MLNX_EN for Linux Release Notes

Rev 4.1-1.0.2.0

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

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

Table of Contents

Release Update History

Release	Date	Description
Rev 4.1-1.0.2.0	August 28, 2017	Added RHEL/CentOS 7.4 and SLES12 SP3 as supported operating systems to Table 1, "Supported Platforms and Operating Systems," on page 5.
	July 9, 2017	Initial release of this version.

1 Overview

These are the release notes of MLNX_EN for Linux Driver, Rev 4.1-1.0.2.0 which operates across all Mellanox network adapter solutions supporting the following uplinks to servers:

Uplink/HCAs	Driver Name	Uplink Speed
ConnectX®-3/ ConnectX®-3 Pro	mlx4	• Ethernet: 10GigE, 40GigE and 56GigE[a]
ConnectX®-4	mlx5	• Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, 56GigE[a], and 100GigE
ConnectX®-4 Lx		• Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, and 50GigE
ConnectX®-5		• Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, and 100GigE
ConnectX®-5 Ex		• Ethernet: 1GigE, 10GigE, 25GigE, 40GigE, 50GigE, and 100GigE

a. 56 GbE is a Mellanox propriety 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 Supported Platforms and Operating Systems

The following are the supported OSs in MLNX_EN Rev 4.1-1.0.2.0 :

Table 1 - Supported Platforms and Operating Systems

Operating System	Platform
RHEL6.2/CentOS6.2	x86_64
RHEL6.3/CentOS6.3	x86_64
RHEL6.4/CentOS6.4	x86_64
RHEL6.5/CentOS6.5	x86_64
RHEL6.7/CentOS6.7	x86_64
RHEL6.8/CentOS6.8	x86_64
RHEL6.9/CentOS6.9	x86_64
RHEL7.0/CentOS7.0	x86_64
RHEL7.1/CentOS7.1	x86_64
RHEL7.2/CentOS7.2	x86_64/PPC64 (Power8)/PPC64LE (Power8)
CentOS7.2 with Kernel 4.9 for NVMEoF	ARMv8(AMD) for Softiron
RHEL7.3/CentOS7.3	x86_64/PPC64 (Power8)/PPC64LE (Power8)/ ARMv8 (Qualcomm)
RHEL7.4/CentOS7.4	x86_64/PPC64/PPC64LE (Power8)

Table 1 - Supported Platforms and Operating Systems

Operating System	Platform
Debian 7.8 + Kernel 4.1.35	x86_64
Debian 7.11	x86_64
Debian 8.7	x86_64
Debian 8.7 Kernel 4.1	x86_64
Debian 8.7 Kernel 4.4	x86_64
Fedora 20	x86_64
Fedora 21	x86_64
Fedora 24	x86_64
OL 6.6	x86_64
OL 6.8	x86_64
OL 7.1	x86_64
SLES11 SP2	x86_64
SLES11 SP3	x86_64
SLES11 SP4	x86_64/PPC64 (Power 8)
SLES12 SP1	x86_64
SLES12 SP2	x86_64/PPC64LE (Power 8)
SLES12 SP3	x86_64/PPC64LE (Power8)
Ubuntu 14.04	x86_64
Ubuntu 16.04 with Kernel 4.9 - Bandera for ARM	ARMv8 (Qualcomm) [**beta**]
Ubuntu 16.04	x86_64/PPC64LE (Power8)
Ubuntu 16.10	x86_64/PPC64LE (Power 8)
Ubuntu 17.04	x86_64/PPC64LE (Power 8)
Kernels 4.10-4.11	x86_64
WindRiver 6.0	x86_64 SR-IOV + PV

32 bit platforms are no longer supported in MLNX_EN

1.1.1 Tested Hypervisors in Paravirtualized and SR-IOV Environments

Table 2 - Tested Hypervisors in Paravirtualized and SR-IOV Environments

Tested Hypervisors	HCAa	Operating System
SR-IOV	ConnectX-3/ ConnectX-3 Pro	SLES11 SP3 KVM
		SLES11 SP4 KVM
		SLES12 SP2 KVM
		Ubuntu 14.04 KVM
		Ubuntu 16.04 KVM
		Ubuntu 16.10 KVM
		Ubuntu 16.10 KVM-PPC
		RHEL 6.9 KVM
		RHEL 7.3 KVM
	ConnectX-4	SLES11 SP3 KVM
		SLES11 SP4 KVM
		SLES12 SP2 KVM
		Ubuntu 14.04 KVM
		Ubuntu 16.04 KVM
		Ubuntu 16.10 KVM
		RHEL6.9 KVM
		RHEL7.3 KVM
	ConnectX-4 Lx	SLES11 SP4 KVM
		Ubuntu 16.04 KVM
		Ubuntu 14.04 KVM
	ConnectX-5	Ubuntu16.10 KVM
		RHEL6.9 KVM
		RHEL7.3 KVM
Paravirtualized	ConnectX-3/ ConnectX-3 Pro	SLES12 SP2 KVM
		Ubuntu 16.10 KVM

