Changes and New Features History

Table of contents

Customer Affecting Changes

This section includes history of changes and new feature of three major (GA) releases back. For older versions' history, please refer to their dedicated release notes.

Supported Cards	Description
All HCAs	Supported in the following adapter cards **unless specifically stated otherwise:** ConnectX-4 / ConnectX -4 Lx / ConnectX-5 / ConnectX-6 / ConnectX-6 Dx / ConnectX-6 Lx / ConnectX-7 / BlueField-2
ConnectX-6 Dx and above	Supported in the following adapter cards **unless specifically stated otherwise:** ConnectX-6 Dx / ConnectX-6 Lx / ConnectX-7 / BlueField-2
ConnectX-6 and above	Supported in the following adapter cards **unless specifically stated otherwise:** ConnectX-6 / ConnectX-6 Dx / ConnectX-6 Lx / ConnectX-7 / BlueField-2
ConnectX-5 and above	Supported in the following adapter cards **unless specifically stated otherwise:** ConnectX-5 / ConnectX-6 / ConnectX-6 Dx / ConnectX-6 Lx / ConnectX-7 / BlueField-2
ConnectX-4 and above	Supported in the following adapter cards **unless specifically stated otherwise:** ConnectX-4 / ConnectX -4 Lx / ConnectX-5 / ConnectX-6 / ConnectX-6 Dx / ConnectX-6 Lx / ConnectX-7 / BlueField-2

5.9-0.5.6.0	
ASAP^2 Features	
Linux Bridge VLAN Filtering of 802.1 Q Packets	[ConnectX-6 Dx] Extended mlx5 Linux bridge VLAN offload to support packets tagged with 802.1 Q VLAN ethertype.
Offloading sFlow Sampling Rules	[ConnectX-5 and above] Added support for sFlow sampling rules offloads. sFlow is an industry standard technology for monitoring high speed switched networks. Open vSwitch integrated sFlow to extend

5.9-0.5.6.0	
	the visibility into virtual servers, ensuring data center visibility and control.
Core Features	
Configuring Shared Buffer Size	[ConnectX-6 Dx and above] Enabled user to control shared buffer size and configuration, implicitly. As with each port buffer command the user triggers, the shared buffer configuration will be updated accordingly by the driver.
Control SF Class	[All HCAs] Added support for Control SF Class. Each PCI, PF, VF, SF function, by default, has netdevice, RDMA, and vdpa-net devices always enabled. This feature enables the user to control which device functionality to enable/disable. **Note:** Requires kernel 5.18 or higher.
NetDev Features	
Support RSS over XSK Queues	[All HCAs] Use default RSS functionality to spread traffic across different XSK queues instead of having to provide explicit steering rules.
TLS TIS Pool	[TLS-Enabled Devices] Per-connection hardware TIS objects is used to maintain the device TLS TX context. Use a SW TIS pool for recycling the TIS objects instead of destroying/creating them. This reduces the interaction with the device via the FW command interface, which increases the TLS connection rate.
RDMA Features	
Expand Rep Counters	[ConnectX-5 and above] Adding RDMA traffic-only counters for rep devices. These counters can now be read from host with ethtool or from sysfs and not only from the cointainer.
UMR QP Recilency	[ConnectX-5 and above] Added a recovery flow for the driver's UMR logic so that other UMR requests can be proccessed after the error UMR was dropped and the UMR QP was reset. Previously, a faulty UMR request would have moved the QP to error state and disable any option to continue issuing UMRs.
General	Bug fixes

Feature/Change	Description
5.8- 1.1.2.1	
General	Bug fixes
5.8- 1.0.1.1	
Remove Dependency Between SR-IOV and eSwitch Mode	[All HCAs] Removed dependency between SR-IOV and eSwitch mode. Currently, there are three eSwitch modes: none, legacy, and switchdev (non of which are the default mode). When disabling SR-IOV, the current eSwitch mode will be changed to none. This feature removes eSwitch mode none and also removes dependency between SR-IOV and eSwitch mode.
DevLink Parallel Command	[All HCAs] Added support for running DevLink commands in parallel on different DevLink devices is possible. For example, burning firmware on a few cards on the same host in parallel using DevLink API is now possible.
Graceful Shutdown of Parent and Page Supplier	[All HCAs] Set default graceful period values for functions based on their type. ECPFs will get graceful period of 3 minutes, PFs get 1 minute, and VFs/SFs get 30 seconds.
N Pulses Per Second (NPPS)	[ConnectX-6 Dx and above] Enhanced NPPS to allow setting a pulse period higher than 1 pulse per second and to allow setting the pulse width. If the width is unset, the driver implicitly sets it to half the given period (the width should be less than the pulse period). In this release, the pulse duration ranges between 65536 NS–524288 NS.
Remote Invalidate Option for MKeys	[All HCAs] Addded support for the option to enable remote invalidation when creating a new mkey. This way the rkey for a memory region can be changed frequently.
GPUDirect Over DMA-BUF	[All HCAs] Added support for GPUDirect support over dma-buf. As such, using the new mechanism nv_peer_mem is no longer required. The following is required for dma-buf support: • Linux kernel version 5.12 or later • OpenRM version 515 or later

Feature/Change	Description
	Perftest support was added as well: Default option in perftest is without dmabuf. To run with this option, add --use_cuda_dmabuf in addition to use_cuda flag.
General	Bug fixes

Feature/ Change	Description
5.7-1.0.2.0	
Support Representor Metering Over SFs	[ConnectX-6 Dx and above and BlueField-2] Extended the support of representor metering from supporting only VFs representor to also supporting SFs representor.
Exposing Error Counters on a VPort Manager	[ConnectX-4 and above] Added support for exposing error counters on a VPort manager function for all other VPorts. These counters can be used to detect malicious users who are exploiting flows that can slow the device. The counters are exposed through debugfs under: /sys/kernel/debug/mlx5/esw/<func>/vnic_diag/
Memory Consumption Minimization	[ConnectX-4 and above] Added support for providing knobs which enable users to minimize memory consumption of mlx5 functions (PF/VF/SF).
XDP Support for Uplink Representors	[ConnectX-5 and ConnectX-6 Dx) Added XDP support for uplink representors in switchdev mode.
Resiliency to tx_port_ts	[ConnectX-6 Dx and above] Added resiliency to the tx_port_ts feature. private-flag may be enabled via ethtool tx_port_ts which provides a more accurate time-stamp. In very rare cases, the said time-stamp was lost,

Feature/ Change	Description
	leading to losing the synchronization altogether. This feature allows for fast recovery and allows to quickly regain synchronization.
Database of Devlink Health Asserts	[ConnectX-4 and above] Health buffer now contains more debug information like the epoch time in sec of the error and the error's severity. The print to dmesg is done with the debug level corresponding to the error's severity. This allows the user to use dmesg attribute: dmesg --level to focus on different severity levels of firmware errors.
Expose FEC Counters via Ethtool	[ConnectX-5 and above] Exposed the following FEC (forward error detection) counters: ETHTOOL_A_FEC_STAT_CORRECTED • fc_fec_corrected_blocks_laneX • rs_fec_corrected_blocks ETHTOOL_A_FEC_STAT_UNCORR • fc_fec_uncorrectable_blocks_laneX • rs_fec_uncorrectable_blocks ETHTOOL_A_FEC_STAT_CORR_BITS • phy_corrected_bits Command: ethtool -I show-fec <ifc>
Application Device Queues	[ConnectX-4 and above] Added driver-level support for Application Device Queues. This feature allows partition defining over the RX/TX queues into groups and isolates traffic of different applications. This mainly improves predictability and tail latency.
Reinjection of Packets Into Kernel	[All HCAs] Added support for a new software steering action, mlx5dv_dr_action_create_dest_root_table(). This action can be used to forward packets back into a level 0 table. As a table with level 0 is the kernel owned table, this will result in injecting packets to the kernel steering pipeline.
DCT LAG	[ConnectX-6 Dx and above] Added firmware support to allow explicit port selection based on steering and not QP affinity. Functionality: 1. Use LAG Hash Mode for the HCA with two ports, if supported.

