <code>/etc/init.d/openibd restart</code>"</p> <p>Keywords: Driver Start</p>
-	<p>Description: Occasionally, when trying to repetitively reload the NES hardware driver on SLES11 SP2, a soft lockups occurs that required reboot.</p> <p>Workaround: N/A</p> <p>Keywords: Driver Start</p>
-	<p>Description: When downgrading from MLNX_OFED 3.0-x.x.x, driver reload might fail with the following error in dmeg: [166271.886407] compat: exports duplicate symbol __eth-tool_get_settings (owned by mlx_compat)</p> <p>Workaround: The issues will be resolved automatically after system reboot or by invoking the following commands: <code>rmmod mlx_compat</code> <code>depmod -a</code> <code>/etc/init.d/openibd restart</code></p> <p>Keywords: Driver Start</p>
773774	<p>Description: When downgrading from MLNX_OFED 3.3-x.x.x, driver reload might fail with the following error in dmeg: <code>rmmod: ERROR: Module mlx_compat is in use by: ib_netlink</code></p> <p>Workaround: The issues will be resolved automatically after system reboot or by invoking the following commands: <code>rmmod ib_netlink</code> <code>depmod -a</code> <code>/etc/init.d/openibd restart</code></p> <p>Keywords: Driver Start</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: 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.</p> <p>Workaround: N/A</p> <p>Keywords: Driver Start</p>
-	<p>Description: 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.</p> <p>Workaround:</p> <ol style="list-style-type: none"> 1. Unmount any mounted Lustre storages: # umount<lustre_mount_point> 2. Unload all Lustre modules: # lustre_rmmod <p>Keywords: Driver Start</p>
-	<p>Description: Driver unload fails with the following error message: Unloading rdma_cm [FAILED] rmmod: ERROR: Module rdma_cm is in use by: xprtrdma</p> <p>Workaround: Make sure that there are no mount points over NFS/RDMA prior to unloading the driver and run: # modprobe -r xprtrdma In case that the xprtrdma module keeps getting loaded automatically even though it is not used, add a pre-stop hook for the openibd service script to always unload it. Create an executable file "/etc/infiniband/pre-stop-hook.sh" with the following content: #!/bin/bash modprobe -r xprtrdma</p> <p>Keywords: Driver Start</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: When loading or unloading the driver on HP ProLiant systems, you may see log messages like: dmar: DMAR:[DMA Write] Request device [07:00.0] fault addr 3df7f000 DMAR:[fault reason 05] PTE Write access is not set This is a known issue with ProLiant systems (see their support notice for Emulex adapters: http://h20564.www2.hp.com/hpsc/doc/public/display?docId=emr_na-c04446026&lang=en-us&cc=us) The messages are harmless, and may be ignored.</p> <p>Workaround: If you are <i>*not*</i> running SR-IOV on your system, you may eliminate these messages by removing the term "intel_iommu=on" from the boot line in file <code>/boot/grub/menu.lst</code>. For SR-IOV systems, this term must remain, you can ignore the log messages.</p> <p>Keywords: Driver Start</p>
677998	<p>Description: False alarm errors may be printed to dmesg.</p> <p>Workaround: N/A</p> <p>Keywords: Driver Start</p>
610395	<p>Description: On RHEL 7.1, after updating to kernel version 3.10.0-229.14.1.el7 or later, driver load fails with unknown symbols errors in dmesg.</p> <p>Workaround: Use the <code>mlnx_add_kernel_support.sh</code> script to compile MLNX_OFED drivers against the new kernel.</p> <p>Keywords: Driver Start</p>
-	<p>Description: Loading the driver using the <code>openibd</code> 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 OL6.x this may change the local system time.</p> <p>Workaround: N/A</p> <p>Keywords: System Time</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: 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:</p> <pre>[4671958.383506] Request for unknown module key 'Mellanox Technologies signing key: 61feb074fc7292f958419386ffdd9d5-ca999e403' err -11</pre> <p>This error can be safely ignored as long as Secure Boot is disabled on the system.</p> <p>Workaround: For further information, please refer to the User Manual section “Enrolling Mellanox's x.509 Public Key On your Systems”.</p> <p>Keywords: UEFI Secure Boot</p>
-	<p>Description: Ubuntu12 requires update of user space open-iscsi to v2.0.873</p> <p>Workaround: N/A</p> <p>Keywords: UEFI Secure Boot</p>
-	<p>Description: The initiator does not respect interface parameter while logging in.</p> <p>Workaround: Configure each interface on a different subnet.</p> <p>Keywords: UEFI Secure Boot</p>
967356	<p>Description: [Ethernet]</p> <ul style="list-style-type: none"> Bare-metal ConnectX-4/ConnectX-4 Lx might suffer up to 15, degradation in some scenarios due to higher CPU utilization. PPC: ConnectX-4 might suffer up to 20, degradation in some scenarios. <p>Workaround: N/A</p> <p>Keywords: Performance</p>
956071	<p>Description: [mlx5] OOB TCP performance for small message sizes may suffer from lower BW than expected.</p> <p>Workaround: Disable adaptive-rx and set higher static moderation: <code>ethtool -C <interface> adaptive-rx off rx-frames 128 rx-usecs 128</code></p> <p>Keywords: Performance</p>
765777	<p>Description: Low VxLAN throughput due to broken GRO offload in most kernels older than kernel v4.6.</p> <p>Workaround: Use kernel version 4.6 or above.</p> <p>Keywords: Performance</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: 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.</p> <p>Workaround: Execute the following script as root: <code>set_irq_affinity.sh <interface or IB device> [2nd interface or IB device]</code></p> <p>Keywords: Performance</p>
414827	<p>Description: 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.</p> <p>Workaround: For additional performance tuning, please refer to Performance Tuning Guide.</p> <p>Keywords: Performance</p>
-	<p>Description: 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.</p> <p>Workaround: Disable adaptive interrupt moderation and set lower values for the interrupt coalescing manually. <code>ethtool -C <eth>X adaptive-rx off rx-usecs 64 rx-frames 24</code></p> <p>Values above may need tuning, depending the system, configuration and link speed.</p> <p>Keywords: Performance</p>
417751	<p>Description: Performance degradation might occur when bonding Ethernet interfaces.</p> <p>Workaround: N/A</p> <p>Keywords: Performance</p>
656415	<p>Description: In RHEL7.0, when the irqbalance service is started or restarted, it incorrectly re-balances the IRQs, including the banned ones.</p> <p>Workaround: N/A</p> <p>Keywords: Performance</p>

Table 7 - Known Issues

Internal Ref	Issue
651322	<p>Description: In RH7.0/RH7.1, performance issue with ConnectX-4 cards over 100GbE link might occur when the process of forwarding the packets between the ports, which is done by the kernel, fib_table_lookup() function is called.</p> <p>For further information, please refer to: http://comments.gmane.org/gmane.linux.network/344243</p>
	<p>Workaround: Use RH7.2 to avoid such performance issues.</p>
	<p>Keywords: Performance</p>
754646	<p>Description: The default RX coalescing values yield to high CPU utilization when using VXLAN on VMs over PV.</p>
	<p>Workaround: Increase the RX microseconds and frames coalescing parameters for a better utilization using the ethtool -C command.</p>
	<p>Keywords: Performance</p>
783496	<p>Description: When using a VF over RH7.X KVM, low throughput is expected.</p>
	<p>Workaround: Install the following packages using the link below:</p> <ul style="list-style-type: none"> • qemu-img-1.5.3-105.el7_2.1.bz1299846.0.x86_64.rpm • qemu-kvm-1.5.3-105.el7_2.1.bz1299846.0.x86_64.rpm • qemu-kvm-common-1.5.3-105.el7_2.1.bz1299846.0.x86_64.rpm <p>http://people.redhat.com/~alwillia/bz1299846/</p>
	<p>Keywords: Performance</p>
780782	<p>Description: CALC operation on PowerPC may report completion with error.</p>
	<p>Workaround: N/A</p>
	<p>Keywords: mlx5 Driver</p>
921252	<p>Description: Support for Memory Window impacts the maximum number of Work Requests (WRs) for connected QPs (for example: RC QP). Applications that create a QP with the maximal size might experience out of memory errors.</p>
	<p>Workaround: Lower the number of Work Requests during the QP creation.</p>
	<p>Keywords: mlx5 Driver</p>

Table 7 - Known Issues

Internal Ref	Issue
860311	Description: An allocation of high-order page in <code>mlx5e_alloc_striding_rx_wqe</code> fails with a call-trace.
	Workaround: No action is required on users end. A fragmented fallback flow will handle this failure.
	Keywords: mlx5 Driver
-	Description: 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)
	Workaround: N/A
	Keywords: mlx5 Driver
435583	Description: EEH events that arrive while the mlx5 driver is loading may cause the driver to hang.
	Workaround: N/A
	Keywords: mlx5 Driver
434570	Description: The mlx5 driver can handle up to 5 EEH events per hour.
	Workaround: If more events are received, cold reboot the machine.
	Keywords: mlx5 Driver
554120	Description: When working with Connect-IB firmware v10.10.5054, the following message would appear in driver start. <code>command failed, status bad system state(0x4), syndrome 0x408b33</code> The message can be safely ignored.
	Workaround: Upgrade Connect-IB firmware to the latest available version.
	Keywords: mlx5 Driver
-	Description: Changing the link speed is not supported in Ethernet driver when connected to a ConnectX-4 card.
	Workaround: N/A
	Keywords: mlx5 Driver
538843	Description: Bonding active-backup mode does not function properly.
	Workaround: N/A
	Keywords: mlx5 Driver

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: Rate, speed and width using IB sysfs/tools are available in RoCE mode in ConnectX-4 only after port physical speed configuration is done.</p> <p>Workaround: N/A</p> <p>Keywords: mlx5 Driver</p>
598092	<p>Description: Since MLNX_OFED's openibd does not unload modules while OpenSM is running, removing mlx5_core manually while OpenSM is running, may cause it to be out of sync when probed again.</p> <p>Workaround: Restart OpenSM</p> <p>Keywords: mlx5 Driver</p>
563022	<p>Description: ConnectX-4 port GIDs table shows a duplicated RoCE v2 default GID.</p> <p>Workaround: N/A</p> <p>Keywords: mlx5 Driver</p>
947542	<p>Description: mlx5 hardware offload is supported when setting up to 4 VxLAN ports (one of these ports must be 4789). When attempting to set more VxLAN ports, these ports will still be supported, but a failure message will appear in the dmesg.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
964991	<p>Description: TX queue rate limit may sometimes exceed the rate that was set by the user by up to 10,.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
911693	<p>Description: In ConnectX-4 Lx and above, the minimal RX ring size is changed to 512, as a result of fundamental changes in receive flow structures.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
948312	<p>Description: [ConnectX-3 Pro] To enable/disable rx-vlan-stag-hw-parse by ethtool, rxvlan should be enabled/disabled accordingly (ethtool -K rxvlan on/off).</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>

Table 7 - Known Issues

Internal Ref	Issue
894547	Description: On SLES12 SP1 and SLES12 SP2, invalid udev rules might cause Ethernet interfaces renaming to fail, leaving some interfaces with names such as renameXY.
	Workaround: Modify the udev rules inside the /etc/udev/rules.d/70-persistent-net.rules file, such that every rule is unique to the target interface. For further details, refer to the Ethernet Related Issues table under the Troubleshooting section in MLNX_OFED User Manual.
	Keywords: Ethernet
754709	Description: mlx5 Ethernet auto-negotiation related issues: <ol style="list-style-type: none"> 1. The command <code>ethtool -s eth4 speed 25000 autoneg on</code> is not a valid <code>ethtool</code> command. <code>speed 25000</code> should not be passed in when <code>autoneg</code> is on. Instead, use <code>advertise 0x100000</code>. 2. <code>ethtool</code> version older than v4.6 does not report neither support nor advertise for new speeds, such as 25G, 100G. 3. When setting auto negotiation with <code>ethtool</code> version older than v4.6, advertised speed will be ignored, and the device will try to reach the highest supported speed available end-to-end.
	Workaround: N/A
	Keywords: Ethernet
843306	Description: [ConnectX-4/ConnectX-4 Lx] When configuring ETS, bandwidth values are limited between 1-100, and 0 is an invalid value.
	Workaround: N/A
	Keywords: Ethernet
704750	Description: [ConnectX-4/ConnectX-4 Lx] First ICMP6 packet may be lost as a result of first IP fragment loss when packets size is significantly bigger than MTU.
	Workaround: N/A
	Keywords: Ethernet
-	Description: 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. kernel warning: <code>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</code>
	Workaround: N/A
	Keywords: Ethernet

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: Kernel panic might occur during FIO splice in kernels before 2.6.34-rc4.</p> <p>Workaround: Use kernel v2.6.34-rc4 which provides the following solution: <code>baff42a net: Fix oops from tcp_collapse() when using splice()</code></p> <p>Keywords: Ethernet</p>
-	<p>Description: In PPC systems when QoS is enabled a harmless Kernel DMA mapping error messages might appear in kernel log (IOMMU related issue).</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
-	<p>Description: 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: <code>do_IRQ: 0.203 No irq handler for vector (irq -1)</code></p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
-	<p>Description: Mixing ETS and strict QoS policies for TCs in 40GbE ports may cause inaccurate results in bandwidth division among TCs.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
-	<p>Description: Affinity hints are not supported in Xen Hypervisor (an irqblancer issue). This causes a non-optimal IRQ affinity.</p> <p>Workaround: To overcome this issues, run: <code>set_irq_affinity.sh eth<x></code></p> <p>Keywords: Ethernet</p>
433366	<p>Description: Reboot might hang in SR-IOV when using the <code>probe_vf</code> parameter with many Virtual Functions. The following message is logged in the kernel log: <code>"waiting for eth to become free. Usage count =1"</code></p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: VXLAN may not be functional when configured over Linux bridge in RH7.0 faceor Ubuntu14.04. The issue is within the bridge modules in those kernels. In Vanilla kernels above 3.16 issues is fixed.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
-	<p>Description: 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.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
-	<p>Description: 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.</p> <p>Workaround: Run: <code>ethtool -K ethX lro off</code></p> <p>Keywords: Ethernet</p>
539117	<p>Description: On SLES12, the bonding interface over Mellanox Ethernet slave interfaces does not get IP address after reboot.</p> <p>Workaround:</p> <ol style="list-style-type: none"> 1. Set "STARTMODE=hotplug" in the bonding slave's ifcfg files. More details can be found in the SUSE documentations page: https://www.suse.com/documentation/sles-12/book_sle_admin/?page=/documentation/sles-12/book_sle_admin/data/sec_bond.html 2. Enable the nanny service to support hot-plugging: Open the "/etc/wicked/common.xml" file. Change: "<code><use-nanny>>false</use-nanny></code>" to "<code><use-nanny>>true</use-nanny></code>" 3. Run: <code># systemctl restart wickedd.service wicked</code> <p>Keywords: Ethernet</p>
989042	<p>Description: <code>ethtool -x</code> command will not function on relatively old kernels that do not support <code>get/set_rxfh*</code> callbacks.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
516136	<p>Description: Ethertype proto 0x806 not supported by ethtool</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: ETS configuration is not supported in the following kernels:</p> <ul style="list-style-type: none"> • 3.7 • 3.8 • 3.9 • 3.10 • 3.2.37-94_fbk17_01925_g8e3b329 • 3.14 • 3.2.55-106_fbk22_00877_g6902630 • 3.2.28-76_fbk14_00230_g3c40d9e <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
-	<p>Description: ETS is not supported in kernels that do not have MQPRIO as QDISC_KIND option in the tc tool.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
592229	<p>Description: When NC-SI is ON, the ports MTU cannot be set to lower than 1500.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
600242	<p>Description: GRO is not functional when using VXLAN in ConnectX-3 adapter cards.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
596075	<p>Description: ethtool -X: The driver supports only the 'equal' mode and cannot be set by using weight flags.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
600752	<p>Description: Q-in-Q infrastructure in the kernel is supported only in kernel version 3.10 and up.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>
596537	<p>Description: When SLES11 SP4 is used as a DHCP client over ConnectX-3 or ConnectX-3 adapters, it might fail to get an IP from the DHCP server.</p> <p>Workaround: N/A</p> <p>Keywords: Ethernet</p>