1.2 Supported HCAs Firmware Versions

MLNX_EN Rev 4.1-1.0.2.0supports the following Mellanox network adapter cards firmware versions:

Table 3 - 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

Table 3 - Supported HCAs Firmware Versions

HCA	Recommended Firmware Rev.	Additional Firmware Rev. Supported
ConnectX®-4	12.20.1010	12.18.2000
ConnectX®-4 Lx	14.20.1010	14.18.2000
ConnectX®-5	16.20.1010	16.19.1200
ConnectX®-5 Ex	16.20.1010	16.19.1200

For the official firmware versions, please see:
http://www.mellanox.com/content/pages.php?pg=firmware_download

2 Changes and New Features in Rev 4.1-1.0.2.0

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

Table 4 - Changes and New Features in Rev 4.1-1.0.2.0

HCAs	Feature/Change	Description
mlx5 Driver	RoCE Diagnostics and ECN Counters	Added support for additional RoCE diagnostics and ECN congestion counters under /sys/class/infiniband/mlx5_0/ports/1/hw_counters/ directory. For further information, refer to the Understanding mlx5 Linux Counters and Status Parameters Community post.
	rx-fcs Offload (ethtool)	Added support for rx-fcs ethtool offload configuration. Normally, the FCS of the packet will be truncated by the ASIC hardware before sending it to the application socket buffer (skb). Ethtool allows to set the rx-fcs not to be truncated, but to pass it to the application for analysis. For more information and usage, refer to Understanding ethtool rx-fcs for mlx5 Drivers Community post.
	DSCP Trust Mode	Added the option to enable PFC based on the DSCP value. Using this solution, VLAN headers will no longer be mandatory for use. For further information, refer to the HowTo Configure Trust Mode on Mellanox Adapters Community post.
	RoCE ECN Parameters	ECN parameters have been moved to the following directory: /sys/kernel/debug/mlx5/<PCI BUS>/cc_params/ For more information, refer to the HowTo Configure DCQCN (RoCE CC) for ConnectX-4 (Linux) Community post.
	Flow Steering Dump Tool	Added support for mlx_fs_dump, which is a python tool that prints the steering rules in a readable manner.
	Secure Firmware Updates	Firmware binaries embedded in MLNX_EN package now support Secure Firmware Updates. This feature provides devices with the ability to verify digital signatures of new firmware binaries, in order to ensure that only officially approved versions are installed on the devices. For further information on this feature, refer to Mellanox Firmware Tools (MFT) User Manual.
	PeerDirect	Added the ability to open a device and create a context while giving PCI peer attributes such as name and ID. For further details, refer to the PeerDirect Programming Community post.

Table 4 - Changes and New Features in Rev 4.1-1.0.2.0

HCAs	Feature/Change	Description
	Probed VFs	Added the ability to disable probed VFs on the hypervisor. For further information, see HowTo Configure and Probe VFs on mlx5 Drivers Community post.
	Local Loopback	Improved performance by rendering Local loopback (unicast and multicast) disabled by mlx5 driver by default while local loopback is not in use. The mlx5 driver keeps track of the number of transport domains that are opened by user-space applications. If there is more than one user-space transport domain open, local loopback will automatically be enabled.
	1PPS Time Synchronization (at **alpha** level)	Added support for One Pulse Per Second (1PPS), which is a time synchronization feature that allows the adapter to send or receive 1 pulse per second on a dedicated pin on the adapter card. For further information on this feature, refer to the HowTo Test 1PPS on Mellanox Adapters Community post.
	Fast Driver Unload	Added support for fast driver teardown in shutdown and kexec flows.
ConnectX-5/ ConnectX-5 Ex	NVMEoF Target Offload	Added support for NVMe over fabrics (NVMEoF) offload, an implementation of the new NVMEoF standard target (server) side in hardware. For further information on NVMEoF Target Offload, refer to HowTo Configure NVMEoF Target Offload.
All	RDMA CM	Changed the default RoCE mode on which RDMA CM runs to RoCEv2 instead of RoCEv1. RDMA_CM session requires both the client and server sides to support the same RoCE mode. Otherwise, the client will fail to connect to the server. For further information, refer to RDMA CM and RoCE Version Defaults Community post.
	Bug Fixes	See Section 4, “Bug Fixes History”, on page 28.