Feature/ Change	Description
	2. Keep port affinity function in LAG Hash Mode if it supports bypass select flow table in non-SwitchDev mode.
AES-XTS in RDMA	Added support for plaintext AES-XTS DEKs.
General	Bug fixes

Customer Affecting Changes

Feature/Change	Description
23.10-1.1.9.0	
Lightweight Local SFs	Following the addition of the Lightweight Local SFs feature in version 23.07, in order to configure the scalable-functions, follow the revised instructions as detailed in the Step-by-Step Guide. **Note:** "Step 2.9 - Set all SF specific device parameters" is now mandatory for local SFs.

23.10-0.5.5.0	
Customer Affecting Change	Description
Debugfs Directory Path Change	The debugfs directory of each interface can now be found under: /sys/kernel/debug/mlx5//, and not directly under the root of the debugfs filesystem (/sys/debug/kernel).
Deprecation of OFED Public Power PC Installation	Starting from this release, MLNX_OFED releases for Power PC are no longer available for download from the public Download Center web page. Instead, you can find it on the following page: https://network.nvidia.com/support/firmware/ibm-systemp/.
Pre-notification: Deprecation of Older Operating Systems	Starting from next release, MLNX_OFED releases will no longer support operating systems with kernels below v4.18. This includes the following systems: • Rhel7.x • Debian9.13

23.10-0.5.5.0	
	• Sles12.x • Xenserver7.1
Customer Affecting Change	Description
23.07-0.5.1.2	
Creating a QKEY with an MSB Set	To allow non-privileged users to create a QKEY with an MSB set, a new module parameter was added. For details, please see "QKEY Mitigation in the Kernel" under New Features.
23.07-0.5.0.0	
IRQ Naming	IRQ renaming is no longer done when bringing the interface up/down. The IRQ name is now constant and is not affected by the interface state.
RPM Packages Verification Key	RPM_GPG-KEY-Mellanox (or its variants) is no longer the public key that verifies RPM packages of MLNX_OFED. Instead, the RPM_GPG-KEY-Mellanox file on the top-level directory of the ISO should be used.
Hairpin sysfs Support	Hairpin sysfs support was restricted to physical and virtual functions only.
mlx5_core node_guid Module Parameter	Removed a non-functional mlx5_core node_guid module parameter.
OpenSM Init	Starting from this release, the opensm init service moves from init.d (/etc/init.d/opensmd start) to systemd (# service opensmd start).
Apt Signing Key	Starting from this release, the public key that signed the apt repository of MLNX_OFED is included in the ISO in a format that can be used directly by the apt for repository signatures verification.
IPoIB ULP Mode Deprecation	Starting from this release, MLNX_OFED supports IPoIB enhanced mode only. The ability to switch back to ULP mode using ipoib_enhanced module parameter is not supported. For more information about the enhanced mode, please refer to the OFED user manual, example: Enhanced IP over InfiniBand.

23.10-0.5.5.0	
Pre-notification: Deprecation of OFED Public Power PC Installation	Starting from next release, MLNX_OFED releases for Power PC will no longer be available for download from the public Download Center web page.

Customer Affecting Change	Description
23.07-0.5.0.0	
Creating a QKEY with an MSB Set	To allow non-privileged users to create a QKEY with an MSB set, a new module parameter was added. For details, please see “QKEY Mitigation in the Kernel“ under New Features.
IRQ Naming	IRQ renaming is no longer done when bringing the interface up/down. The IRQ name is now constant and is not affected by the interface state.
RPM Packages Verification Key	RPM_GPG-KEY-Mellanox (or its variants) is no longer the public key that verifies RPM packages of MLNX_OFED. Instead, the RPM_GPG-KEY-Mellanox file on the top-level directory of the ISO should be used.
Hairpin sysfs Support	Hairpin sysfs support was restricted to physical and virtual functions only.
mlx5_core node_guid Module Parameter	Removed a non-functional mlx5_core node_guid module parameter.
OpenSM Init	Starting from this release, the opensm init service moves from init.d (/etc/init.d/opensmd start) to systemd (# service opensmd start).
Apt Signing Key	Starting from this release, the public key that signed the apt repository of MLNX_OFED is included in the ISO in a format that can be used directly by the apt for repository signatures verification.
IPoIB ULP Mode Deprecation	Starting from this release, MLNX_OFED supports IPoIB enhanced mode only. The ability to switch back to ULP mode using ipoib_enhanced module parameter is not supported. For more information about the enhanced mode, please refer to the OFED user manual, example: Enhanced IP over InfiniBand.

Customer Affecting Change	Description
Pre-notification: Deprecation of OFED Public Power PC Installation	Starting from next release, MLNX_OFED releases for Power PC will no longer be available for download from the public Download Center web page.

Customer Affecting Change	Description
23.04-0.5.3.3	
Netdev Interface Configuration is not Preserved During Reload/Reset/Recovery	As of OFED 23.04, during reset/reload/recovery flows, the netdev interface is destroyed and re-created (rather than just suspended). As a result, the netdev interface configuration is not preserved, and must be re-applied. The way to do this is to use proper network-scripts and/or udev rules files to configure network interface parameters. These are automatically triggered whenever a netdev interface is added, regardless of whether it was added due to a user-initiated operation or an automatic failure recovery operation. Thus, no special processing is required to re-apply the network interface configuration parameters following a reset/reload/recovery operation – it is performed automatically.
Prenotification: Deprecation of OFED Public Power PC Installation	Starting next release, MLNX_OFED releases for Power PC will no longer be available for download from the public Download Center web page.
Prenotification: ULP Mode Deprecation	Starting next release, MLNX_OFED will support IPoIB enhanced mode only. The ability to switch back to ULP mode using ipoib_enhanced module param will not be supported. For more information about the enhanced mode, please refer to OFED user manual, example: Enhanced IP over InfiniBand
Installation, ISO, RedHat	In order to address RHEL kernel symbol changes, ISO images for the following operating systems are built with the updated kernel versions as follows:

Customer Affecting Change	Description
	OS Name Old Kernel New Kernel rhel8.6-aarch64 4.18.0-372.9.1.el8_6.aarch64 4.18.0-372.41.1.el8_6.aarch64 rhel8.6-ppc64le 4.18.0-372.0.1.el8_6.ppc64le 4.18.0-372.41.1.el8_6.ppc64le rhel8.6-x86_64 4.18.0-372.9.1.el8_6.x86_64 4.18.0-372.41.1.el8_6.x86_64 rhel8.7-aarch64 4.18.0-425.3.1.el8.aarch64 4.18.0-425.14.1.el8_7.aarch64 rhel8.7-ppc64le 4.18.0-425.3.1.el8.ppc64le 4.18.0-425.14.1.el8_7.ppc64le rhel8.7-x86_64 4.18.0-425.3.1.el8.x86_64 4.18.0-425.14.1.el8_7.x86_64 rhel9.0-aarch64 5.14.0-70.13.1.el9_0.aarch64 5.14.0-70.46.1.el9_0.aarch64 rhel9.0-ppc64le 5.14.0-70.13.1.el9_0.ppc64le 5.14.0-70.46.1.el9_0.ppc64le rhel9.0-x86_64 5.14.0-70.13.1.el9_0.x86_64 5.14.0-70.46.1.el9_0.x86_64 rhel9.1-aarch64 5.14.0-162.6.1.el9_1.aarch64 5.14.0-162.19.1.el9_1.aarch64 rhel9.1-ppc64le 5.14.0-162.6.1.el9_1.ppc64le 5.14.0-162.19.1.el9_1.ppc64le rhel9.1-x86_64 5.14.0-162.6.1.el9_1.x86_64 5.14.0-162.19.1.el9_1.x86_64 This change comes to support RedHat updated kernels without the need to add --add-kernel-support during OFED installation.
Power Setups on UCX/HPC-X	UCX/HPC-X no longer supports Power setups.
NEO-Host	Starting from this release, OFED will discontinue the provision of NEO-Host.

Customer Affecting Change	Description
	NEO-Host can be manually downloaded and installed using the following guide: https://docs.nvidia.com/networking/display/NEOSDKv26/Installation+and+Initial+Configuration#InstallationandInitialConfiguration-DownloadingtheMellanoxNEOSDKSoftware
dapl	Starting from this release, OFED will discontinue the provision of dapl.
Signing Key for SLES15 sp4 and sp5	As of version 23.04, the builds for SLES15 sp4 and sp5 are being singed with a newer signing key. The corresponding public key can be downloaded from https://www.mellanox.com/downloads/ofed/nv_nbu_kernel_signing_key_pub.der instead of https://www.mellanox.com/downloads/ofed/mlnx_signing_key_pub.der.
dump_pr SM Plugin	Starting from this release, OFED will discontinue the provision of dump_pr subnet manager plugin.
mpi-selector	Starting from this release, OFED will discontinue the provision of mpi-selector.
OpenSM Init	Starting 23.07 release, opensm init service will move from init.d to systemd.