Table 7 - Known Issues

Internal Ref	Issue
560575	Description: When using a hardware that has Time Stamping enabled, the system time might be higher than the expected variance.
	Workaround: N/A
	Keywords: Ethernet
597758	Description: In Q-in-Q, ping failed when sending traffic with package size > 1468
	Workaround: N/A
	Keywords: Ethernet
665131	Description: Call trace may occur when configuring VXLAN or under high traffic stress.
	Workaround: N/A
	Keywords: Ethernet
-	Description: HW LRO does not function in ConnectX-4 adapter cards.
	Workaround: N/A
	Keywords: Ethernet
685069/689607	Description: ethtool header does not currently support the link speeds of 25/50/100. Therefore, these speeds cannot be seen as advertised/supported.
	Workaround: N/A
	Keywords: Ethernet
835239	Description: While running Q-in-Q packets with stag offloading, tcp-snoop/wireshark on host may show svlan ethertype as 0x8100 instead of 0x88A8.
	Workaround: Check the wire or a switch between the hosts, the wire-shark will show 0x88A8 ethertype as expected.
	Keywords: Ethernet
-	Description: After changing port type using connectx_port_config interface ports names can be changed. For example. ib1 -> ib0 if port1 changed to be Ethernet port and port2 left IB.
	Workaround: Use udev rules for persistent naming configuration. For further information, please refer to the User Manual
	Keywords: Port Type Management

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: OpenSM must be stopped prior to changing the port protocol from InfiniBand to Ethernet.</p> <p>Workaround: N/A</p> <p>Keywords: Port Type Management</p>
-	<p>Description: A working IP connectivity between the RoCE devices is required when creating an address handle or modifying a QP with an address vector.</p> <p>Workaround: N/A</p> <p>Keywords: Port Type Management</p>
-	<p>Description: IPv4 multicast over RoCE requires the MGID format to be as follows: :ffff:<Multicast IPv4 Address></p> <p>Workaround: N/A</p> <p>Keywords: Port Type Management</p>
-	<p>Description: IP routable RoCE does not support Multicast Listener Discovery (MLD) therefore, multicast traffic over IPv6 may not work as expected.</p> <p>Workaround: N/A</p> <p>Keywords: Port Type Management</p>
-	<p>Description: DIF: When running IO over FS over DM during unstable ports, block layer BIOS merges may cause false DIF error.</p> <p>Workaround: N/A</p> <p>Keywords: Port Type Management</p>
-	<p>Description: connectx_port_config configurations is not saved after unbind/bind.</p> <p>Workaround: Re-run "connectx_port_config"</p> <p>Keywords: Port Type Management</p>
954924	<p>Description: Accelerated Receive Flow Steering (aRFS) does not work properly with more than 50 streams. Thus, packets are not forwarded based on the location of the application consuming the packet.</p> <p>Workaround: N/A</p> <p>Keywords: Flow Steering</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: Flow Steering is disabled by default in firmware version < 2.32.5100.</p> <p>Workaround: To enable it, set the parameter below as follow: log_num_mgm_entry_size should set to -1</p> <p>Keywords: Flow Steering</p>
-	<p>Description: IPv4 rule with source IP cannot be created in SLES 11.x OSs.</p> <p>Workaround: N/A</p> <p>Keywords: Flow Steering</p>
-	<p>Description: RFS does not support UDP.</p> <p>Workaround: N/A</p> <p>Keywords: Flow Steering</p>
-	<p>Description: 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.</p> <p>Workaround: N/A</p> <p>Keywords: Flow Steering</p>
516136	<p>Description: Setting ARP flow rules through ethtool is not allowed.</p> <p>Workaround: N/A</p> <p>Keywords: Flow Steering</p>
-	<p>Description: QoS is not supported in XenServer, Debian 6.0 and 6.2 with uek kernel</p> <p>Workaround: N/A</p> <p>Keywords: Quality of Service</p>
-	<p>Description: When QoS features are not supported by the kernel, mlnx-_qos tool may crash.</p> <p>Workaround: N/A</p> <p>Keywords: Quality of Service</p>
448981	<p>Description: QoS default settings are not returned after configuring QoS.</p> <p>Workaround: N/A</p> <p>Keywords: Quality of Service</p>

Table 7 - Known Issues

Internal Ref	Issue
940345	Description: In ConnectX-3, when the virtual function (VF) runs on a MLNX_OFED version that is below v4.0, and the physical function runs on MLNX_OFED v4.0 and higher, hardware counters in the VF will be set to zero and will not progress.
	Workaround: N/A
	Keywords: Ethernet Performance Counters
-	Description: In ConnectX-3, 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.
	Workaround: N/A
	Keywords: Ethernet Performance Counters
-	Description: In ConnectX-3, 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.
	Workaround: N/A
	Keywords: Ethernet Performance Counters
-	Description: 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.
	Workaround: N/A
	Keywords: Ethernet Performance Counters
891241	Description: Sysfs for displaying neighbor information is not supported.
	Workaround: N/A
	Keywords: IPoIB
920440	Description: Adaptive RX moderation in not supported.
	Workaround: To improve RX performance, manually configure RX moderation using ethtool. It is recommended to use rx-usecs 16 and rx-frames 88 for datagram mode: Example: on ib0: ethtool -C ib0 rx-usecs 16 rx-frames 88
	Keywords: IPoIB

Table 7 - Known Issues

Internal Ref	Issue
965910	Description: On RHEL7.3, when creating a PKEY using the ifcfg file with ConnectX-3 and ConnectX-3 Pro adapter cards, <code>ifdown</code> and <code>ifup</code> commands must be run respectively.
	Workaround: N/A
	Keywords: IPoIB
854235	Description: IPoIB bonding interface has to be restarted in order to work on some operating systems.
	Workaround: Toggle bonding interface to state down and then to state up.
	Keywords: IPoIB
-	Description: When user increases receive/send a buffer, it might consume all the memory when few child's interfaces are created.
	Workaround: N/A
	Keywords: IPoIB
-	Description: The size of send queue in Connect-IB cards cannot exceed 1K.
	Workaround: N/A
	Keywords: IPoIB
-	Description: 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.
	Workaround: Set <code>ignore-carrier</code> for the corresponding device in <code>NetworkManager.conf</code> . For further information, please refer to " <code>man NetworkManager.conf</code> "
	Keywords: IPoIB
-	Description: IPoIB interface does not function properly if a third party application changes the PKey table.
	Workaround: We recommend modifying PKey tables via OpenSM.
	Keywords: IPoIB
-	Description: Fallback to the primary slave of an IPoIB bond does not work with ARP monitoring. (https://bugs.openfabrics.org/show_bug.cgi?id=1990)
	Workaround: N/A
	Keywords: IPoIB

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: Out-of memory issue might occur due to overload of interfaces created.</p> <p>Workaround: To calculate the allowed memory per each IPoIB interface check the following:</p> <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 <p>Keywords: IPoIB</p>
-	<p>Description: Connect-IB does not reach the bidirectional line rate</p> <p>Workaround: Optimize the IPoIB performance in Connect-IB: <code>cat /sys/class/net/<interface>/device/local_cpus > /sys/class/net/<interface>/queues/rx-0/rps_cpus</code></p> <p>Keywords: IPoIB</p>
-	<p>Description: If the <code>CONNECTED_MODE</code> parameter is set to no or missing from the ifcfg file for Connect-IB IPoIB interface then the "service network restart" will hang.</p> <p>Workaround: Set the <code>CONNECTED_MODE=yes</code> parameter in the ifcfg file for Connect-IB interface.</p> <p>Keywords: IPoIB</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: 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>Workaround: 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> <p>Keywords: IPoIB</p>
-	<p>Description: System might crash in <code>skb_checksum_help()</code> while performing TCP retransmit involving packets with 64k packet size. A similar output to the below will be printed: <pre>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</pre></p> <p>Workaround: Use UD mode in IPoIB.</p> <p>Keywords: IPoIB</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: 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>
	<p>Workaround: 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: <code>SUBSYSTEM=="net", ACTION=="add", DRIVERS=="?*", ATTR{address}=="*<Port GID>", NAME="ibN"</code></p> <p>Where N is the IPoIB required interface index</p>
	<p>Keywords: IPoIB</p>
-	<p>Description: After releasing a bond interface that contains IPoIB slaves, a call trace might be printed to the <code>dmesg</code>.</p>
	<p>Workaround: N/A</p>
	<p>Keywords: IPoIB</p>
-	<p>Description: IPoIB interfaces are loaded without an IP address on SLES 12.</p>
	<p>Workaround:</p> <ol style="list-style-type: none"> 1. Open the <code>/etc/wicked/common.xml</code> file. 2. Change: <code>"<use-nanny>false</use-nanny>"</code> to <code>"<use-nanny>true</use-nanny>"</code> 3. Run: <pre># systemctl restart wickedd.service wicked # ifup all</pre>
	<p>Keywords: IPoIB</p>
-	<p>Description: In RHEL7.0, running <code>ifdown</code> then <code>ifup</code> on an interface after changing <code>CONNECTED_MODE</code> in its <code>ifcfg</code> file, will cause the interface bring up to fail.</p>
	<p>Workaround: Reload the driver <code>"/etc/init.d/openibd restart"</code> or reboot the system.</p>
	<p>Keywords: IPoIB</p>
-	<p>Description: 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.</p>
	<p>Workaround: N/A</p>
	<p>Keywords: IPoIB</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: eth_ipoib module is not loaded after reloading the ib_ipoib module.</p> <p>Workaround: To restart the IPoIB driver, run <code>"/etc/init.d/openibd restart"</code>. Do not restart it by manually restarting each module.</p> <p>Keywords: IPoIB</p>
-	<p>Description: In Ubuntu and Debian, the default of the <code>recv_queue_size</code> and <code>send_queue_size</code> is 128 according to the <code>io_mmu</code> issue.</p> <p>Workaround: N/A</p> <p>Keywords: IPoIB</p>
-	<p>Description: In RHEL6.7, when the Network Manager service is enabled and an IPoIB interface is configured using the <code>"nm-connection-editor"</code> tool, the generated <code>ifcfg</code> file is missing the <code>"DEVICE=<interface name>"</code> parameter. Hence, changing the <code>CONNECTED_MODE</code> in the <code>ifcfg</code> file, will cause a failure in the interface bring up.</p> <p>Workaround: Either disable the Network Manager, or add <code>"DEVICE=<interface name>"</code> to the interface's <code>ifcfg</code> file.</p> <p>Keywords: IPoIB</p>
552840	<p>Description: <code>ifdown</code> command does not function in RH7.x</p> <p>Workaround: N/A</p> <p>Keywords: IPoIB</p>
665143	<p>Description: Kernel Oops may occur after reboot.</p> <p>Workaround: N/A</p> <p>Keywords: IPoIB</p>
555632	<p>Description: Kernel panic may occur while re-assigning LIDs.</p> <p>Workaround: N/A</p> <p>Keywords: IPoIB</p>
556352	<p>Description: ICMP traffic might be lost after Vnic restart</p> <p>Workaround: N/A</p> <p>Keywords: IPoIB</p>
560575	<p>Description: Spikes may occur while running PTP protocol over ConnectX-3/ConnectX-3 Pro.</p> <p>Workaround: N/A</p> <p>Keywords: IPoIB</p>