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

2.1 Unsupported Functionalities/Features/HCAs

The following are the unsupported functionalities/features/HCAs in MLNX_EN:

- ConnectX®-2 Adapter Card

3 Known Issues

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

For the list of old known issues, please refer to MLNX_EN Archived Known Issues file at:
http://www.mellanox.com/pdf/prod_software/MLNX_EN_Archived_Known_Issues.pdf

Table 5 - Archived Known Issues

Internal Reference Number	Issue
995665	**Description**: Connection between NVMEoF host and target cannot be established in a hyper-threaded system with more than 64 CPUs on the NVMEoF host side.
	Workaround: On the host side, connect to NVMEoF subsystem using `--nr-io-queues <num_queues>` flag. Note that `num_queues` must be lower or equal to num_sockets multiplied with num_cores_per_socket.
	Keywords: NVMEoF
1039346	**Description**: Enabling multiple namespaces per subsystem while using NVMEoF target offload is not supported.
	Workaround: To enable more than one namespace, create a subsystem for each one.
	Keywords: NVMEoF Target Offload, namespace
1030301	**Description**: Creating virtual functions on a device that is in LAG mode will destroy the LAG configuration. The boding device over the Ethernet NICs will continue to work as expected.
	Workaround: N/A
	Keywords: LAG, SR-IOV
1047616	**Description**: When node GUID of a device is set to zero (0000:0000:0000:0000), RDMA_CM user space application may crash.
	Workaround: Set node GUID to a nonzero value.
	Keywords: RDMA_CM
1051701	**Description**: New versions of iproute which support new kernel features may misbehave on old kernels that do not support these new features.
	Workaround: N/A
	Keywords: iproute

Table 5 - Archived Known Issues

Internal Reference Number	Issue
1007830	**Description**: When working on Xenserver hypervisor with SR-IOV enabled on it, make sure the following instructions are applied: 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.
	Workaround: N/A
	Keywords: SR-IOV

1005786	**Description:** When using ConnectX-5 adapter cards, the following error might be printed to dmesg, indicating temporary lack of DMA pages:
	`"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"`
	
	Example: This might happen when trying to open more than 64 VFs per port.

	Workaround: N/A
	Keywords: mlx5_core, DMA
1008066	**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 `get_vf_stats` (via iptool) during reboot.
	Workaround: N/A
	Keywords: mlx5_core, reboot

1009488	**Description:** Mounting MLNX_EN to a path that contains special characters, such as parenthesis or spaces is not supported. For example, when mounting MLNX_EN to "/media/CDROM(vcd)/", installation will fail and the following error message will be displayed:
	`# cd /media/CDROM\(vcd\)/`
	`# ./install`
	`sh: 1: Syntax error: "(" unexpected`

	Workaround: N/A
	Keywords: Installation
982144	**Description:** When offload traffic sniffer is on, the bandwidth could decrease up to 50%.
	Workaround: N/A
	Keywords: Offload Traffic Sniffer

Table 5 - Archived Known Issues

Internal Reference Number	Issue
982534	**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: 1. mlx4_core driver does not load. 2. The mlx4_core driver does load, but calls to `ibv_open_device` may return ENOMEM errors.
	Workaround: 1. Add the following parameter in the firmware's ini file under [HCA] section: `log2_uar_bar_megabytes = 7` 2. Re-burn the firmware with the new ini file.
	Keywords: PPC
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 /sys/class/net/<interface>/qos/tc_num sysfs file.
	Keywords: Ethernet, TC
979457	**Description:** When setting IOMMU=ON, a severe performance degradation may occur due to a bug in IOMMU.
	Workaround: Make sure the following patches are found in your kernel: • iommu/vt-d: Fix PASID table allocation • iommu/vt-d: Fix IOMMU lookup for SR-IOV Virtual Functions **Note**: These patches are already available in Ubuntu 16.04.02 and 17.04 OSs.
	Keywords: Performance, IOMMU
942161	**Description:** On some kernels, there might be an issue in csum calculations of tunneled packets when the driver sets CHECKSUM_COMPLETE for the packet. This might print csum error messages to the dmesg 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 MLNX_EN v1.5.x.
	Workaround: Turn off the GRO feature to get better UDP performance. Run: `#ethtool -K <interface> gro off`
	Keywords: GRO, UDP, performance

Table 5 - Archived Known Issues

Internal Reference Number	Issue
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 kernel-mft-mlnx might fail with such an error: `"Error: package kernel-mft-mlnx-kmp-default seems to contain modules for multiple kernel versions"`
	Workaround: Use the following command to remove the kernel-mft packages: `rpm -e --noscripts $(rpm -qa | grep kernel-mft-mlnx)`
	Keywords: Installation
918880	**Description:** The driver version shown in modinfo and ethtool outputs is 3.4-1.0.6 instead of 3.4-2.0.0.
	Workaround: N/A
	Keywords: Installation

690799	**Description:** OpenSM package removal fails with the following error on Ubuntu12.04:
	`Removing opensm ...`
	`/sbin/insserv: No such file or directory`