Customer Affecting Change	Description
5.9-0.5.6.0	
Deprecation, LAG Mode via Sysfs	Setting LAG mode via Sysfs is going to be deprecated in a future release. Instead, LAG Hash mode will be used by default, similar to upstream behavior.

Customer Affecting Change	Description
LAG Configuration, PCI Error	From version 5.9, LAG configuration will be lost in case driver incurs a PCI error. Make sure to reconfigure the bond after driver completes the recovery from the PCI error. In releases prior to 5.9, in case of PCI error (EEH injections on PPC setup), the driver recovers LAG bond and reconfigures it automatically in case it what configured before the appearance of the error.

Customer Affecting Change	Description
5.7-1.0.2.0	
Multi-Block Encryption	Multi-block encryption is currently unsupported, due to a hardware limitation.

Feature/Change	Description
5.6-2.0.9.0	
Operating Systems	Added support for the following Operating Systems: RHEL8.6, RHEL9.0, SLES15-SP4.
General	Bug fixes

Feature/Change	Description
5.6-1.0.3.5	
General	
New Adapter Card Support	Added support for ConnectX-7 adapter cards. ConnectX-7 has the same feature set as the ConnectX-6 adapter card.
ASAP^2 Features	
Bridge Spoof Check	[All HCAs] Added support for spoof check with TC flower rules on representors attached to bridge to mirror spoof check SR-IOV functionality.
Setting VF Group Rate Limit	[ConnectX-5 and above] Added support for setting VF group rate limit using Devlink command.
TC Flows on Shared Block	[ConnectX-5 and above] Added support for creation of TC flows on shared block of VF representors.
Flow Metering	[ConnectX-6 Dx and above] Added support for offloading OpenFlow Meters in OVS-DPDK. Please note the following: • Meter offload can be applied only on port 0 and it's VFs • Only one meter per flow is allowed • Only one meter band per meter is allowed • Only meter band type drop is supported • Meter-stats might not be accurate
Core Features	
Firmware Reset	[BlueField-2] Added support of firmware reset in DPU NIC mode.
Increased Robustness of mlx5_core Driver Recovery	[All HCAs] Increased the firmware pre-initialization timeout from 2 minutes to 2 hours when waiting for firmware during driver health recovery, allowing the driver to passively recover from a firmware reset, even if the reset takes an unusually long time. Additionally, added an exit clause to the wait for firmware loop, allowing immediate response to a user initiated device removal.
NetDev Features	

Feature/Change	Description
Ethtool CQE Mode Control	[ConnectX-4 and above] Replaced the vendor-specific Ethtool API (priv-flag) with a standard Ethtool API (replaced 'ethtool --set-priv-flags ethX rx_cqe_moder on/off tx_cqe_moder on/off' with 'ethtool -C ethX cqe-mode-rx on/off cqe-mode-tx on/off'). This decreases the amount of vendor-specific configurations and aligns mlx5 driver with the upstream Ethtool API.
SyncE	[ConnectX-6 Dx] Added an indication in SyncE Daemon that states whether SyncE engine moved to holdover state due to failure (the reason for failure will be displayed). In addition, added indication whether SyncE engines collected enough frequency samples in order to move to holdover.
Security	
OVS-IPSec Full Offload	[BlueField-2] Added support for configuration of IPsec full offload using OVS by adding VXLAN tunnel to OVS with the PSK option.
Installation	
Installation	New options were added to the ofed_uninstall.sh script: --only-kernel and --only-user. Those can be used to uninstall only kernel packages or only user-space packages (the equivalent of kernel-only install or user-space-only install, respectively). This may be useful to keep different sets of kernel and user-space installations. When running the uninstall script with a combination of --only-kernel and --only-user produced an undefined result.

Feature/Change	Description
5.5-1.0.3.2	
ASAP^2 Features	
Bridge Offloads with VLAN	[ConnectX-4 and above] Added support for bridge offloads with VLAN support that works on top of mlx5 representors in switchdev mode.
Supporting OVS Groups in Fast-Failover Mode	[ConnectX-6 Dx] Improved OVS failover through support for OVS groups in fast-failover mode + VF_LAG configuration with OVS.

Feature/Change	Description
Exposing Hairpin Queues Information	[ConnectX-6 Dx and BlueField-2] Added support for exposing hairpin out of buffer drop counter per device. This feature shows buffer drops related only to hairpin queues which were opened on the queried device. To enable this counting mode (this must be done before any hairpin rules are created), use the following: echo "on <peer_devname>" > /sys/class/net/<dev>/hp_oob_cnt_mode where <peer_devname> is the peer device to which traffic coming to the configured device will be forwarded to for transmission. To read the drop counter, use the following: cat /sys/class/net/<dev>/hp_oob_cnt
Linux Bridge Offload	[ConnectX-6 Dx and BlueField-2] Added bridge offloads to support bonding (VF LAG), attaching bond device to bridge instead of uplink representors.
VLAN Pop/Push	[ConnectX-6 Dx] Added OOB support for VLAN push on Rx (wire to VF) and VLAN pop on Tx (wire to VF) in switchdev mode.
Offload Forwarding to Multiple Destinations	[ConnectX-5 and above] Added support for offloading packet replication to up to 32 destination through the use of TC rule.

Slow Path Metering	[ConnectX-4 and above] Expanding the RDMA statistic tool to support setting vendor-specific optional counters dynamically using netlink. Added to mlx5_ib the following optional counters: cc_rx_ce_pkts,cc_rx_cnp_pkts,cc_tx_cnp_pkts.
	Example:
	$ rdma statistic mode supported link rocep8s0f0/1
	link rocep8s0f0/1 supported optional-counters cc_rx_ce_pkts,cc_rx_cnp_pkts,cc_tx_cnp_pkts
	$ sudo rdma statistic set link rocep8s0f0/1 optional-counters cc_rx_ce_pkts,cc_rx_cnp_pkts
	$ rdma statistic mode link rocep8s0f0/1
	link rocep8s0f0/1 optional-counters cc_rx_ce_pkts,cc_rx_cnp_pkts
	$ sudo rdma statistic set link rocep8s0f0/1 optional-counters cc_rx_ce_pkts
	$ rdma statistic mode link rocep8s0f0/1
	link rocep8s0f0/1 optional-counters cc_rx_ce_pkts
	$ sudo rdma statistic unset link rocep8s0f0/1 optional-counters

Feature/Change	Description
Core Features	
Subfunction Trust Configuration Enhancement	[ConnectX-5 and above] Added support via mlxdevm to mark a given PCI subfunction (SF) or virtual function (VF) as a trusted function. The device/firmware decides how to define privileges and access to resources.
Prevent VF Memory Exhaustion	[All] Added support for preventing VF memory exhaustion. This feature exposes a sysfs (to the system admin) which can set a limit on each VF memory consumption. Note: Currently only supported on Ethernet.
BlueField NIC Separate Reset	[BlueField-2] Added support for resetting the NIC domain of BlueField-2 while keeping ARM alive.
Multiple Steering Priorities for FDB Rules	[ConnectX-6 Dx and BlueField-2] Added support in multiple flow steering priorities for FDB rules.
NetDev Features	
Traffic Engineering: Hierarchical QoS	[ConnectX-5 and above] Added support for offloading the HTB qdisc to the NIC, allowing it to scale better by eliminating a single locking point. The configuration is done with the TC commands. Note: Kernel 5.15 or higher is required. Limited to 256 nodes.
TLS RX Resynchronization Resiliency Feature Description	[ConnectX-6 Dx and above] Added support for driver resiliency against high load of RX resync operations.
Simultaneous PTP and CQE Compression	Added support for the activation of PTP and CQE compression simultaneously. Since CQE compression might harm the accuracy of the PTP, the feature enables PTP packets to be moved to a dedicated queue where they are not subjected to compression. However, this configuration conflicts with setting aRFS. Turning off CQE compression, causes a hiccup in traffic which may cause a loss of synchronization. To overcome this, restart the synchronization. Note: This combination is supported only for Ethernet drivers. Other driver profiles, like IPoIB and representors, do not support this combination.