Table 7 - Known Issues

Internal Ref	Issue
684720	<p>Description: <code>ifdown</code> fails on SLES12SP0/SP1 with the following errors</p> <pre># ifdown ib0 wicked: ifdown: no matching interfaces</pre> <p>The error indicates that there are active interfaces using the interface you are trying to bring down, and you must <code>ifdown</code> all dependent interfaces.</p> <p>Workaround: To see the list of all dependent interfaces, run:</p> <pre># wicked --debug all ifdown ib0 wicked: skipping ib0 interface: unable to ifdown due to low- erdev dependency to: ib0.8001 wicked: ifdown: no matching interfaces wicked: Exit with status: 0</pre> <p>Keywords: IPoIB</p>
766451	<p>Description: Occasionally, in kernel 3.10, under heavy load, the kernel fails to get free page. For more details, please refer to: https://bugs.centos.org/view.php?id=10245</p> <p>Workaround: N/A</p> <p>Keywords: IPoIB</p>
383034	<p>Description: On rare occasions, upon driver restart the following message is shown in the <code>dmesg</code>:</p> <pre>'cannot create duplicate filename '/class/net/eth_ipoib_in- terfaces'</pre> <p>Workaround: N/A</p> <p>Keywords: eIPoIB</p>
-	<p>Description: Legacy API is deprecated, thus when recompiling applications over <code>MLNX_OFED v2.0-3.x.x</code>, warnings such as the below are displayed.</p> <pre>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)</pre> <p>These warnings can be safely ignored.</p> <p>Workaround: N/A</p> <p>Keywords: XRC</p>
-	<p>Description: XRC is not functional in heterogeneous clusters containing non Mellanox HCAs.</p> <p>Workaround: N/A</p> <p>Keywords: XRC</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: XRC options do not work when using qperf tool.</p> <p>Workaround: Use perftest instead</p> <p>Keywords: XRC</p>
-	<p>Description: 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.</p> <p>Workaround: N/A</p> <p>Keywords: XRC</p>
855362	<p>Description: A compilation error will occur in kernel space application when setting the wr_id field upon initializing any of the following structures: ib_wc, ib_send_wr, or ib_recv_wr. This is caused due to wr_id insertion into an anonymous union.</p> <p>Workaround: Assign the enum field explicitly. For example: <code>wr.wr_id = MY_WR_ID;</code></p> <p>Keywords: Verbs</p>
-	<p>Description: Using libnl1_1_3~26 or earlier, requires ibv_create_ah protection by a lock for multi-threaded applications.</p> <p>Workaround: N/A</p> <p>Keywords: Verbs</p>
-	<p>Description: 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.</p> <p>Workaround: When getting an event, poll the CQ until it is empty.</p> <p>Keywords: Verbs</p>
835061	<p>Description: According to the verbs header (/usr/include/infiniband/verbs.h), the static rate field in the Address handler can take value from 0 to 18. The values 11 to 18 (inclusive) are not supported for Connect-X 4 and Connect-X 3.</p> <p>Workaround: Run QP command query to verify the value.</p> <p>Keywords: Verbs</p>

Table 7 - Known Issues

Internal Ref	Issue
935250	Description: When NUM_OF_VFS in firmware capabilities is 32 or higher, ConnectX-4 RoCE LAG will not be supported. This is true even if driver does not have SR-IOV enabled (no VFs are present).
	Workaround: 1. Verify the NUM_OF_VFS value by running: <code>mlxconfig -d /dev/mst/mt4115_pci- conf0 query grep NUM_OF_VFS</code> Output: NUM_OF_VFS 32 2. Change NUM_OF_VFS value by running: <code>mlxconfig -d /dev/mst/mt4115_pci- conf0 set NUM_OF_VFS=0</code> 3. Reboot the machine.
	Keywords: RoCE
935250	Description: RoCE LAG for mlx5 is supported in kernel v4.5 and above.
	Workaround: N/A
	Keywords: RoCE
869158	Description: Occasionally, UC UD traffic over default GIDs, with high iterations may get stuck.
	Workaround: N/A
	Keywords: RoCE
854517	Description: Driver restart while having RDMA-CM running applications may hang.
	Workaround: N/A
	Keywords: RoCE
866072	Description: [mlx5] RoCE v2 multicast traffic using RDMA-CM with IPv4 address will not be received.
	Workaround: N/A
	Keywords: RoCE
-	Description: Not configuring the Ethernet devices or independent VMs with a unique IP address in the physical port, may result in RoCE GID table corruption.
	Workaround: Restart the driver.
	Keywords: RoCE
-	Description: 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.
	Workaround: N/A
	Keywords: RoCE

Table 7 - Known Issues

Internal Ref	Issue
392592	Description: On rare occasions, the driver reports a wrong GID table (read from /sys/class/infiniband/mlx4_*/ports/*/gids/*). This may cause communication problems.
	Workaround: N/A
	Keywords: RoCE
-	Description: MLNX_OFED v2.1-1.0.0 and onwards is not interoperable with older versions of MLNX_OFED.
	Workaround: N/A
	Keywords: RoCE
-	Description: 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: <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) Note also that each IP address occupies 2 entries when RoCE mode is set to 4 (RoCEv1 + RoCE v2). This further reduces the number of allowed IP addresses.
	Workaround: N/A
	Keywords: RoCE
-	Description: A working IP connectivity between the RoCE devices is required when creating an address handle or modifying a QP with an address vector.
	Workaround: N/A
	Keywords: RoCE
-	Description: IPv4 multicast over RoCE requires the MGID format to be as follows ::ffff:<Multicast IPv4 Address>
	Workaround: N/A
	Keywords: RoCE
-	Description: IP RoCEv2 does not support Multicast Listener Discovery (MLD) therefore, multicast traffic over IPv6 may not work as expected.
	Workaround: N/A
	Keywords: RoCE

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: 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 of the presence of the IPv6 address.</p> <p>Workaround: N/A</p> <p>Keywords: RoCE</p>
-	<p>Description: Using IPv6 link local address (GID0) when VLANs are configured is not supported.</p> <p>Workaround: N/A</p> <p>Keywords: RoCE</p>
-	<p>Description: Using GID index 0 (the default GID) on port 2 is currently not supported on kernel 3.14 and below.</p> <p>Workaround: N/A</p> <p>Keywords: RoCE</p>
559276/591244	<p>Description: Dynamically Connected (DC) in RoCE in ConnectX-4 is currently not supported.</p> <p>Workaround: N/A</p> <p>Keywords: RoCE</p>
-	<p>Description: Enslaving a Mellanox device to a bond with already configured IPs (or configured upper devices), prevents these IPs from being configured as GIDs.</p> <p>Workaround:</p> <ol style="list-style-type: none"> 1. Enslave the Mellanox device. 2. Configure IP devices. <p>Keywords: RoCE</p>
517825	<p>Description: <code>ibv_create_ah_from_wc</code> is not supported for multicast messages.</p> <p>Workaround: N/A</p> <p>Keywords: RoCE</p>
609950/649407	<p>Description: Occasionally, when the Bonding Mode is set to other than active/backup mode (mode 1), the GID table is not populated correctly.</p> <p>Workaround: Add slave devices to the master before giving it an IP address.</p> <p>Keywords: RoCE</p>

Table 7 - Known Issues

Internal Ref	Issue
667399	Description: In ConnectX-4 adapter cards, when the port speed is lower than 10Gbps, the IB tools will present a higher rate.
	Workaround: N/A
	Keywords: RoCE
778492	Description: RoCE requires that when a bonding module enslaves 2 Ethernet interfaces, the GID for any IP address on bond0 will appear only on the port of the active interface Due to kernel limitations, the information about active slave is unknown, therefore, any IP address on bond0 will appear on both ports.
	Workaround: Work in fail_over_mac mode (bonding).
	Keywords: RoCE
781383	Description: Creating Address Handler (AH) may run slow or may hang under a heavy load on all nodes cores (for example: MPI All2All cases).
	Workaround: N/A
	Keywords: RoCE
959452	Description: When adding an RXE device to one HCA port, <code>ibdev2net-dev</code> will show that the RXE device has been added to both HCA ports.
	Workaround: N/A
	Keywords: Soft RoCE
963537	Description: When adding an RXE device to an HCA that supports RoCE, this might cause segmentation fault when running verbs applications due to a conflict between the <code>librxe</code> and the HCA library (for example: <code>libmlx5</code>).
	Workaround: Perform one of the following to avoid the issue: <ul style="list-style-type: none"> • Uninstall the actual library (For example: <code>uninstall libmlx5</code>) • Remove the library's configuration file (Run: <code>rm -f /etc/libib-verbs.d/mlx5.driver</code>)
	Keywords: Soft RoCE
-	Description: 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.
	Workaround: N/A
	Keywords: iSCSI over IPoIB

Table 7 - Known Issues

Internal Ref	Issue
664110	Description: SDP is currently not supported in mlx5 driver (Connect-IB and Connect-X 4 adapter cards)
	Workaround: N/A
	Keywords: SDP
-	Description: Older versions of <code>rescan_scsi_bus.sh</code> may not recognize some newly created LUNs.
	Workaround: If encountering such issues, it is recommended to use the '-c' flag.
	Keywords: Storage
-	Description: NVMeoF support is available only for kernels of version 4.8.x and 4.9.x.
	Workaround: N/A
	Keywords: NVMeoF Host/Target
-	Description: When loading the <code>ib_srp</code> driver, the <code>reconnect_delay</code> and <code>fast_io_fail_tmo</code> parameters should be set in this order. Otherwise, the loading may fail.
	Workaround: N/A
	Keywords: SRP
-	Description: SRP daemon does not start at boot.
	Workaround: Add <code>service srpd start</code> to <code>rc.local</code> or start it manually.
	Keywords: SRP
-	Description: <code>srp_daemon</code> fails to connect on ConnectX-4 VF.
	Workaround: N/A
	Keywords: SRP
-	Description: MLNX_OFED SRP installation breaks the <code>ibmvstgt</code> and <code>ibmvscsi</code> symbol resolution in RHEL7.0
	Workaround: N/A
	Keywords: SRP
-	Description: The driver is tested with Storage target vendors recommendations for multipath.conf extensions (ZFS, DDN, TMS, Nimbus, NetApp).
	Workaround: N/A
	Keywords: SRP Interop

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: DDN does not accept non-default P_Key connection establishment.</p> <p>Workaround: N/A</p> <p>Keywords: DDN Storage Fusion 10000 Target</p>
-	<p>Description: Ungraceful power cycle of an initiator connected with Targets DDN, Nimbus, NetApp may result in temporary "stale connection" messages when initiator reconnects.</p> <p>Workaround: N/A</p> <p>Keywords: Oracle Sun ZFS Storage 7420</p>
-	<p>Description: On SLES OSs, the <code>ib_iser</code> module does not load on boot.</p> <p>Workaround: Add a dummy interface using <code>iscsiadm</code>:</p> <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> <p>Keywords: iSER Initiator</p>
-	<p>Description: Ubuntu16 requires update of user space <code>open-iscsi</code> to v2.0.874.</p> <p>Workaround: N/A</p> <p>Keywords: iSER Initiator</p>
-	<p>Description: The initiator does not respect interface parameter while logging in.</p> <p>Workaround: Configure each interface on a different subnet.</p> <p>Keywords: iSER Initiator</p>
-	<p>Description: <code>iSCSID</code> v2.0.873 can enter an endless loop on bind error.</p> <p>Workaround: N/A</p> <p>Keywords: iSER Initiator</p>
-	<p>Description: <code>iSCSID</code> may hang if target crashes during logout sequence (reproducible with TCP).</p> <p>Workaround: N/A</p> <p>Keywords: iSER Initiator</p>

Table 7 - Known Issues

Internal Ref	Issue
683370	<p>Description: iSER small read IO (< 8k) performance degrades compared to previous versions. iSER performs memory registration for each IO and avoids sending a global memory key to the target. Sending the global memory key to the wire should only be done in a trusted environment and is not recommended to use over the Internet protocol.</p> <p>Workaround: Set module param <code>always_register=N</code> <code>\$ modprobe ib_iser always_register=N</code></p> <p>Keywords: iSER Initiator</p>
-	<p>Description: iSER Target currently supports the following OSs (distribution kernel) only:</p> <ul style="list-style-type: none"> • RHEL 7.2/7.3 • SLES 12.1/12.2 • Ubuntu 14.04/15.04/16.04/16.10 <p>Workaround: N/A</p> <p>Keywords: iSER Target</p>
-	<p>Description: RHEL/CentOS 7.0: Discovery over RDMA is not supported.</p> <p>Workaround: N/A</p> <p>Keywords: iSER Target</p>
-	<p>Description: <code>ib_isert</code> is unavailable on custom kernels after running the <code>mlnx_add_kernel_support.sh</code> script.</p> <p>Workaround:</p> <ol style="list-style-type: none"> 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 <code>MLNX_OFED</code> for the custom kernel. <p>Keywords: iSER Target</p>
-	<p>Description: Connection establishment occurs twice which may cause iSER to log a stack trace.</p> <p>Workaround: N/A</p> <p>Keywords: ZFS Appliance</p>
-	<p>Description: The Erasure-coding logical block size must be aligned to 64 bytes</p> <p>Workaround: N/A</p> <p>Keywords: Erasure Coding Verbs</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: Only w=1,2,4,8 are supported (w corresponds to the Galois symbol size - GF(2^w))</p> <p>Workaround: N/A</p> <p>Keywords: Erasure Coding Verbs</p>
-	<p>Description: <code>ibv_exp_ec_mem</code> must pass with the following restrictions:</p> <ul style="list-style-type: none"> <code>num_data_sge</code> must be equal to K (property of the EC calc) <code>num_code_sge</code> must be equal to M (property of the EC calc) <p>Workaround: N/A</p> <p>Keywords: Erasure Coding Verbs</p>
960642	<p>Description: <code>[mlx5] min_tx_rate</code> and <code>max_tx_rate</code> limit per virtual function is not supported on ConnectX-5 and ConnectX-5 Ex adapter cards.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
896859	<p>Description: ConnectX-3 virtual function that runs on a MLNX_OFED version that is older than v4.0 cannot communicate with ConnectX-3 virtual function that runs on MLNX_OFED v4.0.</p> <p>Workaround: Set the OpenSM default alias-guid hop limit to 2 in OpenSM configuration file: <code>aguid_default_hop_limit 2</code></p> <p>Keywords: SR-IOV</p>
858628	<p>Description: PCI error handling is not supported during driver reload. This might cause a kernel panic or calltrace.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
860385	<p>Description: Creating 127 VFs may cause kernel panic in SLES11 SP4 KVM with Kernel 3.0.101-63 because of a IOMMU kernel bug.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
822781	<p>Description: SR-IOV is not supported in systems with a page size greater than 16KB since this is the maximal VF uar size supported.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>

Table 7 - Known Issues

Internal Ref	Issue
795697	Description: [mlx4] While spoof-check filters the incoming traffic to a VM, when this feature is disabled, traffic still does not reach the VM.
	Workaround: The driver must be restarted for the disablement of the feature to take effect and all traffic to be reached to the VM.
	Keywords: SR-IOV
835065	Description: [mlx5] When working with InfiniBand QoS, the bandwidth for VFs that are attached to VMs might not be spread according to the QoS configuration if not enough cores are assigned to the VM.
	Workaround: N/A
	Keywords: SR-IOV
784940	Description: Currently, the firmware cannot process many page requests in parallel as the driver processes page requests serially. Therefore, enabling/disabling a large number of VFs will often cause a driver slow-down.
	Workaround: N/A
	Keywords: SR-IOV
784954	Description: When SR-IOV is disabled, the VF driver receives <code>pci_err_r_detected</code> event and a teardown flow will be started. During the teardown flow, all firmware commands will fail because the function is already deleted.
	Workaround: N/A
	Keywords: SR-IOV
819595	Description: [ConnectX-3 Pro] In case a VF is set to VST mode on the same port following QinQ configuration, that VF will insert C-VLAN not only to untagged packets, but also to tagged packets. The packets that are tagged twice will be dropped by the switch or by the destination host since they have two C-VLANs.
	Workaround: N/A
	Keywords: SR-IOV
775944	Description: Bonding VFs on the same physical port using bonding mode 0 requires configuration of <code>fail_over_mac=1</code> .
	Workaround: N/A
	Keywords: SR-IOV