	Workaround: 1. Create the missing link by running this command: `# ln -s /usr/lib/insserv/insserv /sbin/insserv` 2. Remove the package.
	Keywords: Installation
764204	**Description:** Weak Updates (KMP) support is broken on RHEL PPC64LE with errata kernels. MLNX_EN installation will pass, but no links will be created under the weak-updates directory for the new kernel. Therefore, the driver load will fail.
	Workaround: • As of MLNX_EN v3.3, use the mlnx_add_kernel_support.sh script, or simply provide the --add-kernel-support flag to mlnxofedinstall script. • Update the kmod package using the following link: https://rhn.redhat.com/errata/RHBA-2016-1832.html
	Keywords: Installation

Table 5 - Archived Known Issues

Internal Reference Number	Issue
785119	**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 mlxfwreset instead of power cycle will print messages similar to the following when Passing-Through the device to Virtual Machine: `"-device vfio-pci,host=04:00.0,id=host-dev0,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."`
	Workaround: N/A
	Keywords: Installation
677998	**Description:** False alarm errors may be printed to dmesg.
	Workaround: N/A
	Keywords: Driver Start
967356	**Description:** [**Ethernet**] • 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.
	Workaround: N/A
	Keywords: Performance
956071	**Description:** [**mlx5**] OOB TCP performance for small message sizes may suffer from lower BW than expected.
	Workaround: Disable adaptive-rx and set higher static moderation: `ethtool -C <interface> adaptive-rx off rx-frames 128 rx-usecs 128`
	Keywords: Performance
765777	**Description:** Low VxLAN throughput due to broken GRO offload in most kernels older than kernel v4.6.
	Workaround: Use kernel version 4.6 or above.
	Keywords: Performance
414827	**Description:** Out-of-the-box throughput performance in Ubuntu14.04 is not optimal and may achieve results below the line rate in 40GE link speed.
	Workaround: For additional performance tuning, please refer to Performance Tuning Guide.
	Keywords: Performance

Table 5 - Archived Known Issues

Internal Reference Number	Issue
417751	**Description:** Performance degradation might occur when bonding Ethernet interfaces.
	Workaround: N/A
	Keywords: Performance
656415	**Description:** In RHEL7.0, when the irqbalance service is started or restarted, it incorrectly re-balances the IRQs, including the banned ones.
	Workaround: N/A
	Keywords: Performance
651322	**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. For further information, please refer to: http://comments.gmane.org/gmane.linux.network/344243
	Workaround: Use RH7.2 to avoid such performance issues.
	Keywords: Performance
754646	**Description:** The default RX coalescing values yield to high CPU utilization when using VXLAN on VMs over PV.
	Workaround: Increase the RX microseconds and frames coalescing parameters for a better utilization using the ethtool -C command.
	Keywords: Performance
783496	**Description:** When using a VF over RH7.X KVM, low throughput is expected.
	Workaround: Install the following packages using the link below: • 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 http://people.redhat.com/~alwillia/bz1299846/
	Keywords: Performance
860311	**Description:** An allocation of high-order page in mlx5e_alloc_striding_rx-_wqe fails with a call-trace.
	Workaround: No action is required on users end. A fragmented fallback flow will handle this failure.
	Keywords: mlx5 Driver

Table 5 - Archived Known Issues

Internal Reference Number	Issue
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. `command failed, status bad system state(0x4), syndrome 0x408b33` The message can be safely ignored.
	Workaround: Upgrade Connect-IB firmware to the latest available version.
	Keywords: mlx5 Driver
538843	**Description:** Bonding active-backup mode does not function properly.
	Workaround: N/A
	Keywords: mlx5 Driver
-	**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.
	Workaround: N/A
	Keywords: mlx5 Driver
563022	**Description:** ConnectX-4 port GIDs table shows a duplicated RoCE v2 default GID.
	Workaround: N/A
	Keywords: mlx5 Driver
947542	**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.
	Workaround: N/A
	Keywords: Ethernet

Table 5 - Archived Known Issues

Internal Reference Number	Issue
964991	**Description:** TX queue rate limit may sometimes exceed the rate that was set by the user by up to 10,.
	Workaround: N/A
	Keywords: Ethernet
911693	**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.
	Workaround: N/A
	Keywords: Ethernet
948312	**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).
	Workaround: N/A
	Keywords: Ethernet
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_EN User Manual.
	Keywords: Ethernet
754709	**Description:** mlx5 Ethernet auto-negotiation related issues: 1. The command ethtool -s eth4 speed 25000 autoneg on is not a valid ethtool command. speed 25000 should not be passed in when autoneg is on. Instead, use advertise 0x100000. 2. ethtool 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 ethtool 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