Feature/Change	Description
Installation Features	
Multiple Development Headers Packages	Allowed installing multiple mlnx-ofa_kernel development headers packages (for different kernel versions of the same mlnx-ofa_kernel package version) side by side on the same system.
Kernel Module Signature	Added signature of kernel modules of EulerOS 2.0 SP8-SP10 (x86_64 and aarch64) builds of MLNX_OFED.
Enable sf-cfg-drv by Default in EulerOS2.0	Enabled SF_CFG (SF config dummy driver, --with-sf-cfg-drv) on EulerOS2.0 SP8 and SP10.

Feature/Change	Description
5.4-1.0.3.0	
ASAP^2	
Enlarge Switchdev Tables	[ConnectX-5 and above] Added support for allowing OVS kernel to support up to 128 matches (groups) per table and 16M entries per group.
Offloading Extended ct_state Flags	[ConnectX-5 and above] Added support to offload ct_state flags rpl, inv, and rel. • For rpl, support was added for both set and not set matching offload (i.e., +rpl and -rpl). • For inv and rel, support was added only for the not set option (i.e., -rel and -inv).
Core	
Auxiliary Bus in mlx5 Driver	[ConnectX-4] Updated mlx5 driver to use auxiliary bus in order to integrate different driver components into driver core and optimize module load/unload sequences.
Installation	
Script Removal	[General] Moved all Python scripts and some other common scripts out of the mlnx-ofa_kernel packages. This removed the python dependency

Feature/Change	Description
from mlnx-ofa_kernel	from that package when rebuilding it and avoided unnecessary errors when rebuilding them for custom kernels.
Netdev	
What-Just-Happened (WJH) in NICs	[ConnectX-4] Added support for WJH in NICs. WJH allows for visibility of dropped packets (i.e., receiving notice of drop counters increase, seeing content of the dropped packets, debugging, and more). WJH is a service in devlink context and it is already implemented in the switch. **Note:** processing dropped packets (even for visibility purposes) may cause a degradation in performance and leaves the driver vulnerable for malicious attacks. The feature is disabled by default. **Supported traps:** • VLAN mismatch: existing generic trap DEVLINK_TRAP_GENERIC_ID_DMAC_MISMATCH Traps received packets with wrong VLAN tag • DMAC mismatch: new generic trap DEVLINK_TRAP_GENERIC_ID_DMAC_MISMATCH Traps received packets with wrong destination MAC Support added in user-space (N/A or package name + version): Devlink infrastructure (man7.org/linux/man-pages/man8/devlink-trap.8.html) Devlink provides an infrastructure called devlink trap which allow a device to register/unregister and to enable/disable traps. Devlink traps also provide traps grouping and policing. The trapped packets are monitored and then forward to the drop monitor. Drop monitor is used to send notifications to user space about dropped packets.**Note:** For this release, NIC WJH will not implement the policy.
ethtool Extended Link State	[General] Added ethtool extended link state to mlx5e. ethtool can be used to get more information to help troubleshoot the state. For example, if there is no link due to missing cable, run the following: $ ethtool eth1 ... Link detected: no (No cable) Besides the general extended state, drivers can pass additional information about the link state using the sub-state field. Example:

Feature/Change	Description
	$ ethtool eth1 ... Link detected: no (Autoneg, No partner detected) The extended state is available only for some cases of no link. In other cases, ethtool will print only "Link detected: no" as it did before.
RDMA	
DV "Signature API"	[ConnectX-5 and above] Added support for "Signature API" which, on supported devices, allows application-level data-integrity checks via a signature handover mechanism. Various signature types, including CRC32 and T10-DIF, can be automatically calculated and checked, stripped, or appended during the transfer at full wire speed.
ibv_query_qp_data_in_order() verb	[General] Added support for ibv_query_qp_data_in_order() API. This API enables an application to check if the given QP data is guaranteed to be in order, enabling poll for data instead of poll for completion.
Relaxed Ordering for Kernel ULPs	[ConnectX-4] Added support for enabling Relaxed Ordering for Kernel ULPs. Using relaxed ordering can improve performance in some setups. Since kernel ULPs are expected to support RO, it is enabled for them by default so they can benefit from it.
ah_to_qp Mapping	[ConnectX-6 Dx] Added support for mapping a QP to AH over DEVX API, which enables DC/UD QPs to use multiple CC algorithms in the same data center.
Steering UserSpace	
Matching on RAW Tunnel Headers	[ConnectX-5 and above] Added DR support for matching on RAW tunnel headers using the misc5 parameters, This feature allows matching on each bit of the header, inducing reserved fields.
Software Steering Insertion Rate Optimizations	[ConnectX-6 Dx] Added support for better insertion rate in software steering. This includes multi-QP which skips areas in the code that may be for debug only.
Software Steering Rule Optimization	[ConnectX-6 Dx] Improved rate of updating steering rules, insertion, and deletion. The feature includes definers, multi-qp approach, and better memory usage.

Feature/Change	Description
Duplicate Rules Insertion	[ConnectX-5 and above] Added support for ability to allow or prevent insertion of duplicate rules, so the user can choose one of the following behaviors: 1. Prevent duplicate rules, so that already-existing rule and fail can be detected. 2. Allow duplicate rules, to enable updating the rule's action (this will only take effect once the previous rule is deleted). By default, duplicate rules are allowed.
Improved Software Steering Rule Creation Stability	[ConnectX-6 Dx] Made it so that all rule's insertion occur in a defined time using defined (export) size of Htble and decreased use of dynamic allocation.

Feature/Change	Description
MLNX_EN 5.2-1.0.4.0	
Rx Multi-strides CQE Compression	[ConnectX-5 and above] Added CQE compression support for Rx multi-strides packets.
Multi-application QoS	[ConnectX-5 and above] Added support for configuring QoS on a single QP or on a group of QPs.
MPLS-over-UDP Hardware Offload Support	[ConnectX-5 and above] Added support for encap/decap hardware offload of IPv4 traffic over MPLS-over-UDP. This can be used in networks with MPLS routers to achieve more efficient routing.

Feature/Change	Description
Connection Tracking with Hairpin	[ConnectX-5 and above] Added support for adding connection tracking rules on VFs to forward traffic from one VF to the other.
sFlow Sampling Rules Offload	[ConnectX-5 and above] Added support for offloading sFlow sampling rules. sFlow is an industry standard technology for monitoring high speed switched networks. Open vSwitch integrated sFlow to extend the visibility into virtual servers, ensuring data center visibility and control. Added support for offloading sFlow sampling rules.
mlx5dv_dr Software Steering Parallel Rules Insertion	[ConnectX-5 & ConnectX-6 Dx/ BlueField & BlueField-2] Added support for a locking mechanism to enable parallel insertion of rules into the software steering using the mlx5dv_dr API. The parallel insertion improves the insertion rate and takes place when adding Rx and Tx rules via the FDB domain.
mlx5dv_dr API Matching on Geneve Tunnel	[ConnectX-5 & ConnectX-6 Dx/ BlueField & BlueField-2] Added support for the option to match mlx5dv_dr API on Geneve tunnel using a dynamic flex parser. The option header consists of class, type, length and data. The parser should be configured using devx command, after which a rule can be created to match on parser ID and data.
OVS-DPDK Geneve Encap/Decap	[ConnectX-5 & ConnectX-6 Dx/ BlueField & BlueField-2] Added support for Geneve tunneling offload, including matching on extension header.
OVS-DPDK Parallel Offloads	[ConnectX-5 & ConnectX-6 Dx/ BlueField & BlueField-2] Added support for parallel insertion and deletion of offloaded rules using multiple OVS threads.
GTP-U TEID Modification	[BlueField-2 & ConnectX-6 Dx] [**Beta**] Added support to modify GTP-U TEID. This support requires flex parser configuration.
OVS-DPDK E2E Cache Support	[BlueField-2 & ConnectX-6 Dx] [**Beta**] Improved performance of OVS Connection Tracking flows by enabling the merge of the multi-table flow matches and actions into one joint flow.
Tx Port Time-	[ConnectX-6 Dx and above] Transmitted packet timestamping accuracy can be improved when using a timestamp generated at the port level