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: 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 <code>mlx4_core enable_64b_cqe_eqe=0</code> as a line in the file <code>/etc/modprobe.d/mlx4_core.conf</code>.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
381764	<p>Description: <code>mlx4_port1_mtu sysfs</code> entry shows a wrong MTU number in the VM.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
426988	<p>Description: When at least one port is configured as InfiniBand, and the <code>num_vfs</code> is provided but the <code>probe_vf</code> is not, HCA initialization fails.</p> <p>Workaround: Use both the <code>num_vfs</code> and the <code>probe_vf</code> in the <code>modprobe</code> line.</p> <p>Keywords: SR-IOV</p>
385750/378528	<p>Description: 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.</p> <p>Workaround: Unload the module <code>mlx4_ib</code> and reload it in the VM.</p> <p>Keywords: SR-IOV</p>
-	<p>Description: Attaching or detaching a Virtual Function on SLES11 SP3 to a guest Virtual Machine while the <code>mlx4_core</code> driver is loaded in the Virtual Machine may cause a kernel panic in the hypervisor.</p> <p>Workaround: Unload the <code>mlx4_core</code> module in the hypervisor before attaching or detaching a function to or from the guest.</p> <p>Keywords: SR-IOV</p>
392172	<p>Description: 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</p> <p>Workaround: Shut down the driver in the VM before detaching the VF.</p> <p>Keywords: SR-IOV</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: Enabling SR-IOV requires appending the <code>intel_iommu=on</code> option to the relevant OS in file <code>/boot/grub/grub.conf</code> or <code>/boot/grub2/grub.cfg</code>, depending on the OS installed. Without that SR-IOV cannot be loaded.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
-	<p>Description: 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.</p> <p>Workaround: Attach/detach VFs to/from a VM only while that VM is down.</p> <p>Keywords: SR-IOV</p>
-	<p>Description: The known PCI BDFs for all VFs in kernel command line should be specified by adding <code>xen-pciback.hide</code>. For further information, please refer to http://wiki.xen.org/wiki/Xen_PCI_Passthrough</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
-	<p>Description: The inbox qemu version (2.0) provided with Ubuntu 14.04 does not work properly when more than 2 VMs are run over an Ubuntu 14.04 Hypervisor.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
-	<p>Description: 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.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
-	<p>Description: Attempting to attach a PF to a VM when SR-IOV is already enabled on that PF may result in a kernel panic.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: osmttest 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.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
-	<p>Description: 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.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
-	<p>Description: QPerf test is not supported on SR-IOV guests in Connect-IB cards.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
-	<p>Description: On ConnectX-3 HCAs with firmware version 2.32.5000 and later, SR-IOV VPI mode works only with Port 1 = ETH and Port 2 = IB.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
-	<p>Description: Occasionally, the <code>lspci grep Mellanox</code> command shows incorrect or partial information due to the current <code>pci.ids</code> file on the machine.</p> <p>Workaround:</p> <ol style="list-style-type: none"> 1. Locate the file: <code>\$locate pci.ids</code> 2. Manually update the file according to the latest version available online at: https://pci-ids.ucw.cz/v2.2/pci.ids This file can also be downloaded using the following command: <code>update-pciids</code> <p>Keywords: SR-IOV</p>
-	<p>Description: SR-IOV is not supported in AMD architecture.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
506512	<p>Description: Setting 1 Mbit/s rate limit on Virtual Functions (Qos Per VF feature) may cause TX queue transmit timeout.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: DC transport type is not supported on SR-IOV VMs.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
567908	<p>Description: Attaching a VF to a VM before unbinding it from the hypervisor and then attempting to destroy the VM, may cause the system to hang for a few minutes.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
-	<p>Description: When using SR-IOV make sure to set interface to down and unbind BEFORE unloading driver/removing VF/restarting VM or kernel will lock. (reboot needed) Basically, clean-up might not work perfectly so user should do it manually.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
601749	<p>Description: Since the guest MAC addresses are configured to be all zeroes by default, in ConnectX-4 the administrator must explicitly set the VFs MAC addresses. otherwise the Guest VM will see MAC zero and traffic is not passed.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
649366	<p>Description: Restarting the PF (Hypervisor) driver while Virtual Functions are assigned is not allowed in RH7 and above due to a <code>vfio-pci</code> bug.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
639046	<p>Description: Due to an issue with SR-IOV loopback, prevention "Duplicate IPv6 detected" are seen in the VF driver.</p> <p>Workaround: N/A</p> <p>Keywords: SR-IOV</p>
655410	<p>Description: [ConnectX-4/Connect-IB] Failed to enable SR-IOV due to errors in PCI or BIOS.</p> <p>Workaround:</p> <ol style="list-style-type: none"> 1. Add <code>pci=realloc=on</code> to the grub command line. 2. Add more memory to the server. 3. Upgrade BIOS version. <p>Keywords: SR-IOV</p>

Table 7 - Known Issues

Internal Ref	Issue
651119	Description: Kernel panic may occur while running IPv6 UDP on SR-IOV ConnectX-4 environment
	Workaround: N/A
	Keywords: SR-IOV
669910	Description: Bind/Unbind over ConnectX-4 Hypervisor may cause system lockup.
	Workaround: N/A
	Keywords: SR-IOV
650458	Description: Occasionally, IPv6 might not function properly and cause lockup on SR-IOV ConnectX-4 environment.
	Workaround: N/A
	Keywords: SR-IOV
688551	Description: In ConnectX-3 adapter cards, the extended counter <code>port_rcv_data_64</code> on the VF may not be updated in some flows.
	Workaround: N/A
	Keywords: SR-IOV
690656/690674	Description: When the physical link is down, any traffic from the PF to any VF on the same port will be dropped.
	Workaround: N/A
	Keywords: SR-IOV
691661	Description: When in LAG mode and the Virtual Functions are present (VF LAG), the IP address given to the bonding interface (in the hypervisor) cannot be used for RoCE as well.
	Workaround: Probe one of the VFs in the hypervisor and use for RoCE.
	Keywords: SR-IOV
691661	Description: Ethernet SR-IOV in ConnectX-4 requires firmware version 12.14.1100 and higher
	Workaround: N/A
	Keywords: SR-IOV
737434	Description: VF vport statistics are not cleared upon <code>ifconfig</code> up/down.
	Workaround: N/A
	Keywords: SR-IOV

Table 7 - Known Issues

Internal Ref	Issue
738464	Description: In SLES11 SP4, user cannot open all VFs announced in <code>sriov_totalvfs</code> . However he can set the <code>num_vfs</code> up to maximum <code>sriov_totalvfs-1</code> vfs.
	Workaround: N/A
	Keywords: SR-IOV
784127	Description: While disabling SR-IOV, all firmware teardown flow commands are expected to fail and error messages will be reported in the <code>dmesg</code> .
	Workaround: N/A
	Keywords: SR-IOV
784146	Description: Creating/destroying as many as 64 VFs may sometimes take longer time than usual on some setups.
	Workaround: N/A
	Keywords: SR-IOV
766105	Description: Due to a bug in some QEMU versions, interrupts do not function properly for Virtual Functions. This causes the driver initialization to fail, and such error message will be printed: <code>"mlx4_core 0000:0b:00.0: command 0x31 timed out (go bit not cleared) mlx4_core 0000:0b:00.0: NOP command failed to generate interrupt (IRQ 57), aborting"</code> .
	Workaround: Upgrade to the latest version of QEMU in the hypervisor.
	Keywords: SR-IOV
-	Description: When working with InfiniBand QoS, the bandwidth for Virtual Functions (VF) that are attached to Virtual Machines (VM) might not be spread according to the QoS configuration if not enough cores are assigned to the Virtual Machine.
	Workaround: N/A
	Keywords: SR-IOV
413372	Description: SR-IOV non persistent configuration (such as VGT, VST, Host assigned GUIDs, and QP0-enabled VFs) may be lost upon Reset Flow.
	Workaround: Reset Admin configuration post Reset Flow
	Keywords: Reset Flow
-	Description: Upon Reset Flow or after running restart driver, Ethernet VLANs are lost.
	Workaround: Reset the VLANs using the <code>ifup</code> command.
	Keywords: Reset Flow

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: Restarting the driver or running <code>connectx_port_config</code> when Reset Flow is running might result in a kernel panic</p> <p>Workaround: N/A</p> <p>Keywords: Reset Flow</p>
-	<p>Description: Networking configuration (e.g. VLANs, IPv6) should be statically defined in order to have them set after Reset Flow as of after restart driver.</p> <p>Workaround: N/A</p> <p>Keywords: Reset Flow</p>
-	<p>Description: After recovering from an EEH event, <code>mlx5_core/mlx4_core</code> unload may fail due to a bug in some kernel versions. The bug is fixed in Kernel 3.15</p> <p>Workaround: N/A</p> <p>Keywords: Reset Flow</p>
965588	<p>Description: Lustre is not supported on MLNX_OFED v4.0.</p> <p>Workaround: Contact your Lustre provider.</p> <p>Keywords: General</p>
926137	<p>Description: IPV6 ping does not use device specified with <code>-I</code> parameter in <code>iputils</code> version s20160308.</p> <p>Workaround: Use <code>iputils</code> version s20161105 or above.</p> <p>Keywords: General</p>
936768	<p>Description: When querying the HCA core clock, data will be presented in KHz for all cards.</p> <p>Workaround: N/A</p> <p>Keywords: General</p>

Table 7 - Known Issues

Internal Ref	Issue
856033	<p>Description: The following PCIe bus error on Qualcomm ARM processor might appear when mapping a large number of DMA addresses: AER: Corrected error received: id=0000 PCIe Bus Error: severity=Corrected, type=Transaction Layer, id=0000(Receiver ID) device [17cb:0400] error status/mask=00002000/00004000 [13] Advisory Non-Fatal mlx5_warn:mlx5_0:dump_cqe:257:(pid 0): dump error cqe 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000 12007806 25000063 8728c8d3</p> <p>Workaround: Edit the kernel parameters (in grub) and add <code>qiommu.identity_map_qiommus=PCIE0_MMU,PCIE4_MMU</code> (The bus numbers depend on the ConnectX-4 slot.) Reboot the server.</p> <p>Keywords: General</p>
-	<p>Description: On 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 <code>ibstat</code>). <code>Mlxburn/flint</code> return <code>0xffff</code> 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.</p> <p>Workaround: N/A. Please use the GUID value returned by the fabric/driver utilities (not <code>0xffff</code>).</p> <p>Keywords: General</p>
552870/548518	<p>Description: On rare occasions, under extremely heavy MAD traffic, MAD (Management Datagram) storms might cause soft-lockups in the UMAD layer.</p> <p>Workaround: N/A</p> <p>Keywords: General</p>
-	<p>Description: Packets are dropped on the SM server on big clusters.</p> <p>Workaround: Increase the <code>recv_queue_size</code> of <code>ib_mad</code> module parameter for SM server to 8K. The <code>recv_queue_size</code> default size (4K)</p> <p>Keywords: General</p>

Table 7 - Known Issues

Internal Ref	Issue
663434	Description: On ConnectX-4/ConnectX-4 Lx, when running "lspci" in RH7.0/7.1, the device information is displayed incorrect or the device is unnamed.
	Workaround: Run <code>update-pciids</code>
	Keywords: General
767016	Description: Resetting hardware counters after netdev goes up can break statistics scripts.
	Workaround: N/A
	Keywords: General
959842	Description: User space libraries (for example: libibverbs, libmlx4/5) provided by MLNX_OFED v4.0 cannot work with kernel modules provided by an older MLNX_OFED version.
	Workaround: N/A
	Keywords: ABI Compatibility
-	Description: MLNX_OFED v2.3-1.0.1 is not ABI compatible with previous MLNX_OFED/OFED versions.
	Workaround: Recompile the application over the new MLNX_OFED version
	Keywords: ABI Compatibility
919836	Description: ucamtose fails when using a local loopback IP.
	Workaround: Use the device's interface IP instead of loopback IP.
	Keywords: Connection Manager (CM)
-	Description: When 2 different ports have identical GIDs, the CM might send its packets on the wrong port.
	Workaround: All ports must have different GIDs.
	Keywords: Connection Manager (CM)
781382	Description: The number of local ports that rdma_cm ID can bind to is limited. This limitation depends on the OS dynamics.
	Workaround: Modify the range of available ports for binding, run: <code>sysctl net.ipv4.ip_local_port_range="MIN MAX"</code> The MIN and MAX values can range from 0 to 65535.
	Note: Modifying the range also affects the range of available ports for socket applications (TCP/IP) even though the pool is not mutual between the RDMA stack and the TCP/IP stack.
	Keywords: Connection Manager (CM)

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: 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.</p> <p>Workaround: N/A</p> <p>Keywords: Fork Support</p>
-	<p>Description: On rare occasions, ConnectX-3 Pro adapter card may fail to link up when performing parallel detect to 40GbE.</p> <p>Workaround: Restart the driver</p> <p>Keywords: Uplinks</p>
-	<p>Description: 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.</p> <p>Workaround: N/A</p> <p>Keywords: Resources Limitation</p>
387061	<p>Description: <code>mlx4_core</code> can allocate up to 64 MSI-X vectors, an MSI-X vector per CPU.</p> <p>Workaround: N/A</p> <p>Keywords: Resources Limitation</p>
-	<p>Description: 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.</p> <p>Workaround: N/A</p> <p>Keywords: Resources Limitation</p>
553657	<p>Description: Registering a large amount of Memory Regions (MR) may fail because of DMA mapping issues on RHEL 7.0.</p> <p>Workaround: N/A</p> <p>Keywords: Resources Limitation</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: Occasionally, a user process might experience some memory shortage and not function properly due to Linux kernel occupation of the systems free memory for its internal cache.</p> <p>Workaround: To free up memory and allow it to be allocated in a user process, run the <code>drop_caches</code> procedure below. 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> <p>Keywords: Resources Limitation</p>
-	<p>Description: On ConnectX-4 Lx, the following may not be supported when using Multi-Packet WR flag (<code>IBV_EXP_QP_BURST_CREATE_ENABLE_MULTI_PACKET_SEND_WR</code>) on QP-burst family creation:</p> <ul style="list-style-type: none"> • ACLs • SR-IOV (eSwitch offloads) • priority and dscp forcing • Loopback decision. • VLAN insertion • encapsulation (encap/decap) • sniffer • Signature <p>Workaround: N/A</p> <p>Keywords: Accelerated Verbs</p>
-	<p>Description: perfest package in <code>MLNX_OFED v2.2-1.0.1</code> and onwards does not work with older versions of the driver.</p> <p>Workaround: N/A</p> <p>Keywords: Performance Tools</p>

Table 7 - Known Issues

Internal Ref	Issue
-	<p>Description: 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.</p>
	<p>Workaround: Run <code>ibdiagnet --skip nodes_info</code></p>
	<p>Keywords: Diagnostic Utilities</p>
-	<p>Description: <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.</p>
	<p>Workaround: Enable IPoIB Flow Steering and restart the driver. For further information, please refer to MLNX_OFED User Manual section Enable/Disable Flow Steering.</p>
	<p>Keywords: Diagnostic Utilities</p>
736136	<p>Description: The maximum number of HCAs shown by <code>ibstat</code> is 32 HCAs.</p>
	<p>Workaround: N/A</p>
	<p>Keywords: Diagnostic Utilities</p>
-	<p>Description: Running <code>ibdump</code> in InfiniBand mode with firmware older than v2.33.5000, may cause the server to hang due to a firmware issue.</p>
	<p>Workaround: Run <code>ibdump</code> with firmware v2.33.5000 and higher</p>
	<p>Keywords: Tools</p>

4 Bug Fixes History

This table lists the bugs fixed in this release.