Table 5 - Archived Known Issues

Internal Reference Number	Issue
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
433366	**Description:** Reboot might hang in SR-IOV when using the `probe_vf` parameter with many Virtual Functions. The following message is logged in the kernel log: `"waiting for eth to become free. Usage count =1"`
	Workaround: N/A
	Keywords: Ethernet
539117	**Description:** On SLES12, the bonding interface over Mellanox Ethernet slave interfaces does not get IP address after reboot.
	Workaround: 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: `"<use-nanny>false</use-nanny>"` to `"<use-nanny>true</use-nanny>"` 3. Run: `# systemctl restart wickedd.service wicked`
	Keywords: Ethernet
989042	**Description:** `ethtool -x` command will not function on relatively old kernels that do not support get/set_rxfh* callbacks.
	Workaround: N/A
	Keywords: Ethernet
516136	**Description:** Ethertype proto 0x806 not supported by ethtool
	Workaround: N/A
	Keywords: Ethernet
592229	**Description:** When NC-SI is ON, the ports MTU cannot be set to lower than 1500.
	Workaround: N/A
	Keywords: Ethernet

Table 5 - Archived Known Issues

Internal Reference Number	Issue
600242	**Description:** GRO is not functional when using VXLAN in ConnectX-3 adapter cards.
	Workaround: N/A
	Keywords: Ethernet
596075	**Description:** ethtool -X: The driver supports only the 'equal' mode and cannot be set by using weight flags.
	Workaround: N/A
	Keywords: Ethernet
600752	**Description:** Q-in-Q infrastructure in the kernel is supported only in kernel version 3.10 and up.
	Workaround: N/A
	Keywords: Ethernet
596537	**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.
	Workaround: N/A
	Keywords: Ethernet
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
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

Table 5 - Archived Known Issues

Internal Reference Number	Issue
835239	**Description:** While running Q-in-Q packets with stag offloading, tcpsump/ wireshark on host may show svlan ethertype as 0x8100 instead of 0x88A8.
	Workaround: Check the wire or a switch between the hosts, the wireshark will show 0x88A8 ethertype as expected.
	Keywords: Ethernet
954924	**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.
	Workaround: N/A
	Keywords: Flow Steering
516136	**Description:** Setting ARP flow rules through ethtool is not allowed.
	Workaround: N/A
	Keywords: Flow Steering
448981	**Description:** QoS default settings are not returned after configuring QoS.
	Workaround: N/A
	Keywords: Quality of Service
940345	**Description:** In ConnectX-3, when the virtual function (VF) runs on a MLNX_EN version that is below v4.0, and the physical function runs on MLNX_EN 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
683370	**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.
	Workaround: Set module `param always_register=N` `$ modprobe ib_iser always_register=N`
	Keywords: iSER Initiator
960642	**Description:** [**mlx5**] `min_tx_rate` and `max_tx_rate` limit per virtual function is not supported on ConnectX-5 and ConnectX-5 Ex adapter cards.
	Workaround: N/A
	Keywords: SR-IOV

Table 5 - Archived Known Issues

Internal Reference Number	Issue
858628	**Description:** PCI error handling is not supported during driver reload. This might cause a kernel panic or calltrace.
	Workaround: N/A
	Keywords: SR-IOV
860385	**Description:** Creating 127 VFs may cause kernel panic in SLES11 SP4 KVM with Kernel 3.0.101-63 because of a IOMMU kernel bug.
	Workaround: N/A
	Keywords: SR-IOV
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
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 an driver slowdown.
	Workaround: N/A
	Keywords: SR-IOV
784954	**Description:** When SR-IOV is disabled, the VF driver receives `pci_err_-detected` 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 `fail_over_mac=1`.
	Workaround: N/A
	Keywords: SR-IOV

Table 5 - Archived Known Issues

Internal Reference Number	Issue
381764	**Description:** `mlx4_port1_mtu` sysfs entry shows a wrong MTU number in the VM.
	Workaround: N/A
	Keywords: SR-IOV
385750/ 378528	**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.
	Workaround: Unload the module mlx4_ib and reload it in the VM.
	Keywords: SR-IOV
392172	**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
	Workaround: Shut down the driver in the VM before detaching the VF.
	Keywords: SR-IOV
506512	**Description:** Setting 1 Mbit/s rate limit on Virtual Functions (Qos Per VF feature) may cause TX queue transmit timeout.
	Workaround: N/A
	Keywords: SR-IOV
567908	**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.
	Workaround: N/A
	Keywords: SR-IOV
601749	**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.
	Workaround: N/A
	Keywords: SR-IOV
649366	**Description:** Restarting the PF (Hypervisor) driver while Virtual Functions are assigned is not allowed in RH7 and above due to a `vfio-pci` bug.
	Workaround: N/A
	Keywords: SR-IOV

Table 5 - Archived Known Issues

Internal Reference Number	Issue
639046	**Description:** Due to an issue with SR-IOV loopback, prevention "`Duplicate IPv6 detected`" are seen in the VF driver.
	Workaround: N/A
	Keywords: SR-IOV
655410	**Description:** [**ConnectX-4/Connect-IB**] Failed to enable SR-IOV due to errors in PCI or BIOS.
	Workaround: 1. Add `pci=realloc=on` to the grub command line. 2. Add more memory to the server. 3. Upgrade BIOS version.
	Keywords: SR-IOV
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 `port_rcv_data_64` 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