Feature/Change	Description
Stamping	instead of a timestamp generated upon CQE creation. Tx port time-stamping better reflects the actual time of a packet's transmission. This feature is disabled by default. The feature can be enabled or disabled using the following command. ethtool --set-priv-flags <ifs-name> tx_port_ts on / off For further information on this feature, please see Tx Port Time-Stamping.
Tunnel Rules Offload	[ConnectX-6 Dx and above] Added support for offloading tunnel rules when the source interface is VF (in addition to uplink) in the Hypervisor.
	[ConnectX-6 Dx and above] Added support for offloading tunnel rules when the source interface is OpenvSwitch bridge (internal port).
Connection Tracking Mirroring Offload	[ConnectX-6 Dx and above] Added support for using Mirroring Offload with Connection Tracking.
mlx5dv_dr API ASO Flow Meter	[ConnectX-6 Dx and above] Added support for ASO flow meter using the mlx5dv_dr API, which allows for monitoring the packet rate for specific flows. When a packet hits a flow that is connected to a flow meter, the rate of packets through this meter is evaluated, and the packet is marked with a color copied into one of the C registers, according to the current rate compared to the reference rate.
mlx5dv_dr API ASO First Hit	[ConnectX-6 Dx and above] Added support for ASO first hit using the mlx5dv_dr API, which allows for tracking rule hits by packets. When a packet hits a rule with the ASO first hit action, a flag is set indicating this event, and the original value of the flag is copied to one of the C registers.
mlx5dv_dr API GTP-U Extension Header	[ConnectX-6 Dx and above] Added mlx5dv_dr API support for matching on a new field "gtpu_first_ext_dw_0". This field enables packet filtering based on the GTP-U first extension header (first dword only). To enable parsing of tunnel GTP-U extension header, run the following command. ./cloud_fw_reset.py FLEX_PARSER_PROFILE_ENABLE=3
IPsec Offload	[ConnectX-6 Lx and above] Added IPsec full offload support for extended sequence number, replay protection window and lifetime packet limit.
Firmware Upgrade	[All HCAs] Firmware upgrade during MLNX_EN installation is now done on all supported devices simultaneously rather than consecutively.

Feature/Change	Description
RDMA-CM Disassociate Support	[All HCAs] Added support for connecting kernel and RDMA-CM in a reliable way based on device index.
New Query GID API	[All HCAs] Added support for a new query GID API that allows for querying a single GID entry by its port and GID index, or querying for all GID tables of a specific device. This API works over ioctl instead of sysfs, which accelerates the querying process.
Multi-Host Firmware Reset	[All HCAs] Added support for performing multi-host firmware reset in order to upgrade the device firmware. Firmware reset loads the new firmware in case it was burnt on the flash and was pending activation, and reloads the current firmware image from the flash in case no new firmware was pending.
Firmware Live Patching	[All HCAs] [**Alpha**] Added support for firmware live patching in the driver. Live patching updates the firmware without the need to perform firmware reset. However, it can only be applied in scenarios where the difference between the current and new firmware versions are minor, which is decided upon by the firmware itself.
Devlink Firmware Reset	[All HCAs] Added support in the devlink tool for performing firmware reset in order to upgrade the device firmware. Firmware reset loads the new firmware in case it was burnt on the flash and was pending activation, and reloads the current firmware image from the flash in case no new firmware was pending. For further information, please refer to the the devlink man page. **Note**: In order for the firmware reset to run successfully, the following conditions should be met. • Each function should have the driver up and active with a version that supports this feature • None of the functions has the devlink parameter enable_remote_dev_reset set to False. **Note** The current MLNX_EN does not include the latest iproute2 version that provides support for this feature. Therefore, to be able to work with it, make

Feature/Change	Description
	sure to install the latest iproute2 version available on Github.
Command Interface Resiliency	[All HCAs] Added a resiliency mechanism for the driver to manually poll the command event queue (EQ) in case of a command timeout. In case the resiliency mechanism finds unhandled event queue entry (EQE) due to a lost interrupt, the driver will handle it, after which the command interface returns to a healthy state.
Offloaded Traffic Sniffer	[All HCAs] Setting a sniffer private flag is deprecated and no longer required. In order to capture offloaded/RoCE traffic, tcpdump can now be run on the RDMA device.
Devlink Port Health Reporters	[All HCAs] Added per-port reporters to devlink health to manage per-port health activities. Users can now access the devlink port reporters by specifying the port index in addition to the device devlink name through the devlink health commands API. This update was first introduced in iproute2 v5.8. As part of this feature, mlx5e Tx and Rx reporters are now redefined as devlink port reporters. For examples, please see devlink-health manpage.
Memory Registration Optimization	[All HCAs] Optimized memory consumption of memory registration in huge page systems. As an example, in a 2MB huge page system, 600 MB would be saved for 100 GB memory registration.
mlx5dv API	[All HCAs] Added support for mlx5dv API to modify the configured UDP source port for RoCE packets of a given RC/UC QP when QP is in RTS state.
Bug Fixes	See Bug Fixes.
Innova IPsec NIC Support	[Innova IPsec] Removed support for the network adapter Innova IPsec (EN).

Category	Description
Rev 5.1-1.0.4.0	

Category	Description
IP-in-IP RSS Offload	[ConnectX-4 and above] Added support for receive side scaling (RSS) offload in IP-in-IP (IPv4 and IPv6).
Devlink Port Support in Non-representor Mode	[ConnectX-4 and above] Added support for viewing the mlx5e physical devlink ports using the 'devlink port' command. This also may affect network interface names, if predictable naming scheme is configured. Suffix indicating a port number will be added to interface name.
Devlink Health State Notifications	[ConnectX-4 and above] Added support for receiving notifications on devlink health state changes when an error is reported or recovered by one of the reporters. These notifications can be seen using the userspace ‘devlink monitor’ command.
Legacy SR-IOV VF LAG Load Balancing	[ConnectX-4 and above] When VF LAG is in use, round-robin the Tx affinity of channels among the different ports, if supported by the firmware, enables all SQs of a channel to share the same port affinity. This allows the distribution of traffic sent from a VF between two ports, as well as round-robin the starting port among VFs to distribute traffic originating from single-core VMs.
RDMA-CM DevX Support	[ConnectX-4 and above] Added support for DevX in RDMA-CM applications.
RoCEv2 Flow Label and UDP Source Port Definition	[ConnectX-4 and above] This feature provides flow label and UDP source port definition in RoCE v2. Those fields are used to create entropy for network routes (ECMP), load balancers and 802.3ad link aggregation switching that are not aware of RoCE headers.
RDMA Tx Steering	[ConnectX-4 and above] Enabled RDMA Tx steering flow table. Rules in this flow table will allow for steering transmitted RDMA traffic.
Custom Parent-Domain Allocators for CQ	[ConnectX-4 and above] Enabled specific custom allocations for CQs.
mlx5dv Helper APIs for Tx	[ConnectX-4 and above] Added support for the following mlx5dv helper APIs which enable the user application to query or set a RAW QP's Tx affinity port number in a LAG configuration.

Category	Description
Affinity Port Selection	• mlx5dv_query_qp_lag_port • mlx5dv_modify_qp_lag_port
RDMA-CM Path Alignment	[ConnectX-4 and above] Added support for RoCE network path alignment between RDMA-CM message and QP data. The drivers and network components in RoCE calculate the same hash results for egress port selection both on the NICs and the switches.
CQ and QP Context Exposure	[ConnectX-4 and above] Exposed QP, CQ and MR context in raw format via RDMA tool.
In-Driver xmit_more	[ConnectX-4 and above] Enabled xmit_more feature by default in kernels that lack Rx bulking support (v4.19 and above) to ensure optimized IP forwarding performance when stress from Rx to Tx flow is insufficient. In kernels with Rx bulking support, xmit_more is disabled in the driver by default, but can be enabled to achieve enhanced IP forwarding performance.
Relaxed Ordering	[ConnectX-4 and above] Relaxed ordering is a PCIe feature which allows flexibility in the transaction order over the PCIe. This reduces the number of retransmissions on the lane, and increases performance up to 4 times. By default, mlx5e buffers are created with Relaxed Ordering support when firmware capabilities are on and the PCI subsystem reports that CPU is not on the kernel's blocklist. **Note**: Some CPUs which are not listed in the kernel's blocklist may suffer from buggy implementation of relaxed ordering, in which case the user may experience a degradation in performance and even unexpected behavior. To turn off relaxed ordering and restore previous behavior, run setpci command as instructed here. Example: "RlxdOrd-“ : setpci -s82:00.0 CAP_EXP+8.w=294e
ODP Huge Pages Support	[ConnectX-4 and above] Enabled ODP Memory Region (MR) to work with huge pages by exposing IBV_ACCESS_HUGETLB access flag to indicate that the MR range is mapped by huge pages. The flag is applicable only in conjunction with IBV_ACCESS_ON_DEMAND.
Offloaded Traffic Sniffer	[ConnectX-4 and above] Removed support for Offloaded Traffic Sniffer feature and replaced its function with Upstream solution tcpdump tool.