Table 8 - Bug Fixes History

Internal Ref	Issue
919545	Description: Fixed the issue of when the Kernel becomes out of memory upon driver start, it could crash on SLES 12 SP2.
	Keywords: mlx Eth Driver
	Discovered in Release: 3.4-2.0.0.0
	Fixed in Release: 4.0-2.0.0.1
970668	Description: Fixed the issue where very high stress on DC QP transport might have triggered NMI messages on specific servers.
	Keywords: mlx5 Driver
	Discovered in Release: 4.0-1.0.1.0
	Fixed in Release: 4.0-2.0.0.1
966134	Description: Allowed Ethernet VFs to open Raw Ethernet QPs even if RoCE is not supported for the VF.
	Keywords: mlx4_ib
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 4.0-2.0.0.1
864063	Description: Fixed the issue of when Spoof-check may have been turned on for MAC address 00:00:00:00:00:00
	Keywords: mlx4
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 4.0-2.0.0.1
869209	Description: Fixed an issue that caused TCP packets to be received in an out of order manner when Large Receive Offload (LRO) is on.
	Keywords: mlx5_en
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 4.0-2.0.0.1
913319	Description: Fixed the issue of low performance when creating many address handles.
	Keywords: libibverbs
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 4.0-1.0.1.0

Table 8 - Bug Fixes History

Internal Ref	Issue
912897	Description: Added debug prints to <code>ib_umem_get</code> function to fix lack of error indication when this function fails.
	Keywords: InfiniBand
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 4.0-1.0.1.0
945887	Description: [ConnectX-3] Fixed the issue where multicast traffic over Raw Ethernet QP on virtual functions were received on the same QP (loopback).
	Keywords: SR-IOV
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 4.0-1.0.1.0
920292	Description: Fixed three issues in <code>libmlx5</code> that were found by NVIDIA in the patches that are part of <code>MLNX_OFED v3.4</code> : <ol style="list-style-type: none"> <code>mlx5_exp_peer_commit_qp</code> returns number of entries = 4 instead of 3. Peer capability check is wrong - should fail the check when there is neither NOR nor GEQ support. Missing break in <code>mlx5_exp_peer_peek_cq</code>. There is now fallback in the <code>IBV_EXP_PEER_PEEK_ABSOLUTE</code> case.
	Keywords: <code>libmlx5</code>
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 4.0-1.0.1.0
890285	Description: Fixed the issue where memory allocation for CQ buffers used to fail when increasing the RX ring size.
	Keywords: <code>mlx5_core</code>
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 4.0-1.0.1.0
867094	Description: Fixed the issue where <code>MLNX_OFED</code> used to fail to load on 4K page ARM architecture.
	Keywords: ARM
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 4.0-1.0.1.0
873538	Description: Fixed the issue where <code>biosdavename</code> running on Redhat 6.x with <code>MLNX_OFED</code> may show the same name to ConnectX-3 Eth port 1 and ConnectX-3 Eth port 2.
	Keywords: <code>biosdavename</code>
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 3.4-2.0.0.0

Table 8 - Bug Fixes History

Internal Ref	Issue
876329	Description: Fixed the issue of when the error flow was re-factored, the reading of the device caps was excluded from the error recovery flow.
	Keywords: mlx5 Driver
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 3.4-2.0.0.0
876419	Description: Fixed the issue where kernel panic was observed on openibd stop as a result of querying non-existent bond slave.
	Keywords: mlx4_en
	Discovered in Release: 3.3-2.0.0.0
	Fixed in Release: 3.4-2.0.0.0
868665	Description: Fixed the issue where kernel panic in <code>mlx4_en_get_phys_port_id</code> may occur during server reboot.
	Keywords: mlx4_en
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-2.0.0.0
882227	Description: Fixed the issue of when EEH was injected and the <code>mlx4</code> tear down code was called, the eqs were not released, causing a page fault.
	Keywords: mlx4_en
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 3.4-2.0.0.0
887348	Description: Fixed the issue of when <code>prof_sel</code> was invalid, <code>mlx5_core</code> failed upon debug print.
	Keywords: mlx5_core
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 3.4-2.0.0.0
898161	Description: Fixed the issue where a compilation error in kernels of v4.6 or above used to occur due to a large stack size in the <code>get_numa_phys_mask</code> function.
	Keywords: mlx5_core
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 3.4-2.0.0.0
880269	Description: Fixed the issue of when OFED was run on kernel v4.6 or higher, in which a memory management subsystem change was embedded, a kernel failure used to occur.
	Keywords: ib_core
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 3.4-2.0.0.0

Table 8 - Bug Fixes History

Internal Ref	Issue
887245	Description: Fixed the issue where the system used to pick the dummy <code>ibisert</code> module instead of the real module on RHEL with errata kernel.
	Keywords: <code>ibisert</code>
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 3.4-2.0.0.0
854344	Description: Fixed the issue where <code>mlnx_affinity</code> script on RHEL/CentOS7.x host did not disable or enable <code>irqbalancer</code> .
	Keywords: <code>irqbalancer</code>
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.0
824736	Description: Fixed wrong <code>skprio2UP</code> mapping by removing it and its scripts, such as <code>tc_wrap</code> , from the driver. This mapping should now be done using the kernel's <code>set_egress_map</code> commands. Note: Only for RDMACM over old kernels, the original <code>skprio2UP</code> mapping in <code>tc_wrap</code> remains valid as these kernels do not support <code>set_egress_map</code> .
	Keywords: QoS
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.0
824775	Description: Fixed the issue where starting <code>ibacm</code> daemon failed on Debian based distributions with the following message: <code>/etc/init.d/ibacm: line 37: /sbin/start_daemon: No such file or directory.</code>
	Keywords: <code>ibacm</code>
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.0
799004	Description: Fixed the issues of when establishing IPoIB CM connection, a race could occur if there were many CM connections taking place while the driver was going up and down. This race in the IPoIB driver could have caused memory corruption.
	Keywords: IPoIB
	Discovered in Release: 3.0-2.0.0.0
	Fixed in Release: 3.4-1.0.0.0
777733/778099	Description: Fixed the issue where in ARM architecture, multiple kernel panics of <code>mlx4</code> and <code>mlx5</code> drivers were observed as a result of undefined behavior of <code>vmap(virt_to_page(dma_alloc_coherent))</code> call sequence on driver load, by allocating contiguous memory instead of vmapping it.
	Keywords: ARM
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.0

Table 8 - Bug Fixes History

Internal Ref	Issue
826686	Description: Fixed the issue where server reboot could get stuck because of kernel panic in <code>mlx4_en_get_drvinfo()</code> that is called from asynchronous event handler.
	Keywords: <code>mlx4_en</code>
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.0
824130	Description: Fixed the issue where <code>ethtool</code> self test used to fail on interrupt test after timeout if <code>mlx4_ib</code> module was not loaded.
	Keywords: <code>mlx4_en</code>
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.0
855311	Description: Fixed the issue of when using RDMA READ with a higher value than 30 SGEs in the WR, this might have lead to local length error.
	Keywords: <code>mlx4 driver</code>
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.4-1.0.0.0
786720	Description: Fixed a crash that used to occur when trying to bring the interface up in a kernel that did not support accelerated RFS (aRFS).
	Keywords: <code>mlx5 driver</code>
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.0
781747	Description: Fixed the issue of when attempting to disable SR-IOV while there are any VF netdevs open, the operation would fail and the driver would hang.
	Keywords: SR-IOV
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.0
568602	Description: Fixed the issue of when repeating change of the <code>mlx5_num_vfs</code> value from 0 to non-zero might have caused kernel panic in the PF driver.
	Keywords: SR-IOV
	Discovered in Release: 3.0-2.0.0
	Fixed in Release: 3.4-1.0.0.0

Table 8 - Bug Fixes History

Internal Ref	Issue
814941	Description: Fixed the issue where as a result of the <code>ndo_set_mac_address</code> support in IPoIB, a memory corruption issue was exposed in the bonding driver over IPoIB interface. This memory corruption issue could have caused unpredictable behavior, such as ports becoming dysfunctional in RedHat, kernel panicking in Ubuntu devices, and other behaviors.
	Keywords: IPoIB
	Discovered in Release: 3.3-1.0.0
	Fixed in Release: 3.3-1.0.4.0
781183	Description: Fixed a potential security breach in the InfiniBand stack that was caused due to wrong reliance on the write system call. For more information, please see CVE-2016-4565 tracker.
	Keywords: IB Core
	Discovered in Release: 3.2-2.0.0 Fixed in Release: 3.3-1.0.0
592652	Description: Fixed the issue where the InfiniBand error counters found under <code>/sys/class/infiniband/<mlx5_dev>/ports/<port>/</code> did not function properly in ConnectX-4 adapter cards.
	Keywords: RoCE
	Discovered in Release: 3.1-0.0.7 Fixed in Release: 3.3-1.0.0
765162	Description: Fixed the issue where <code>dapl</code> package was missing in <code>MLNX_OFED</code> for Ubuntu PPC64LE.
	Keywords: Installation
	Discovered in Release: 2.0.2.0.5 Fixed in Release: 3.3-1.0.0.0
748308	Description: Changed TX queue counter format to: <code>xq_[tc]*[ring/channel]</code> .
	Keywords: TX Queue Counter
	Discovered in Release: 3.2-2.0.0 Fixed in Release: 3.3-1.0.0.0
751097	Description: Fixed RDMA sniffer functionality issues.
	Keywords: RDMA Sniffer
	Discovered in Release: 3.2-2.0.0 Fixed in Release: 3.3-1.0.0.0

Table 8 - Bug Fixes History

Internal Ref	Issue
751096	Description: Fixed IPoIB Connected Mode in ConnectX-3 functionality issues.
	Keywords: IPoIB
	Discovered in Release: 3.2-2.0.0
	Fixed in Release: 3.3-1.0.0.0
769688	Description: Fixed the issue where in order to change the IPoIB mode (connected/ datagram), the interface had to be taken down (via <code>ifconfig ibx down</code> or <code>ifdown ibx</code>). Now, the mode can be changed regardless of the interfaces state (up or down).
	Keywords: IPoIB
	Discovered in Release: 3.2-2.0.0
	Fixed in Release: 3.3-1.0.0.0
704756	Description: Added DCB PFC support through CEE netlink commands to prevent Priority Flow Control mode functionality issues on the host side.
	Keywords: mlx4_en
	Discovered in Release: 3.2-2.0.0
	Fixed in Release: 3.3-1.0.0.0
648680/655070	Description: Fixed an issue which added error messages to the dmesg when a VF used ethtool facilities.
	Keywords: SR-IOV
	Discovered in Release: 3.1-1.0.5
	Fixed in Release: 3.3-1.0.0.0
690772/690656	Description: Fixed an issue which caused any traffic from PF to any VF on the same port to drop when the physical link was down.
	Keywords: SR-IOV
	Discovered in Release: 3.2-1.0.1.1
	Fixed in Release: 3.3-1.0.0.0
708299	Description: Fixed kernels back-ports of XPS and affinity that did not have <code>CONFIG_CPUMASK_OFFSTACK</code>
	Keywords: mlx5 driver
	Discovered in Release: 3.2-1.0.1.1
	Fixed in Release: 3.2-2.0.0.0
685082	Description: Added support for Rate Limit 0 to enable unlimited rate limiter and to prevent max rate zero traffic lose.
	Keywords: mlx5 driver
	Discovered in Release: 3.2-1.0.1.1
	Fixed in Release: 3.2-2.0.0.0

Table 8 - Bug Fixes History

Internal Ref	Issue
667559	Description: Fixed an issue which enabled SR-IOV on RHEL 6.7 although SR-IOV was already enabled. A check was added to make sure SR-IOV is not enabled before enabling it.
	Keywords: SR-IOV
	Discovered in Release: 3.2-1.0.1.1
	Fixed in Release: 3.2-2.0.0.0
682750	Description: Fixed race between the udev that changes the interface name of eth_i-poib driver and the eIPoIB daemon that configured the same interface.
	Keywords: eIPoIB
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.2-2.0.0.0
692520	Description: Fixed an issue which prevented ConnectX-4/ConnectX-4 Lx adapter cards from running Ethernet traffic on Big Endian arch machines.
	Keywords: Ethernet traffic
	Discovered in Release: 3.2-1.0.1.1
	Fixed in Release: 3.2-2.0.0.0
668346	Description: Set close NUMA node as default for RSS.
	Keywords: Performance
	Discovered in Release: 3.2-1.0.1.1
	Fixed in Release: 3.2-2.0.0.0
696150	Description: Fixed an issue where the ARP request packets destined for a proxy VXLAN interface were not handled correctly when GRO was enabled.
	Keywords: mlx4_en
	Discovered in Release: 3.2-1.0.1.1
	Fixed in Release: 3.2-2.0.0.0
698795	Description: Fixed an issue which prevented the calculated software counters (the correct ones) from being shown and provided the error counters that were previously inactive.
	Keywords: Counters
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.2-2.0.0.0
597110	Description: Fixed an issue which prevented the driver from reaching VLAN when the VLAN was created over a Linux bridge.
	Keywords: Virtualization
	Discovered in Release: 3.1-1.0.3
	Fixed in Release: 3.2-1.0.1.1