Table 5 - Archived Known Issues

Internal Reference Number	Issue
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 ifconfig up/down.
	Workaround: N/A
	Keywords: SR-IOV
738464	**Description:** In SLES11 SP4, user cannot open all VFs announced in `sriov_totalvfs`. However he can set the num_vfs up to maximum sriov_totalvfs-1 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 dmesg.
	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: `"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".`
	Workaround: Upgrade to the latest version of QEMU in the hypervisor.
	Keywords: SR-IOV

Table 5 - Archived Known Issues

Internal Reference Number	Issue
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
856033	**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
	Workaround: Edit the kernel parameters (in grub) and add qiommu.identity_map_qiommus=PCIE0_MMU,PCIE4_MMU (The bus numbers depend on the ConnectX-4 slot.) Reboot the server.
	Keywords: General
552870/ 548518	**Description:** On rare occasions, under extremely heavy MAD traffic, MAD (Management Datagram) storms might cause soft-lockups in the UMAD layer.
	Workaround: N/A
	Keywords: General
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 `update-pciids`
	Keywords: General
767016	**Description:** Resetting hardware counters after netdev goes up can break statistics scripts.
	Workaround: N/A
	Keywords: General

Table 5 - Archived Known Issues

Internal Reference Number	Issue
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: `sysctl net.ipv4.ip_local_port_range="MIN MAX"` 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)
387061	**Description:** `mlx4_core` can allocate up to 64 MSI-X vectors, an MSI-X vector per CPU.
	Workaround: N/A
	Keywords: Resources Limitation
553657	**Description:** Registering a large amount of Memory Regions (MR) may fail because of DMA mapping issues on RHEL 7.0.
	Workaround: N/A
	Keywords: Resources Limitation

4 Bug Fixes History

This table lists the bugs fixed in this release.

For the list of old bug fixes, please refer to Mellanox OFED Archived Bug Fixes file at:
http://www.mellanox.com/pdf/prod_software/MLNX_EN_Archived_Bug_Fixes.pdf

Table 6 - Bug Fixes History

Internal Ref	Issue
965591	**Description:** Added support for Lustre.
	Keywords: Lustre
	Discovered in Release: 4.0-2.0.0.1
	Fixed in Release: 4.1-1.0.2.0
1038933	**Description:** Fixed a backport issue where IPv6 procedures were called while they were not supported in the underlying kernel.
	Keywords: iw_cm
	Discovered in Release: 4.0-2.0.0.1
	Fixed in Release: 4.1-1.0.2.0
1064722	**Description:** Added log debug prints when changing HW configuration via DCB. To enable log debug prints, run: `ethtool -s <devname> msglvl hw on/off`
	Keywords: DCB, msglvl
	Discovered in Release: 4.0-2.0.0.1
	Fixed in Release: 4.1-1.0.2.0
1022251	**Description:** Fixed SKB memory leak issue that was introduced in kernel 4.11, and added warning messages to the Soft RoCE driver for easy detection of future SKB leaks.
	Keywords: Soft RoCE
	Discovered in Release: 4.0-2.0.0.1
	Fixed in Release: 4.1-1.0.2.0
1044546	**Description:** Fixed the issue where a kernel crash used to occur when RXe device was coupled with a virtual (dummy) device.
	Keywords: Soft RoCE
	Discovered in Release: 4.0-2.0.0.1
	Fixed in Release: 4.1-1.0.2.0
1047617	**Description:** Fixed the issue where a race condition in the RoCE GID cache used to cause for the loss of IP-based GIDs.
	Keywords: RoCE, GID
	Discovered in Release: 4.0-2.0.0.1
	Fixed in Release: 4.1-1.0.2.0

Table 6 - Bug Fixes History

Internal Ref	Issue
1006768	**Description:** Fixed the issue where an rdma_cm connection between a client and a server that were on the same host was not possible when working over VLAN interfaces.
	Keywords: RDMACM
	Discovered in Release: 4.0-2.0.0.1
	Fixed in Release: 4.1-1.0.2.0
801807	**Description:** Fixed an issue where RDMACM connection used to fail upon high connection rate accompanied with the error message: `RDMA_CM_EVENT_UNREACHABLE`.
	Keywords: RDMACM
	Discovered in Release: 3.0-2.0.1
	Fixed in Release: 4.1-1.0.2.0
869768	**Description:** Fixed the issue where SR-IOV was not supported in systems with a page size greater than 16KB.
	Keywords: SR-IOV, mlx5, PPC
	Discovered in Release: 4.0-2.0.0.1
	Fixed in Release: 4.1-1.0.2.0
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_5 Eth Driver
	Discovered in Release: 3.4-2.0.0.0
	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