Category	Description
Connection Tracking Offload	[ConnectX-5 and above] Added support for offloading TC filters containing connection tracking matches and actions.
Dual-Port RoCE Support	[ConnectX-5 and above] Enabled simultaneous operation of dual-port RoCE and Ethernet in SwitchDev mode.
IP-in-IP Tunnel Offload for Checksum and TSO	[ConnectX-5 and above] Added support for the driver to offload checksum and TSO in IP-in-IP tunnels.
Packet Pacing DevX Support	[ConnectX-5 and above] Enabled RiverMax to work over DevX with packet pacing functionality by exposing a few DV APIs from rdma-core to enable allocating/destroying a packet pacing index. For further details on usage, see man page for: mlx5dv_pp_alloc() and mlx5dv_pp_free().
Software Steering Support for Memory Reclaiming	[ConnectX-5 and above] Added support for reclaiming device memory to the system when it is not in use. This feature is disabled by default and can be enabled using the command mlx5dv_dr_domain_set_reclaim_device_memory().
SR-IOV Live Migration	[ConnectX-5 and above] [**Beta**] Added support for performing a live migration for a VM with an SR-IOV NIC VF attached to it and with minimal to no traffic disruption. This feature is supported in SwitchDev mode; enabling users to fully leverage VF TC/OVS offloads, where the failover inbox driver is in the Guest VM, and the bonding driver is in the Hypervisor. Note that you must use the latest QEMU and libvirt from the Upstream github.com sources.
Uplink Representor Modes	[ConnectX-5 and above] Removed support for new_netdev mode in SwitchDev mode. The new default behaviour is to always keep the NIC netdev.
OVS-DPDK Offload Statistics	[ConnectX-5 and above] Added support for dumping connection tracking offloaded statistics.

Category	Description
OVS-DPDK Connection Tracking Labels Exact Matching	[ConnectX-5 and above] Added support for labels exact matching in OVS-DPDK CT openflow rules.
Kernel Software Managed Flow Steering (SMFS) Performance	[ConnectX-5 and above] Improved the performance of Kernel software steering by reducing its memory consumption.
OVS-DPDK LAG Support	[ConnectX-5 & ConnectX-6 Dx] Added support for LAG (modes 1,2,4) with OVS-DPDK.
Get FEC Status on PAM4/50G	[ConnectX-6 and above] Allowed configuration of Reed Solomon and Low Latency Reed Solomon over PAM4 link modes.
RDMA-CM Enhanced Connection Establishment (ECE)	[ConnectX-6 and above] Added support for allowing automatic enabling/disabling of vendor specific features during connection establishment between network nodes, which is performed over RDMA-CM messaging interface.
RoCE Selective Repeat	[ConnectX-6 and above] This feature introduces a new QP retransmission mode in RoCE in which dropped packet recovery is done by re-sending the packet instead of re-sending the PSN window only (Go-Back-N protocol). This feature is enabled by default when RDMA-CM is being used and both connection nodes support it.
IPsec Full Offload	[ConnectX-6 Dx & BlueField-2] [**Beta**] Added support for IPsec full offload (VxLAN over ESP transport).
Hardware vDPA on OVS-DPDK	[ConnectX-6 Dx & BlueField-2] Added support for configuring hardware vDPA on OVS-DPDK. This support includes the option to fall back to Software vDPA in case the NIC installed on the driver does not support hardware vDPA.
IPsec Crypto Offloads	[ConnectX-6 Dx] Support for IPsec Crypto Offloads feature over ConnectX-6 Dx devices and up is now at GA level.

<table>
<tr><th>Category</th><th>Description</th></tr>
<tr><td>TLS Tx Hardware Offload</td><td>[ConnectX-6 Dx] Support for TLS Tx Hardware Offload feature over ConnectX-6 Dx devices and up is now at GA level.</td></tr>
<tr><td>TLS Rx Hardware Offload</td><td>[ConnectX-6 Dx] [Alpha] Added support for hardware offload decryption of TLS Rx traffic over crypto-enabled ConnectX-6 Dx NICs and above.</td></tr>
<tr><td>Userspace Software Steering ConnectX-6 Dx Support</td><td>[ConnectX-6 Dx] Support for software steering on ConnectX-6 Dx adapter cards in the user-space RDMA-Core library through the mlx5dv_dr API is now at GA level.</td></tr>
<tr><td>Kernel Software Steering ConnectX-6 Dx Support</td><td>[ConnectX-6 Dx] [Beta] Added support for kernel software steering on ConnectX-6 Dx adapter cards.</td></tr>
<tr><td>Adapters</td><td>[ConnectX-6 Lx] Added support for ConnectX-6 Lx adapter cards.</td></tr>
<tr><td>RDMA-Core Migration</td><td>[All HCAs] As of MLNX_EN v5.1, Legacy verbs libraries have been fully replaced by RDMA-Core library.
For the list of new APIs used for various MLNX_EN features, please refer to the Migration to RDMA-Core document.</td></tr>
<tr><td>Firmware Reactivation</td><td>[All HCAs] Added support for safely inserting consecutive firmware images without the need to reset the NIC in between.</td></tr>
<tr><td>UCX-CUDA Support</td><td>[All HCAs] UCX-CUDA is now supported on the following OSs and platforms.

<table><tr><th>OS</th><th>Platform</th></tr><tr><td>RedHat 7.6 ALT</td><td>PPC64LE</td></tr><tr><td>RedHat 7.7</td><td>x86_64</td></tr><tr><td>RedHat 7.8</td><td>PPC64LE/x86_64</td></tr><tr><td>RedHat 7.9</td><td>x86_64</td></tr><tr><td>RedHat 8.1</td><td>x86_64</td></tr></table></td></tr>
</table>

Category	Description
	<table><tr><th>OS</th><th>Platform</th></tr><tr><td>RedHat 8.2</td><td>x86_64</td></tr></table>
HCOLL-CUDA	[All HCAs] The hcoll package includes a CUDA plugin (hmca_gpu_cuda.so). As of MLNX_EN v5.1, it is built on various platforms as the package hcoll-cuda. It will be installed by default if the system has CUDA 10-2 installed. **Notes**: • If you install MLNX_EN from a package repository, you will need to install the package hcoll-cuda explicitly to be able to use it. • HCOLL-CUDA is supported on the same OSs that include support for UCX-CUDA (listed in the table above), except for RedHat 8.1 and 8.2.
GPUDirect Storage (GDS)	[All HCAs] [**Beta**] Added support for the new technology of GDS (GPUDirect Storage) which enables a direct data path between local or remote storage, such as NFS, NVMe or NVMe over Fabric (NVMe-oF), and GPU memory. Both GPUDirect RDMA and GPUDirect Storage avoid extra copies through a bounce buffer in the CPU's memory. They enable the direct memory access (DMA) engine near the NIC or storage to move data on a direct path into or out of GPU memory, without burdening the CPU or GPU. To enable the feature, run ./install --with-nfsrdma –-with-nvmf --enable-gds --add-kernel-support To get access to GDS Beta, please reach out to the GDS team at GPUDirectStorageExt@nvidia.com. For the list of operating systems on which GDS is supported, see here.
Bug Fixes	See Bug Fixes.

Category	Description
Rev 5.0-1.0.0.0	
Adapters	[ConnectX-6 Dx] Added support for ConnectX-6 Dx adapter cards.
Userspace Software Steering	[**Beta**] Added support for software steering on ConnectX-6 Dx adapter cards in the user-space RDMA-Core library through the mlx5dv_dr API.