Table 8 - Bug Fixes History

Internal Ref	Issue
656298	Description: Fixed an issue in the driver (in ConnectX-4) that discarded s-tag VLAN packets when in Promiscuous Mode.
	Keywords: mlx5 driver
	Discovered in Release: 3.1-1.0.3
	Fixed in Release: 3.2-1.0.1.1
647865	Description: Fixed an issue which prevented PORT_ERR event to be propagated to the user-space application when the port state was changed from Active to Initializing.
	Keywords: mlx5 driver
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.2-1.0.1.1
663975	Description: Fixed a rare issue which allowed the knem package to run depmod on the wrong kernel version.
	Keywords: HPC Acceleration packages
	Discovered in Release: 3.1-1.0.3
	Fixed in Release: 3.2-1.0.1.1
666992	Description: Fixed a race condition in the IB/umad layer that caused NULL pointer dereference.
	Keywords: IB/Core
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.2-1.0.1.1
657718	Description: Fixed an IPoIB issue that caused connectivity lost after servers restart in a cluster.
	Keywords: IPoIB
	Discovered in Release: 3.1-1.0.3
	Fixed in Release: 3.2-1.0.1.1
619272	Description: Fixed an issue causing MLNX_OFED to remove the mutt package upon driver uninstall.
	Keywords: Driver un-installation
	Discovered in Release: 3.1-1.0.3
	Fixed in Release: 3.2-1.0.1.1
613514	Description: Added a warning message in dmesg, notifying the user that the PFC RX/TX cannot be enabled simultaneously with Global Pauses. In this case Global Pauses will be disabled.
	Keywords: PFC
	Discovered in Release: 3.1-1.0.3
	Fixed in Release: 3.2-1.0.1.1

Table 8 - Bug Fixes History

Internal Ref	Issue
606916	Description: Fixed an issue causing MADs to drop in large scale clusters.
	Keywords: IB MAD
	Discovered in Release: 3.1-1.0.0
	Fixed in Release: 3.1-1.0.3
367410	Description: Fixed InfiniBand counters which were unavailable in the VM.
	Keywords: SR-IOV
	Discovered in Release: 2.1-1.0.0
	Fixed in Release: 3.1-1.0.0
549687	Description: Fixed InfiniBand traffic counters that are found under /sys/class/infini-band/<mlx5_dev>/ports/<port>/ which dis not function properly in ConnectX-4 adapter cards.
	Keywords: RoCE
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.1-1.0.0
589247/591877	Description: Fixed VXLAN functionality issues.
	Keywords: Virtualization
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
-	Description: TCP/UDP latency on ConnectX-4 was higher than expected.
	Keywords: Performance
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
-	Description: TCP throughput on ConnectX-4 achieved full line rate.
	Keywords: Performance
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
568718	Description: Fixed an issue causing inconsistent performance with ConnectX-3 and PowerKVM 2.1.1.
	Keywords: Performance
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
552658	Description: Fixed ConnectX-4 traffic counters.
	Keywords: Performance
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0

Table 8 - Bug Fixes History

Internal Ref	Issue
572068	Description: Updated the desired num_entries in each iteration, and accordingly updated the offset of the WC in the given WC array.
	Keywords: num_entries
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.1-1.0.0
536981/554293	Description: Fixed incorrect port rate and port speed values in RoCE mode in ConnectX-4.
	Keywords: mlx5 driver
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
551898	Description: In RedHat7.1 kernel 3.10.0-299, when sending ICMP/TCP/UDP traffic over Connect-IB/ConnectX-4 in UD mode, the packets were dropped with the following error: UDP: bad checksum...
	Keywords: IPoIB
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
596458	Description: Fixed an issue which prevented openibd from starting correctly during boot.
	Keywords: openibd
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
589207	Description: Added a new module parameter to control the number of IRQs allocated to the device.
	Keywords: Ethernet
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
576326	Description: Fixed an issue on PPC servers which prevented PCI from reloading after EEH error recovery.
	Keywords: mlx5 driver
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
569369	Description: Fixed an issue which prevented the OpenSM package from being fully removed when uninstalling MLNX_OFED v3.0-2.0.1
	Keywords: OpenSM
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0

Table 8 - Bug Fixes History

Internal Ref	Issue
568169	Description: Added the option to toggle LRO ON/OFF using the "-κ" flags. The priv flag hw_lro will determine the type of LRO to be used, if the flag is ON, the hardware LRO will be used, otherwise the software LRO will be used.
	Keywords: mlx5_en
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
568168	Description: Added the option to toggle LRO ON/OFF using the "-κ" flags.
	Keywords: mlx5_en
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
551075	Description: Fixed race when updating counters.
	Keywords: mlx5_en
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
550275	Description: Fixed scheduling while sending atomic dmesg warning during bonding configuration.
	Keywords: mlx5_en
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
550824	Description: Added set_rx_csum callback implementation.
	Keywords: mlx5_en
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
535884	Description: Fixed mismatch between SL and VL in outgoing QP1 packets, which caused buffer overruns in attached switches at high MAD rates.
	Keywords: mlx4_ib
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.1-1.0.0
542722	Description: Fixed a problem on VFs where the RoCE driver registered a zero MAC into the port's MAC table (during QP1 creation) because the ETH driver had not yet generated a non-zero random MAC for the ETH port.t
	Keywords: SR-IOV/RoCE
	Discovered in Release: 2.3-1.0.1
	Fixed in Release: 3.1-1.0.0

Table 8 - Bug Fixes History

Internal Ref	Issue
561866	Description: Removed BUG_ON assert when checking if the ring is full.
	Keywords:
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.1-1.0.0
541149	Description: Added libvma support for Debian 8.0 x86_64 and Ubuntu 15.04
	Keywords: libvma
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 3.1-1.0.0
-	Description: Fixed an issue which prevented the failure to destroy QP upon IPoIB unload on debug kernel.
	Keywords: IPoIB
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.0-2.0.0
-	Description: Fixed an issue which prevented the driver version to be reported to the Remote Access Controller tools (such as iDRAC)
	Keywords: Configuration
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.0-2.0.0
-	Description: Passed the correct port number in port-change event to single-port VFs, where the actual physical port used is port 2.
	Keywords: SR-IOV
	Discovered in Release: 2.4-1.0.0
	Fixed in Release: 3.0-2.0.0
-	Description: Enabled OpenSM, running over a ConnectX-3 HCA, to manage a mixed ConnectX-3/ConnectX-4 network (by recognizing the "Well-known GID" in mad demux processing).
	Keywords: SR-IOV
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.0-2.0.0
-	Description: Fixed double-free memory corruption in case where SR-IOV enabling failed (error flow).
	Keywords: SR-IOV
	Discovered in Release: 3.0-1.0.1
	Fixed in Release: 3.0-2.0.0

Table 8 - Bug Fixes History

Internal Ref	Issue
-	<p>Description: Fixed a crash in EQ's initialization error flow.</p> <p>Keywords: Start-up sequence</p> <p>Discovered in Release: 3.0-1.0.1</p> <p>Fixed in Release: 3.0-2.0.0</p>
554253	<p>Description: MLNX_OFED v3.0-1.0.1 installation using yum fails on RH7.1</p> <p>Keywords: Installation</p> <p>Discovered in Release: 3.0-1.0.1</p> <p>Fixed in Release: 3.0-2.0.0</p>
542686	<p>Description: In PPC systems, when working with ConnectX-4 adapter card configured as Ethernet, driver load fails with BAD INPUT LENGTH. dmesg: "command failed, status bad input length(0x50), syndrome 0x9074aa"</p> <p>Keywords: mlx5 driver</p> <p>Discovered in Release: 3.0-1.0.1</p> <p>Fixed in Release: 3.0-2.0.0</p>
-	<p>Description: Error counters such as: CRC error counters, RX out range length error counter, are missing in the ConnectX-4 Ethernet driver.</p> <p>Keywords: mlx5 driver</p> <p>Discovered in Release: 3.0-1.0.1</p> <p>Fixed in Release: 3.0-2.0.0</p>
-	<p>Description: Changing the RX queues number is not supported in Ethernet driver when connected to a ConnectX-4 card.</p> <p>Keywords: mlx5 driver</p> <p>Discovered in Release: 3.0-1.0.1</p> <p>Fixed in Release: 3.0-2.0.0</p>
-	<p>Description: Hardware checksum call trace may appear when receiving IPV6 traffic on PPC systems that uses CHECKSUM COMPLETE method.</p> <p>Keywords: Ethernet</p> <p>Discovered in Release: 3.0-1.0.1</p> <p>Fixed in Release: 3.0-2.0.0</p>
-	<p>Description: Fixed ping/traffic issue occurred when RXVLAN offload was disabled and CHECKSUM COMPLETE was used on ingress packets.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.4-1.0.4</p> <p>Fixed in Release: 3.0-1.0.1</p>

Table 8 - Bug Fixes History

Internal Ref	Issue
-	<p>Description: CVE-2014-8159 Fix: Prevented integer overflow in IB-core module during memory registration.</p> <p>Keywords: Security</p> <p>Discovered in Release: 2.0-2.0.5</p> <p>Fixed in Release: 2.4-1.0.4</p>
-	<p>Description: Fixed the return value of max inline received size in the created QP.</p> <p>Keywords: mlx5_ib</p> <p>Discovered in Release: 2.3-2.0.1</p> <p>Fixed in Release: 2.4-1.0.0</p>
-	<p>Description: Resolved soft lock on massive amount of user memory registrations</p> <p>Keywords: mlx5_ib</p> <p>Discovered in Release: 2.3-2.0.1</p> <p>Fixed in Release: 2.4-1.0.0</p>
-	<p>Description: Occasionally, port_rcv_data and port_xmit_data counters may not function properly.</p> <p>Keywords: InfiniBand Counters</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.4-1.0.0</p>
-	<p>Description: LRO fixes and improvements for jumbo MTU.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.3-2.0.1</p> <p>Fixed in Release: 2.4-1.0.0</p>
-	<p>Description: Fixed a crash occurred when changing the number of rings (ethtool set-channels) when interface connected to netconsole.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.2-1.0.1</p> <p>Fixed in Release: 2.4-1.0.0</p>
-	<p>Description: Fixed ping issues with IP fragmented datagrams in MTUs 1600-1700.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.2-1.0.1</p> <p>Fixed in Release: 2.4-1.0.0</p>

Table 8 - Bug Fixes History

Internal Ref	Issue
-	<p>Description: 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.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.4-1.0.0</p>
-	<p>Description: 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 (contiguous pages).</p> <p>Keywords: mlx5_ib</p> <p>Discovered in Release: 2.3-2.0.1</p> <p>Fixed in Release: 2.3-2.0.5</p>
-	<p>Description: Fixed an issue in verbs API: fallback to glibc on contiguous memory allocation failure</p> <p>Keywords: mlx5_ib</p> <p>Discovered in Release: 2.3-2.0.1</p> <p>Fixed in Release: 2.3-2.0.5</p>
-	<p>Description: Fixed a memory corruption issue in multi-core system due to intensive IPoIB transmit operation.</p> <p>Keywords: IPoIB</p> <p>Discovered in Release: 2.3-2.0.1</p> <p>Fixed in Release: 2.3-2.0.5</p>
-	<p>Description: Fixed an issue to prevent process starvation due to MAD packet storm.</p> <p>Keywords: IB MAD</p> <p>Discovered in Release: 2.3-2.0.1</p> <p>Fixed in Release: 2.3-2.0.5</p>
433348	<p>Description: Fixed an issue which prevented the spread of events among the closet NUMA CPU when only a single RX queue existed in the system.</p> <p>Keywords: IPoIB</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Returned the CQ to its original state (armed) to prevent traffic from stopping</p> <p>Keywords: IPoIB</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>

Table 8 - Bug Fixes History

Internal Ref	Issue
-	<p>Description: Fixed a TX timeout issue in CM mode, which occurred under heavy stress combined with ifup/ifdown operation on the IPoIB interface.</p> <p>Keywords: IPoIB</p> <p>Discovered in Release: 2.1-1.0.0</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Fixed "sleeping while atomic" error occurred when the driver ran many firmware commands simultaneously.</p> <p>Keywords: mlx4_core</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Fixed an issue related to spreading of completion queues among multiple MSI-X vectors to allow better utilization of multiple cores.</p> <p>Keywords: mlx4_ib</p> <p>Discovered in Release: 2.1-1.0.0</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Fixed an issue that caused an application to fail when attaching Shared Memory.</p> <p>Keywords: mlx4_ib</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Fixed dmesg warnings: "NOHZ: local_softirq_pending 08".</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Fixed erratic report of hardware clock which caused bad report of PTP hardware Time Stamping.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.1-1.0.0</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Fixed race when async events arrived during driver load.</p> <p>Keywords: mlx5_core</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>

Table 8 - Bug Fixes History

Internal Ref	Issue
-	<p>Description: Fixed race in mlx5_eq_int when events arrived before eq->dev was set.</p> <p>Keywords: mlx5_core</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Enabled all pending interrupt handlers completion before freeing EQ memory.</p> <p>Keywords: mlx5_core</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Defined mlnx.conf as a configuration file in mlnx-ofa_kernel RPM</p> <p>Keywords: mlnx.conf</p> <p>Discovered in Release: 2.1-1.0.0</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Fixed counter index allocation for VFs which enables Ethernet port statistics.</p> <p>Keywords: SR-IOV</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Fixed iSER DIX sporadic false DIF errors caused in large transfers when block merges were enabled.</p> <p>Keywords: iSER</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: RoCE v2 was non-functional on big Endian machines.</p> <p>Keywords: RoCE v2</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: Fixed registration memory failure when fork was enabled and contiguous pages or ODP were used.</p> <p>Keywords: Verbs</p> <p>Discovered in Release: 2.3-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>