Table 6 - Bug Fixes History

Internal Ref	Issue
890285	**Description:** Fixed the issue where memory allocation for CQ buffers used to fail when increasing the RX ring size.
	Keywords: mlx5_core
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 4.0-1.0.1.0
867094	**Description:** Fixed the issue where MLNX_EN 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 `biosdavename` running on Redhat 6.x with MLNX_EN may show the same name to ConnectX-3 Eth port 1 and ConnectX-3 Eth port 2.
	Keywords: biosdavename
	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 `mlx4_en_get_phys_port_id` 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 mlx4 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 `prof_sel` was invalid, `mlx5_core` 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

Table 6 - Bug Fixes History

Internal Ref	Issue
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 `get_numa_phys_mask` function.
	Keywords: mlx5_core
	Discovered in Release: 3.4-1.0.0.0
	Fixed in Release: 3.4-2.0.0.0
854344	**Description:** Fixed the issue where `mlnx_affinity` script on RHEL/CentOS7.x host did not disable or enable `irqbalancer`.
	Keywords: irqbalancer
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.03.4-1.0.0.3
824736	**Description:** Fixed wrong skprio2UP mapping by removing it and its scripts, such as tc_wrap, from the driver. This mapping should now be done using the kernel's set_egress_map commands. Note: Only for RDMACM over old kernels, the original skprio2UP mapping in tc_wrap remains valid as these kernels do not support set_egress_map.
	Keywords: QoS
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.03.4-1.0.0.3
826686	**Description:** Fixed the issue where server reboot could get stuck because of kernel panic in mlx4_en_get_drvinfo() that is called from asynchronous event handler.
	Keywords: mlx4_en
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.03.4-1.0.0.3
824130	**Description:** Fixed the issue where ethtool self test used to fail on interrupt test after timeout if mlx4_ib module was not loaded.
	Keywords: mlx4_en
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.03.4-1.0.0.3
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: mlx5 driver
	Discovered in Release: 3.3-1.0.0.0
	Fixed in Release: 3.4-1.0.0.03.4-1.0.0.3

Table 6 - Bug Fixes History

Internal Ref	Issue
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: 3.4-1.0.0.03.4-1.0.0.3
568602	**Description:** Fixed the issue of when repeating change of the mlx5_num_vfs 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: 3.4-1.0.0.03.4-1.0.0.3

5 Change Log History

Table 7 - Change Log History

Category	Description
4.0-2.0.0.1	
PCIe Error Counting	[**ConnectX-4/ConnectX-4 Lx**] Added the ability to expose physical layer statistical counters to ethtool.
Standard ethtool	[**ConnectX-4/ConnectX-4 Lx**] Added support for flow steering and rx-all mode.
SR-IOV Bandwidth Share for Ethernet/ RoCE (**beta**)	[**ConnectX-4/ConnectX-4 Lx]** Added the ability to guarantee the minimum rate of a certain VF in SR-IOV mode.
Adapter Cards	Added support for ConnectX-5 and ConnectX-5 Ex HCAs.
NFS over RDMA (NFSoRDMA)	Removed support for NFSoRDMA drivers. These drivers are no longer provided along with the MLNX_EN package.
3.4-1.0.0.3	
Installation	**[ConnectX®-3/ConnectX®-3 Pro/ConnectX®-4/ConnectX®-4 Lx]** Installation script was renamed from `install.sh` to `install`.
	[ConnectX®-3/ConnectX®-3 Pro/ConnectX®-4/ConnectX®-4 Lx] The package is now shipped with pre-built binary RPMs per OS distribution. By default, the package will install drivers supporting Ethernet only. In addition, the package will include the following new installation options: • Full VMA support which can be installed using the installation option "`--vma`". • Infrastructure to run DPDK using the installation option "`--dpdk`". Notes: • DPDK itself is not included in the package. Users would still need to install DPDK separately after the MLNX_EN installation is completed. • RoCE support can be enabled by installing the VMA package. For further information, please refer to the Installation section in the User Manual.
	The package can be set as a local yum/apt-get repository. Refer to the User Manual for the updated installation instructions.
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.
Package Content	**[ConnectX®-3/ConnectX®-3 Pro]** SR-IOV enabled firmware binaries for ConnectX-3 has been removed from MLNX_EN package (the installation flag "`--enable-sriov`" has been deprecated). To configure SR-IOV, please use the "mlxconfig" or "mstconfig" utilities.