Category	Description
ConnectX-6 Dx Support	
Virtual Output Queuing (VoQ) Counters	[ConnectX-6 Dx and above] Exposed rx_prio[p]_buf_discard, rx_prio[p]_wred_discard and rx_prio[p]_marked firmware counters that count the number of packets that were dropped due to insufficient resources.
IPsec Crypto Offloads	[ConnectX-6 Dx and above] [**Beta**] IPsec crypto offloads are now supported on ConnectX-6 Dx devices and up. The offload functions use the existing ip xfrm tool to activate offloads on the device. It supports transport/tunnel mode with AES-GCM IPsec scheme.
TLS TX Hardware Offload	[ConnectX-6 Dx and above, not including ConnectX-6 Lx] [Alpha] Added support for hardware offload encryption of TLS traffic.
VirtIO Acceleration through Datapath I/O Processor (vDPA)	[ConnectX-6 Dx and above] Added support to enable mapping the VirtIO access region (VAR) to be used for doorbells by vDPA applications. Specifically, the following DV APIs were introduced (see man page for more details): • mlx5dv_alloc_var() • mlx5dv_free_var()
Resource Allocation on External Memory	[ConnectX-5 and above] Added support to enable overriding mlx5 internal allocations in order to let applications allocate some resources on external memory, such as that of the GPU. The above is achieved by extending the parent domain object with custom allocation callbacks. Currently supported verbs objects are: QP, DBR, RWQ, SRQ.
Hardware Clock Exposure	[ConnectX-5 and above] Added support for querying the adapter clock via mlx5dv_query_device.
ODP Diagnostic Counters	[ConnectX-5 and above] Added ODP diagnostics counters for the following items per MR (memory region) within IB/mlx5 driver: 1. Page faults: Total number of faulted pages. 2. Page invalidations: Total number of pages invalidated by the OS during all invalidation events. The translations can no longer be

Category	Description
	valid due to either non-present pages or mapping changes. 3. Prefetched pages: When prefetching a page, a page fault is generated in order to bring the page to the main memory.
Devlink Health CR-Space Dump	[ConnectX-5 and above] Added the option to dump configuration space via the devlink tool in order to improve debug capabilities.
Multi-packet TX WQE Support for XDP Transmit Flows	[ConnectX-5 and above] The conventional TX descriptor (WQE or Work Queue Element) describes a single packet for transmission. Added driver support for the HW feature of multi-packet TX WQEs in XDP transmit flows. With this, the HW becomes capable of working with a new and improved WQE layout that describes several packets. In effect, this feature saves PCI bandwidth and transactions, and improves transmit packet rate.
OVS-Kernel ToS Rewrite	[ConnectX-5 and above] Added support for Type of Service (ToS) rewrite in the OVS-Kernel.
OVS-Kernel Mirroring	[ConnectX-5 and above] Added support for mirroring output in SwitchDev mode in the OVS-Kernel. The mirroring port may either be a local or a remote VF, using VxLAN or GRE encapsulations.
GENEVE Encap/Decap Rules Offload	[ConnectX-5 and above] Added support for GENEVE encapsulation/decapsulation rules offload.
GPRS Tunneling Protocol (GTP) Header	[ConnectX-5 and above] [**Beta**] Added support for matching (filtering) GTP header-based packets using mlx5dv_dr API over user-space RDMA-Core library.
Multi Packet Tx WQE Support for XDP Transmit Flows	[ConnectX-5 and above] Added driver support for the hardware feature of multi-packet Tx to work with a new and improved WQE layout that describes several packets instead of a single packet for XDP transmission flows. This saves PCI bandwidth and transactions, and improves transmit packet rate.
Userspace Software Steering	[ConnectX-5 and above] [**Beta**] Added support for software steering to dump flows for debugging purposes in the user-space RDMA-Core library through the mlx5dv_dr API.

Category	Description
Debugging API	
Kernel Software Steering for Connection Tracking (CT)	[ConnectX-5 and above] [**Beta**] Added support for updating CT rules using the software steering mechanism.
Kernel Software Steering Remote Mirroring	[ConnectX-5 and above] [**Beta**] Added support for updating remote mirroring rules using the software steering mechanism.
Discard Counters	[ConnectX-4 and above] Exposed rx_prio[p]_discards discard counters per priority that count the number of received packets dropped due to lack of buffers on the physical port.
MPLS Traffic	[ConnectX-4 and above] Added support for reporting TSO and CSUM offload capabilities for MPLS tagged traffic and, allowed the kernel stack to use these offloads.
mlx5e Max Combined Channels	[ConnectX-4 and above] Increased the driver's maximal combined channels value from 64 to 128 (however, note that OOB value will not cross 64). 128 is the upper bound. Lower maximal value can be seen on the host, depending on the number of cores and MSIX's configured by the firmware.
RoCE Accelerator Counters	[ConnectX-4 and above] Added the following RoCE accelerator counters: • **roce_adp_retrans** - counts the number of adaptive retransmissions for RoCE traffic • **roce_adp_retrans_to** - counts the number of times RoCE traffic reached timeout due to adaptive retransmission • **roce_slow_restart** - counts the number of times RoCE slow restart was used • **roce_slow_restart_cnps** - counts the number of times RoCE slow restart generated CNP packets

Category	Description
	• **roce_slow_restart_trans** - counts the number of times RoCE slow restart changed state to slow restart
Migration to RDMA-Core	[All HCAs] The default installation of the userspace is now the RDMA-Core library instead of the legacy verbs. This achieves most of the legacy experimental verbs' functionalities, and more. For NVIDIA VMA or NVIDIA RiverMax, use experimental verbs (prefix "ibv_exp"). For further information on the migration to RDMA-Core and the list of new APIs used for various MLNX_EN features, please refer to the Migration to RDMA-Core document.
ibdev2netdev Tool Output	[All HCAs] ibdev2netdev tool output was changed such that the bonding device now points at the bond instead of the slave interface.
Memory Region	[All HCAs] Added support for the user to register memory regions with a relaxed ordering access flag. This can enhance performance, depending on architecture and scenario.
Devlink Health Reporters	[All HCAs] Added support for monitoring and recovering from errors that occur on the RX queue, such as cookie errors and timeout.
GSO Optimization	[All HCAs] Improved GSO (Generic Segmentation Offload) workload performance by decreasing doorbells usage to the minimum required.
TX CQE Compression	[All HCAs] Added support for TX CQE (Completion Queue Element) compression. Saves on outgoing PCIe bandwidth by compressing CQEs together. Disabled by default. Configurable via private flags of ethtool.
Firmware Versions Query via Devlink	[All HCAs] Added the option to query for running and stored firmware versions using the devlink tool.
Firmware Flash Update via Devlink	[All HCAs] Added the option to update the firmware image in the flash using the devlink tool. Usage: devlink dev flash <dev> file <file_name>.mfa2 For further information on how to perform this update, see "Updating Firmware Using ethtool/devlink and .mfa2 File" section in MFT User Manual.

Category	Description
Devlink Health WQE Dump	[All HCAs] Added support for WQE (Work Queue Element) dump, triggered by an error on Rx/Tx reporters. In addition, some dumps (not triggered by an error) can be retrieved by the user via devlink health reporters.
GENEVE Tunnel Stateless Offload	[All HCAs] Added support for GENEVE tunneled hardware offloads of TSO, CSUM and RSS.
TCP Segmentation and Checksum Offload	[All HCAs] Added TCP segmentation and checksum offload support for MPLS-tagged traffic.
Bug Fixes	See Bug Fixes.

Category	Description
Rev 4.7-3.2.9.0	
Uplink Representor Modes	Added support for the following Uplink Representor modes: 1. new_netdev: default mode - when found in this mode, the uplink representor is created as a new netdevice 2. nic_netdev: when found in this mode, the NIC netdevice acts as an uplink representor device **Example**: echo nic_netdev > /sys/class/net/ens1f0/compat/devlink/uplink_rep_mode **Notes:** • The mode can only be changed when found in Legacy mode • The mode is not saved when reloading mlx5_core
mlx5_core	Added new mlx5_core module parameter "num_of_groups", which controls the number of large groups in the FDB flow table. **Note**: In MLNX_OFED v4.6-3.1.9.0.14, the default value of num_of_groups was 15, while in the current MLNX_OFED v4.7-3, the default value is 4. In order to achieve the same OOB experience, make sure to set the num_of_groups module parameter to 15 prior to driver load.