Table 8 - Bug Fixes History

Internal Ref	Issue
-	<p>Description: Using both '-c --config' and '--add-kernel-support' flags simultaneously when running the mlnxofedinstall.sh script caused installation failure with the following on screen message "--config does not exist".</p> <p>Keywords: Installation</p> <p>Discovered in Release: 2.2-1.0.1</p> <p>Fixed in Release: 2.3-2.0.0</p>
-	<p>Description: 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.</p> <p>Keywords: IPoIB</p> <p>Discovered in Release: 2.1-1.0.0</p> <p>Fixed in Release: 2.3-1.0.1</p>
-	<p>Description: XRC over ROCE in SR-IOV mode is not functional</p> <p>Keywords: XRC</p> <p>Discovered in Release: 2.0-3.1.0</p> <p>Fixed in Release: 2.2-1.0.1</p>
-	<p>Description: Fixed wrong calculation of packet true-size reporting in LRO flow.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.1-1.0.0</p> <p>Fixed in Release: 2.2-1.0.1</p>
-	<p>Description: Fixed kernel panic on Debian-6.0.7 which occurred when the number of TX channels was set above the default value.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.1-1.0.0</p> <p>Fixed in Release: 2.2-1.0.1</p>
-	<p>Description: Fixed a crash incidence which occurred when enabling Ethernet Time-stamping and running VLAN traffic.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.0-2.0.5</p> <p>Fixed in Release: 2.2-1.0.1</p>
-	<p>Description: Fixed the QP attribute mask upon smac resolving</p> <p>Keywords: IB Core</p> <p>Discovered in Release: 2.1-1.0.0</p> <p>Fixed in Release: 2.1-1.0.6</p>

Table 8 - Bug Fixes History

Internal Ref	Issue
-	Description: Fixed a send WQE overhead issue
	Keywords: mlx5_ib
	Discovered in Release: 2.1-1.0.0
	Fixed in Release: 2.1-1.0.6
-	Description: Fixed a NULL pointer de-reference on the debug print
	Keywords: mlx5_ib
	Discovered in Release: 2.1-1.0.0
	Fixed in Release: 2.1-1.0.6
-	Description: Fixed arguments to kzalloc
	Keywords: mlx5_ib
	Discovered in Release: 2.1-1.0.0
	Fixed in Release: 2.1-1.0.6
-	Description: Fixed the locks around completion handler
	Keywords: mlx4_core
	Discovered in Release: 2.1-1.0.0
	Fixed in Release: 2.1-1.0.6
-	Description: Restored port types as they were when recovering from an internal error.
	Keywords: mlx4_core
	Discovered in Release: 2.0-2.0.5
	Fixed in Release: 2.1-1.0.0
-	Description: Added an N/A port type to support port_type_array module param in an HCA with a single port
	Keywords: mlx4_core
	Discovered in Release: 2.0-2.0.5
	Fixed in Release: 2.1-1.0.0
-	Description: Fixed memory leak in SR-IOV flow.
	Keywords: SR-IOV
	Discovered in Release: 2.0-2.0.5
	Fixed in Release: 2.0-3.0.0
-	Description: Fixed communication channel being stuck
	Keywords: SR-IOV
	Discovered in Release: 2.0-2.0.5
	Fixed in Release: 2.0-3.0.0

Table 8 - Bug Fixes History

Internal Ref	Issue
-	<p>Description: Fixed ALB bonding mode failure when enslaving Mellanox interfaces</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.0-3.0.0</p> <p>Fixed in Release: 2.1-1.0.0</p>
-	<p>Description: Fixed leak of mapped memory</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.0-3.0.0</p> <p>Fixed in Release: 2.1-1.0.0</p>
-	<p>Description: Fixed TX timeout in Ethernet driver.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.0-2.0.5</p> <p>Fixed in Release: 2.0-3.0.0</p>
-	<p>Description: Fixed ethtool stats report for Virtual Functions.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.0-2.0.5</p> <p>Fixed in Release: 2.0-3.0.0</p>
-	<p>Description: Fixed an issue of VLAN traffic over Virtual Machine in paravirtualized mode.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.0-2.0.5</p> <p>Fixed in Release: 2.0-3.0.0</p>
-	<p>Description: Fixed ethtool operation crash while interface down.</p> <p>Keywords: mlx4_en</p> <p>Discovered in Release: 2.0-2.0.5</p> <p>Fixed in Release: 2.0-3.0.0</p>
-	<p>Description: Fixed memory leak in Connected mode.</p> <p>Keywords: IPoIB</p> <p>Discovered in Release: 2.0-2.0.5</p> <p>Fixed in Release: 2.0-3.0.0</p>
-	<p>Description: Fixed an issue causing IPoIB to avoid pkey value 0 for child interfaces.</p> <p>Keywords: IPoIB</p> <p>Discovered in Release: 2.0-2.0.5</p> <p>Fixed in Release: 2.0-3.0.0</p>

5 Change Log History

Table 9 - Change Log History

Category	Description
4.0-2.0.0.1	
PCIe Error Counting	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added the ability to expose physical layer statistical counters to ethtool.
Multiprotocol Label Switching (MPLS) Tagged Packets Classification	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Enabled packet flow steering rules with IPv4/IPv6 classification (for raw packet QP (DPDK) only) to work on IPv4/IPv6 over MPLS (Ethernet 0x8847 and 0x8848) encapsulated packets.
RoCE VFs	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added the ability to enable/disable RoCE on VFs.
RoCE LAG	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for RoCE over LAG interface.
Standard ethtool	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for flow steering and rx-all mode.
SR-IOV Bandwidth Share for Ethernet/RoCE (beta)	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added the ability to guarantee the minimum rate of a certain VF in SR-IOV mode.
Adapter Cards	[ConnectX-5/ConnectX-5 Ex] Added support for ConnectX-5 and ConnectX-5 Ex HCAs.
DSCP ConfigFS Control for RDMA-CM QPs	Added the ability to configure ToS/DSCP for RDMA-CM QPs only.
Soft RoCE (beta)	Add software implementation of RoCE that allows RoCE to run on any Ethernet network adapter whether it offers hardware acceleration or not.
NVMe over Fabrics (NVMeoF)	NVMeoF related module installation has been disabled by default. In order to enable it, add the "--with-nvme" installation option to the "mlnxofedinstall" script.
3.4-2.0.0.0	
NVMeoF	Added support for NVMeoF in host/target systems over RDMA.
3.4-1.0.0.0	
VST Q-in-Q	[ConnectX®-3/ConnectX®-3 Pro] Added support for Q-in-Q encapsulation per VF in Linux (VST) for ConnectX-3 Pro adapter cards.

Table 9 - Change Log History

Category	Description
Package Content	<p>[ConnectX®-3/ConnectX®-3 Pro] SR-IOV enabled firmware binaries for ConnectX-3 has been removed from MLNX_OFED package (the installation flag “--enable-sriov” has been deprecated). To configure SR-IOV, please use the “mlxconfig” or “mstconfig” utilities.</p> <p>[ConnectX®-3/ConnectX®-3 Pro] MLNX_OFED repository metadata files has been moved to the folder holding the binary packages (named “RPMS” in rpm based OS, and “DEBS” in Debian based OS). Please update your repository configuration file accordingly (refer to the MLNX_OFED User Manual for more details about setting up MLNX_OFED as a repository).</p>
Raw Ethernet Programming	<p>[ConnectX®-4/ConnectX®-4 Lx] Added new APIs for enhanced raw Ethernet programming:</p> <ul style="list-style-type: none"> • Packet Pacing • TCP Segmentation Offload (TSO) • ToS based steering • Flow ID based steering (beta) • VxLAN based steering (beta) <p>For further information, refer to the “Programming” section in OFED User Manual.</p>
Enhanced PCIe Error Recovery	<p>[ConnectX®-4/ConnectX®-4 Lx] Enhanced PCIe error recovery by adding the following behaviors to the flow:</p> <ul style="list-style-type: none"> • In case SR-IOV is enabled during the recovery process, it will not get automatically disabled and will require the administrator that enabled it to disable it. • When the driver goes down, VF PCI function will not be removed. • Ethernet interface attributes (MTU, state, ring size, etc...) will be recovered after the error recovery stage is completed. • The net device kernel layer will not be aware of any ongoing PCI error recovery process.
SR-IOV Max Rate Limit Ethernet/RoCE (beta level)	<p>[ConnectX®-4/ConnectX®-4 Lx] Added the ability to rate-limit traffic per Virtual Function in SR-IOV mode.</p>
Dynamically tuned Interrupt Moderation (DIM)	<p>[ConnectX®-4/ConnectX®-4 Lx] Added support for dynamically controlling the interrupts per channel to ensure maximum packet rate with minimum interrupt rate. This feature is enabled by default.</p>
Dump Configuration	<p>[ConnectX®-4/ConnectX®-4 Lx] Added support for dump configuration which helps dumping driver and firmware configuration using ethtool. It creates a backup of the configuration files into a specified dump file.</p>
Ethernet Counters	<p>[ConnectX®-4/ConnectX®-4 Lx] Updated the list of counters the can be retrieved via ethtool for mlx5 driver, changed counters names and added new counters.</p>

Table 9 - Change Log History

Category	Description
Mellanox PeerDirect Async (beta level)	<p>[ConnectX®-3/ConnectX®-3 Pro/ConnectX®-4/ConnectX®-4 Lx] The experimental PeerDirect Async APIs have been changed to make the implementation of peer clients simpler. Note the following:</p> <ul style="list-style-type: none"> • These changes are not backward compatible. • The code adds CQ polling support for peer devices that do not support the NOR operation, replacing it with a GEQ operation. <p>To see the API changes, refer to the man page.</p>
ABI Incompatibility	<p>[ConnectX®-3/ConnectX®-3 Pro/ConnectX®-4/ConnectX®-4 Lx] Added the ability to fix the issue of preventing the load of MLNX_EN modules when a new kernel is not compatible with these modules.</p>
Mellanox Scalable Hierarchical Aggregation Protocol (SHARP™)	<p>[Connect-IB/ConnectX®-3/ConnectX®-3 Pro/ConnectX®-4] IB only: This technology improves the performance of MPI operation by offloading collective operations from the CPU and dispatching to the switch network, and eliminating the need to send data multiple times between endpoints. This approach decreases the amount of data traversing the network as aggregation nodes are reached, and dramatically reduces the MPI operation time. For further information on SHARP and its configuration, see SHARP Deployment Guide.</p>
3.3-1.0.0.0	
VF MAC Address Anti-Spoofing	<p>[ConnectX-4/ConnectX-4 Lx] Also known as MAC spoof-check, the VF MAC Address Anti-Spoofing prevents malicious VFs from faking their MAC addresses.</p>
VF All-multi Mode	<p>[ConnectX-4/ConnectX-4 Lx] Added support for the VF to enter all-multi RX mode, meaning that in addition to the traffic originally targeted to the VF, it will receive all the multicast traffic sent from/to the other functions on the same physical port. Note: Only privileged/trusted VFs can enter the all-multi RX mode.</p>
VF Promiscuous Mode	<p>[ConnectX-4/ConnectX-4 Lx] Added support for the VF to enter promiscuous RX mode, meaning that in addition to the traffic originally targeted to the VF, it will receive the unmatched traffic and all the multicast traffic that reaches the physical port. The unmatched traffic is any traffic's DMAC that does not match any of the VFs' or PFs' MAC addresses. Note: Only privileged/trusted VFs can enter the promiscuous RX mode.</p>
Privileged VF	<p>[ConnectX-4/ConnectX-4 Lx] Added support for determining privileged/trusted VFs so security sensitive features can be enabled for these VFs, such as entering promiscuous and all-multi RX modes.</p>
DCBX	<p>[ConnectX-4/ConnectX-4 Lx] Added support for standard DCBX CEE API.</p>
Per Priority Counters	<p>[ConnectX-4/ConnectX-4 Lx] Exposed performance counters per priority.</p>
IB Error Counters	<p>[ConnectX-4/ConnectX-4 Lx] Exposed IB sysfs error counters for mlx5 driver.</p>
Accelerated Receive Flow Steering (aRFS)	<p>[ConnectX-4/ConnectX-4 Lx] Boosts the speed of RFS by adding hardware assistance. RFS is an in-kernel-logic responsible for load balancing between CPUs by attaching flows to CPUs that are used by flow's owner applications.</p>

Table 9 - Change Log History

Category	Description
Packet Pacing for UDP/TCP	[ConnectX-4/ConnectX-4 Lx] Performs rate limit per UDP/TCP connection.
OFED Scripts	Renamed the UP name that appears in mlnx_perf report to “TC”, as the mlnx_perf script counts the packets and calculates the bandwidth on rings that belong to the same Traffic Class (TC).
Physical Memory Allocation	Added support for Physical Address Memory Region (PA-MR) which allows managing physical memory used for posting send and receive requests.
MAD Congestion Control	Added an SA MAD congestion control mechanism that is configurable using sysfs entries.
IB Router	Added the ability to send traffic between two or more subnets.
PeerDirect Async	Mellanox PeerDirect Async™ sub-system gives peer hardware devices, such as GPU cards, and dedicated AS accelerators the ability to take control over HCA in critical path offloading CPU.
Physical MR	Allows the user to use physical addresses instead of virtual addresses in critical path. Thus enhances performance since there is no need in addresses translation.
RoCE v1 (Layer 2) Compatibility	Added the option to connect between nodes running MLNX_OFED and nodes running RoCE with Layer 2 GID format.
3.2-2.0.0.0	
API Changes	<ul style="list-style-type: none"> • Support FCS scattering for Raw Packet QPs and WQs. • Indication of L4 packet type on the receive side completions • Support CVLAN insertion for WQs
IPoIB	<ul style="list-style-type: none"> • Added support for the following IPoIB UD QP offloads: <ul style="list-style-type: none"> • RX check summing (AKA RX csu) • Large Send Offloads (AKA LSO) <p>To see the new IPoIB UD mode, run: "ethtool -k <interface>"</p>
3.2-1.0.1.1	
VXLAN Hardware Stateless Offloads	[ConnectX-4 / ConnectX-4 Lx] Provides scalability and security challenges solutions.
Priority Flow Control (PFC)	[ConnectX-4 / ConnectX-4 Lx] Applies pause functionality to specific classes of traffic on the Ethernet link.
Offloaded Traffic Sniffer/TCP Dump	[ConnectX-4 / ConnectX-4 Lx] Allows bypass kernel traffic (such as, RoCE, VMA, DPDK) to be captured by existing packet analyzer such as tcpdump.
Ethernet Time Stamping	[ConnectX-4 / ConnectX-4 Lx] Keeps track of the creation of a packet. A time-stamping service supports assertions of proof that a datum existed before a particular time.
Custom RoCE Counters	[ConnectX-4 / ConnectX-4 Lx] Provide a clear indication on RDMA send/receive statistics and errors.
LED Beaconing	[ConnectX-4 / ConnectX-4 Lx] Enables visual identification of the port by LED blinking.