Table 7 - Change Log History

Category	Description
Enhanced PCIe Error Recovery	[**ConnectX®-4/ConnectX®-4 Lx]** Enhanced PCIe error recovery by adding the following behaviors to the flow: • 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)	[**ConnectX®-4/ConnectX®-4 Lx]** Added the ability to rate-limit traffic per Virtual Function in SR-IOV mode.
Dynamically tuned Interrupt Moderation (DIM)	[**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.
Dump Configuration	[**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.
Ethernet Counters	[**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.
3.3-1.0.0.0	
VF MAC Address Anti-Spoofing	[**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.
VF All-multi Mode	[**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.
VF Promiscuous Mode	[**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.
Privileged VF	[**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.
DCBX	[**ConnectX-4/ConnectX-4 Lx**] Added support for standard DCBX CEE API.
Per Priority Counters	[**ConnectX-4/ConnectX-4 Lx**] Exposed performance counters per priority.

Table 7 - Change Log History

Category	Description
Accelerated Receive Flow Steering (aRFS)	[**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.
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).
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.
LED Beaconing	[ConnectX-4 / ConnectX-4 Lx] Enables visual identification of the port by LED blinking.
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.
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.
3.1-1.0.4	
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.
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
3.0-1.0.1	
NICs	Added support for ConnectX®-4 Single/Dual-Port Adapter supporting up to 100Gb/s.

Table 7 - Change Log History

Category	Description
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.
Ethtool	Updated ethtool to incorporate ConnectX®-4 adapter card functionalities.
2.3-2.0.1	
Bug Fixes	See “Bug Fixes History” on page 28.
Reset Flow	Added support for Enhanced Error Handling for PCI (EEH), a recovery strategy for I/O errors that occur on the PCI bus.
2.3-1.0.0	
Ethernet	Added support for arbitrary UDP port for VXLAN. From upstream 3.15-rc1 and onward, it is possible to use arbitrary UDP port for VXLAN. This feature requires firmware version 2.32.5100 or higher. Additionally, the following kernel configuration option `CONFIG_MLX4_EN_VXLAN=y` must be enabled.
	MLNX_EN no longer changes the OS sysctl TCP parameters.
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.
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
Performance	Out of the box performance improvements: • 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
2.0-3.0.0	
Operating Systems	Additional OS support: • SLES11SP3 • Fedora16, Fedora17
Hardware	Added ConnectX®-3 Pro support
1.5.10	
General	See Section 4, “Bug Fixes History”, on page 28.
1.5.9	
Operating Systems	Added support for kernel.org 3.5

Table 7 - Change Log History

Category	Description
Performance	Improved latency by optimizing RX repost mechanism
1.5.8.3	
Operating Systems	Added support for RHEL6.3
1.5.8.2	
Operating Systems	Added support for new kernels: 3.1, 3.2, 3.3
1.5.8.2	
Performance	Moved to interrupt mode to handle TX completions
	Added IRQ affinity control scripts (please see README file for more details)
	Optimized Numa aware memory allocations
	Optimized interrupt usage for TX/RX completions
Installation	Added KMP compliant installation process
Linux Tools	Added support for Ethtool
1.5.7.2	
Operating Systems	Added support for new OS's:
	RHEL6.2
	RHEL5.8
	SLES11SP2
Performance	Added recording RX queue for GRO packets
	Added the usage of Toeplitz hash function for RSS calculation
Reports/Statistics	Enabled RXHASH report on supported systems
1.5.7	
Operating Systems	Added support for new OS's:
	RHEL6.1
	RHEL5.5
	RHEL5.7
	kernel.org (2.6.37, 2.6.38, 2.6.39, 3.0)
	RHEL6.1 KVM

Table 7 - Change Log History

Category	Description
Performance	Improved performance on PPC systems (Using GRO where LRO is not efficient)
	Added IPv6 support to LRO
	Incremented number of TX and RX queues
	Enabled NAPI usage at any given time
	Enabled TX completions spread among multiple MSI-X vectors
	Improved small packets packet rate
	Added 40GigE support (including Ethtool report)
	Added NUMA support
	Added general performance improvements
1.5.6	
Operating Systems	Added support for new OS's:
	RHEL6.0
	RHEL5.6
	SLES11SP1
	kernel.org (2.6.35, 2.6.36)
Performance	Added blue flame support for kernels > 2.6.28 (improves TX latency by 0.4 usec)
	Added RX acceleration feature that supports recvmsg and recvmmsg system calls. See MLNX_EN_Linux_README for further details.
	Added option to use interrupts for TX completion (polling is the default)
	Added option to disable NAPI (enabled by default)
	Added support for control number of RX rings from module parameter
	Added interrupt vector per each RX ring. See /proc/interrupts
	Adaptive moderation improvements
	Added system tuning option to achieve better performance (idle loop polling)
Linux Tools	Added hardware revision report via Ethtool
Multicast Filtering	Added exact match multicast filtering
Driver Load	Link is brought up upon driver load
1.5.1.3	
Operating Systems	Added support for new OS's:
	RHEL5.5
	kernel.org (2.6.16 - 2.6.32)
Performance	Added UDP RSS support (on ConnectX-2 HW only)
	Improved VLAN tagging performance
Linux Tools	Ethtool -e support