Table 9 - Change Log History

Category	Description
Enhanced Transmission Selection Standard (ETS)	[ConnectX-4 / ConnectX-4 Lx] Exploits the time periods in which the offered load of a particular Traffic Class (TC) is less than its minimum allocated bandwidth.
Strided WQE User Space	[ConnectX-4 / ConnectX-4 Lx] Striding RQ is a receive queue comprised by work queue elements (i.e. WQEs), where multiple packets of LRO segments (i.e. message) are written to the same WQE.
VLAN Stripping in Linux Verbs	[ConnectX-4 / ConnectX-4 Lx] Adds access to the device's ability to offload the Customer VLAN (cVLAN) header stripping from an incoming packet.
iSER: Remote invalidation support (target and initiator)	[ConnectX-4 / ConnectX-4 Lx] Improves performance by enabling the hardware to perform implicit memory region invalidation.
iSER: Zero-Copy ImmediateData	[ConnectX-4 / ConnectX-4 Lx] Reduces the latency of small writes by avoiding an extra memory copy in the iSER target stack.
iSER: Indirect Memory Registration	[ConnectX-4 / ConnectX-4 Lx] Uses ConnectX®-4 adapter card's Indirect Memory Registration capabilities to avoid bounce buffer strategy implementation and to reduce the latency of highly unaligned vectored IO operations, and also in cases of BIO merging.
Vector Calculation/Erasure coding offload	[ConnectX-4 / ConnectX-4 Lx] Uses the HCA for offloading erasure coding calculations.
Virtual Guest Tagging (VGT+)	[ConnectX-3 / ConnectX-3 Pro] VGT+ is 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.
Link Aggregation for Virtual Functions	[ConnectX-3 / ConnectX-3 Pro] Protects a VM with an attached ConnectX-3 VF from VF port failure, when VFs are present and RoCE Link Aggregation is configured in the Hypervisor.
3.1-1.0.3	
User Access Region (UAR)	Allows the ConnectX-3 driver to operate on PPC machines without requiring a change to the MMIO area size.
CQE Compression	Saves PCIe bandwidth by compressing a few CQEs into a smaller amount of bytes on PCIe
Bug fixes	See Section 4, "Bug Fixes History" , on page 62
3.1-1.0.0	
Wake-on-LAN (WOL)	Wake-on-LAN (WOL) is a technology that allows a network professional to remotely power on a computer or to wake it up from sleep mode.
Hardware Accelerated 802.1ad VLAN (Q-in-Q Tunneling)	Q-in-Q tunneling allows the user to create a Layer 2 Ethernet connection between two servers. The user can segregate a different VLAN traffic on a link or bundle different VLANs into a single VLAN.
ConnectX-4 ECN	ECN in ConnectX-4 enables end-to-end congestions notifications between two end-points when a congestion occurs, and works over Layer 3.

Table 9 - Change Log History

Category	Description
RSS Verbs Support for ConnectX-4 HCAs	Receive Side Scaling (RSS) technology allows spreading incoming traffic between different receive descriptor queues. Assigning each queue to different CPU cores allows better load balancing of the incoming traffic and improve performance.
Minimal Bandwidth Guarantee (ETS)	The amount of bandwidth (BW) left on the wire may be split among other TCs according to a minimal guarantee policy.
SR-IOV Ethernet	SR-IOV Ethernet at Beta level
3.0-2.0.1	
Virtualization	Added support for SR-IOV for ConnectX-4/Connect-IB adapter cards.
3.0-1.0.1	
HCAs	Added support for ConnectX®-4 Single/Dual-Port Adapter supporting up to 100Gb/s.
RoCE per GID	RoCE per GID provides the ability to use different RoCE versions/modes simultaneously.
RoCE Link Aggregation (RoCE LAG): ConnectX-3/ConnectX-3 Pro only	RoCE Link Aggregation (available in kernel 4.0 only) provides failover and link aggregation capabilities for mlx4 device physical ports. In this mode, only one IB port that represents the two physical ports, is exposed to the application layer.
Resource Domain Experimental Verbs	Resource domain is a verb object which may be associated with QP and/or CQ objects on creation to enhance data-path performance.
Alias GUID Support in InfiniBand	Enables the <code>query_gid</code> verb to return the admin desired value instead of the value that was approved by the SM, to prevent a case where the SM is unreachable or a response is delayed, or if the VF is probed into a VM before their GUID is registered with the SM.
Denial Of Service (DOS) MAD Prevention	Denial Of Service MAD prevention is achieved by assigning a threshold for each agent's RX. Agent's RX threshold provides a protection mechanism to the host memory by limiting the agents' RX with a threshold.
QoS per VF (Rate Limit per VF)	Virtualized QoS per VF, (supported in ConnectX-3/ConnectX-3 Pro adapter cards only with firmware v2.33.5100 and above), limits the chosen VFs' throughput rate limitations (Maximum throughput). The granularity of the rate limitation is 1Mbits.
Ignore Frame Check Sequence (FCS) Errors	Upon receiving packets, the packets go through a checksum validation process for the FCS field. If the validation fails, the received packets are dropped. Using this feature, enables you to choose whether or not to drop the frames in case the FCS is wrong and use the FCS field for other info.
Sockets Direct Protocol (SDP)	Sockets Direct Protocol (SDP) is a byte-stream transport protocol that provides TCP stream semantics. and utilizes InfiniBand's advanced protocol offload capabilities.

Table 9 - Change Log History

Category	Description
Scalable Subnet Administration (SSA)	The Scalable Subnet Administration (SSA) solves Subnet Administrator (SA) scalability problems for Infiniband clusters. It distributes the needed data to perform the path-record-calculation needed for a node to connect to another node, and caches these locally in the compute (client) nodes. SSA ^a requires AF_IB address family support (3.12.28-4 kernel and later).
SR-IOV in ConnectX-3 cards	Changed the Alias GUID support behavior in InfiniBand.
LLR max retransmission rate	Added LLR max retransmission rate as specified in Vendor Specific MAD V1.1, Table 110 - PortLLRStatistics MAD Description ibdiagnet presents the LLR max_retransmission_rate counter as part of the PM_INFO in db_csv file.
Experimental Verbs	Added the following verbs: <ul style="list-style-type: none"> • <code>ibv_exp_create_res_domain</code> • <code>ibv_exp_destroy_res_domain</code> • <code>ibv_exp_query_intf</code> • <code>ibv_exp_release_intf</code> Added the following interface families: <ul style="list-style-type: none"> • <code>ibv_exp_qp_burst_family</code> • <code>ibv_exp_cq_family</code>
2.4-1.0.4	
Bug Fixes	See “Bug Fixes History” on page 62.
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	

Table 9 - Change Log History

Category	Description
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 62.
2.3-2.0.1	
Bug Fixes	See "Bug Fixes History" on page 62.
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 <code>mr_cache</code> subtree from <code>debugfs</code> to <code>mlx5_ib</code> while preserving all its semantics.
InfiniBand Utilities	Updated the <code>ibutils</code> package. Added to the <code>ibdiagnet</code> tool the " <code>ibdiagnet2.mlnx_cntrs</code> " option to enable reading of Mellanox diagnostic counters.
Bug Fixes	See "Bug Fixes History" on page 62.
2.3-1.0.1	
Ethernet	<p>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.</p> <p>MLNX_OFED no longer changes the OS <code>sysctl</code> TCP parameters.</p> <p>Added Explicit Congestion Notification (ECN) support</p> <p>Added Flow Steering: A0 simplified steering support</p> <p>Added RoCE v2 support</p>

Table 9 - Change Log History

Category	Description
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.
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
2.2-1.0.1	
Reset Flow	Reset Flow is not activated by default. It is controlled by the mlx4_core' internal_err_reset' module parameter.

Table 9 - Change Log History

Category	Description
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.
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 num_cpu 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 pf_msix_table_size (multiple of 4 - 1) INI parameter Virtual Functions – the remaining MSI-X vectors are spread equally between all VFs, according to the num_vfs mlx4_core 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 therefor, 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
2.1-1.0.6	
IB Core	Added allocation success verification process to ib_alloc_device.
dapl	dapl is recompiled with no FCA support.
openibd	Added the ability to bring up child interfaces even if the parent's ifcfg file is not configured.
libmlx4	Unmapped the hca_clock_page parameter from mlx4_uninit_context.
scsi_transport_srp	scsi_transport_srp cannot be cleared up when rport reconnecting fails.
mlnxofedinstall	Added support for the following parameters: <ul style="list-style-type: none"> '--umad-dev-na' '--without-<package>'

Table 9 - Change Log History

Category	Description
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 • perfest to v2.0-0.76.gbf9a463 • hcoll to v2.0.472-1 • Openmpi to v1.6.5-440ad47 • dapl to v2.0.40
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 dma_mapping errors.
eIPoIB	Added non-virtual environment support.
2.0-3.0.0	
Operating Systems	Additional OS support: <ul style="list-style-type: none"> • SLES11SP3 • Fedora16, Fedora17
Hardware	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.
Storage	iSER (at GA level)

Table 9 - Change Log History

Category	Description
2.0-2.0.5	
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. SSA is tested on SLES 12 only (x86-64 architecture).

API Change Log History

Table 10 - API Change Log History

Name	Description
3.2-2.0.0.0	

Table 10 - API Change Log History

Name	Description
libibverbs	<ul style="list-style-type: none"> • Support FCS scattering for Raw Packet QPs and WQs. <ul style="list-style-type: none"> • Query: <code>ibv_exp_query_device</code> reports <code>IBV_EXP_DEVICE_SCATTER_FCS</code> when it is supported. • Enablement of this feature is done in the creation: <ol style="list-style-type: none"> 1. For Raw Packet QPs: Set <code>IBV_EXP_QP_CREATE_SCATTER_FCS</code> in <code>exp_create_flags</code>. 2. For WQs: Set <code>IBV_EXP_CREATE_WQ_FLAG_SCATTER_FCS</code> in flags of <code>ibv_exp_wq_init_attr</code>. • Indication of L4 packet type on the receive side completions: <ul style="list-style-type: none"> • Query: <code>ibv_exp_query_device</code> reports <code>IBV_EXP_DEVICE_RX_TCP_UDP_PKT_TYPE</code> when it is supported. <code>ibv_exp_cq_family_flags</code> was extended with two flags <code>IBV_EXP_CQ_RX_TCP_PACKET</code> and <code>IBV_EXP_CQ_RX_UDP_PACKET</code> to support L4 packet type when using <code>poll_length_flags()</code>. • Support CVLAN insertion for WQs: <ul style="list-style-type: none"> • Query: <code>IBV_EXP_RECEIVE_WQ_CVLAN_INSERTION</code> is set in <code>ibv_exp_vlan_offloads</code> when CVLAN insertion is supported. • Enablement: The <code>ibv_exp_qp_burst_family</code> was extended to support CVLAN insertion: <ol style="list-style-type: none"> 1. <code>send_pending_vlan</code>: Put one message in the provider send queue and insert <code>vlan_tci</code> to header. 2. <code>send_pending_inline_vlan</code>: Put one inline message in the provider send queue and insert <code>vlan_tci</code> to header. 3. <code>send_pending_sg_list_vlan</code>: Put one scatter-gather(sg) message in the provider send queue and insert <code>vlan_tci</code> to header.
3.2-1.0.1.1	
libibverbs	<ul style="list-style-type: none"> • [Beta] Added API and primitives for Erasure Coding calculations. <ul style="list-style-type: none"> • Verbs: <ul style="list-style-type: none"> • <code>Ibv_exp_alloc_ec_calc</code> • <code>Ibv_exp_dealloc_ec_calc</code> • <code>Ibv_exp_ec_encode_sync</code> • <code>Ibv_exp_ec_decode_sync</code> • <code>Ibv_exp_ec_encode_async</code> • <code>Ibv_exp_ec_decode_async</code> • <code>Ibv_exp_ec_encode_send</code> <p>For further information, please refer to the manual page of the verbs.</p> • Structs: <ul style="list-style-type: none"> • <code>Ibv_exp_ec_calc</code> • <code>Ibv_exp_ec_mem</code> • <code>Ibv_exp_ec_stripe</code> • <code>Ibv_exp_ec_comp</code> • Added version 1 of the CQ family with support for: <ul style="list-style-type: none"> • Multi-Packet RQ (also called striding RQ) • Cvlan stripping offload • Added enhanced masked-atomic device capability • Added a flag to the create QP/WQ option to enable end of RX message padding

Table 10 - API Change Log History

Name	Description
3.1-1.0.3	
libibverbs	<ul style="list-style-type: none"> • Added <code>ibv_exp_wq_family</code> interface family (Supported only by ConnectX®-4) • Added flag to the QP-burst family to enable Multi-Packet WR • Added return error statuses to the <code>ibv_exp_query_intf</code> to notify that common-flags/family-flags are not supported. • Added <code>ibv_exp_query_gid_attr</code> verb. For further information, please refer to the manual page of the verb.
3.0-1.0.0	
libibverbs	<ul style="list-style-type: none"> • Added the following new APIs: <ul style="list-style-type: none"> • <code>ibv_exp_create_res_domain</code> - create resource domain • <code>ibv_exp_destroy_res_domain</code> - destroy resource domain • <code>ibv_exp_query_intf</code> - query for family of verbs interface for specific QP/CQ • <code>ibv_exp_release_intf</code> - release the queried interface • Updated the following APIs: <ul style="list-style-type: none"> • <code>ibv_exp_create_qp</code> - Add resource-domain to the verb parameters • <code>ibv_exp_create_cq</code> - Add resource-domain to the verb parameters
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>
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>
2.2-1.0.1	

Table 10 - API Change Log History

Name	Description
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>
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>
2.1-1.0.0	

Table 10 - API Change Log History

Name	Description
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>
<p>Verbs Extension API: Verbs extension API defines OFA APIs extension scheme to detect ABI compatibility and enable backward and forward compatibility support.</p>	<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_xrzd</code> • <code>ibv_close_xrzd</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>
2.1-1.0.0	
<p>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.</p>	<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>
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>
2.0-2.0.5	

Table 10 - API Change Log History

Name	Description
Libibverbs - Extended speeds	<ul style="list-style-type: none"> • Missing the <code>ext_active_speed</code> attribute from the struct <code>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 • The following enumerations were renamed: <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>