Category	Description
	For further information, please refer to Performance Tuning Based on Traffic Patterns section in MLNX_OFED User Manual.
VFs Groups Minimum Bandwidth Rate	Added support for setting a minimum bandwidth rate on a group of VFs (BW guarantee) to ensure this group is able to transmit at least the amount of bandwidth specified on the wire.
Direct Verbs Support for Batch Counters on Root Table	Added support for mlx5dv_dr API to set batch counters for root tables.
Modify Header	Added support for mlx5dv_dr_actions to support up to 32 modify actions.
mlx5dv_dr Memory Consumption	Reduced the mlx5dv_dr API memory consumption by improving the memory allocator.
mlx5dv_dr Memory Allocation	Reduced memory allocation time when using the mlx5dv_dr API. This is particularly significant for the first inserted rules on which memory is allocated.
Mediated Devices	Added support for mediated devices that allows the creation of accelerated devices without SR-IOV on the Bluefield® system. For further information on mediated devices and how to configure them, please refer to Mediated Devices section in MLNX_EN User Manual.

Category	Description
Rev 4.7-1.0.0.1	
Counters Monitoring	[ConnectX-4 and above] Added support for monitoring selected counters and generating a notification event (Monitor_Counter_Change event) upon changes made to these counters. The counters to be monitored are selected using the SET_MONITOR_COUNTER command.
EEPROM Device	[ConnectX-4 and above] Added support to read additional EEPROM information from high pages of modules such as SFF-8436 and SFF-8636. Such information can be: 1. Application Select table 2. User writable

Category	Description
Thresholds via Ethtool	EEPROM 3. Thresholds and alarms - Ethtool dump works on active cables only (e.g. optic), but thresholds and alarms can be read with “offset” and “length” parameters in any cable by running: ethtool -m <DEVNAME> offset X length Y
RDMA_RX RoCE Steering Support	[ConnectX-4 and above] Added the ability to create rules to steer RDMA traffic, with two destinations supported: DevX object and QP. Multiple priorities are also supported.
ASAP^2	[ConnectX-5 and above] Incorporated the documentation of *Accelerated Switching And Packet Processing (ASAP^2): Hardware Offloading for vSwitches* into MLNX_OFED Release Notes and User Manual.
MLNX_OFED Installation via Repository	[All HCAs] The repository providing legacy verbs has been moved from RPMS or DEBS folders to RPMS/MLNX_LIBS and DEBS/MLNX_LIBS. In addition, a new repository providing RDMA-Core based userspace has been added to RPMS/UPSTREAM_LIBS and DEBS/UPSTREAM_LIBS.

Category	Description
Rev 4.6-1.0.1.1	
Devlink Configuration Parameters Tool	[ConnectX-3/ConnectX-3 Pro] Added support for a set of configuration parameters that can be changed by the user through the Devlink user interface.
ODP Pre-fetch	[ConnectX-4 and above] Added support for pre-fetching a range of an on-demand paging (ODP) memory region (MR), this way reducing latency by making pages present with RO/RW permissions before the actual IO is conducted.
DevX Privilege Enforcement	[ConnectX-4 and above] Enforced DevX privilege by firmware. This enables future device functionality without the need to make driver changes unless a new privilege type is introduced.
DevX Interoperability APIs	[ConnectX-4 and above] Added support for modifying and/or querying for a verb object (including CQ, QP, SRQ, WQ, and IND_TBL APIs) via the DevX interface. This enables interoperability between verbs and DevX.

Category	Description
DevX Asynchronous Query Commands	[ConnectX-4 and above][ConnectX-4 and above] Added support for running QUERY commands over the DevX interface in an asynchronous mode. This enables applications to issue many commands in parallel while firmware processes the commands.
DevX User-space PRM Handles Exposure	[ConnectX-4 and above] Exposed all PRM handles to user-space so DevX user application can mix verbs objects with DevX objects. For example: Take the cqn from the created ibv_cq and use it on a devx)create(QP).
Indirect Mkey ODP	[ConnectX-4 and above] Added the ability to create indirect Mkeys with ODP support over DevX interface.
XDP Redirect	[ConnectX-4 and above] Added support for XDP_REDIRECT feature for both ingress and egress sides. Using this feature, incoming packets on one interface can be redirected very quickly into the transmission queue of another capable interface. Typically used for load balancing.
RoCE Disablement	[ConnectX-4 and above] Added the option to disable RoCE traffic handling. This enables forwarding of traffic over UDP port 4791 that is handled as RoCE traffic when RoCE is enabled. When RoCE is disabled, there is no GID table, only Raw Ethernet QP type is supported and RoCE traffic is handled as regular Ethernet traffic.
Forward Error Correction (FEC) Encoding	[ConnectX-4 and above] Added the ability to query and modify Forward Error Correction (FEC) encoding, as well as disabling it via Ethtool.
RAW Per-Lane Counters Exposure	[ConnectX-4 and above] Exposed RAW error counters per cable-module lane via ethtool stats. The counters show the number of errors before FEC correction (if enabled). For further information, please see phy_raw_errors_lane[i] under **Physical Port Counters** section in Understanding mlx5 ethtool Counters Community post.
VF LAG	[ConnectX-4 Lx and above] Added support for High Availability and load balancing for Virtual Functions of different physical ports in SwitchDev SR-IOV mode.
ASAP^2 Offloading VXLAN	[ConnectX-5 and above] Added support for performing hardware Large Receive Offload (HW LRO) on VFs with HW-decapsulated VXLAN.

Category	Description
Decapsulation with HW LRO	For further information on the VXLAN decapsulation feature, please refer to ASAP^2 User Manual under nvidia.com/en-us/networking/.com Products Software ASAP^2.
PCI Atomic Operations	[ConnectX-5 and above] Added the ability to run atomic operations on local memory without involving verbs API or compromising the operation's atomicity.
Virtual Ethernet Port Aggregator (VEPA)	[ConnectX-5] Added support for activating/deactivating Virtual Ethernet Port Aggregator (VEPA) mode on a single virtual function (VF). To turn on VEPA on the second VF, run: echo ON > /sys/class/net/enp59s0/device/sriov/1/vepa
VFs Rate Limit	[ConnectX-5] Added support for setting a rate limit on groups of Virtual Functions rather on an individual Virtual Function.
ConnectX-6 Support	[ConnectX-6] [**Beta**] Added support for ConnectX-6 (VPI only) adapter cards. **NOTE**: In HDR installations that are built with remotely managed Quantum-based switches, the switch's firmware must be upgraded to version 27.2000.1142 prior to upgrading the HCA's (ConnectX-6) firmware to version 20.25.1500. When using ConnectX-6 HCAs with firmware v20.25.1500 and connecting them to Quantum-based switches, make sure the Quantum firmware version is 27.2000.1142 in order to avoid any critical link issues.
Ethtool 200Gbps	[ConnectX-6] ConnectX-6 hardware introduces support for 200Gbps and 50Gbps-per-lane link mode. MLNX_OFED supports full backward compatibility with previous configurations. Note that in order to advertise newly added link-modes, the full bitmap related to the link modes must be advertised from ethtool man page. For the full bitmap list per link mode, please refer to MLNX_OFED User Manual. **NOTE**: This feature is firmware-dependent. Currently, ConnectX-6 Ethernet firmware supports up to 100Gbps only. Thus, this capability may not function properly using the current driver and firmware versions.
PCIe Power State	[ConnectX-6] Added support for the following PCIe power state indications to be printed to dmesg: 1. **Info message #1**: PCIe slot power capability was not advertised.

Category	Description
	2. **Warning message**: Detected insufficient power on the PCIe slot (xxxW). 3. **Info message #2**: PCIe slot advertised sufficient power (xxxW). When indication #1 or #2 appear in dmesg, user should make sure to use a PCIe slot that is capable of supplying the required power.
Message Signaled Interrupts-X (MSI-X) Vectors	[mlx5 Driver] Added support for using a single MSI-X vector for all control event queues instead of one MSI-X vector per queue in a virtual function driver. This frees extra MSI-X vectors to be used for completion event queue, allowing for additional traffic channels in the network device.
Send APIs	[mlx5 Driver] Introduced a new set of QP Send operations (APIs) which allows extensibility for new Send opcodes.
BlueField Support	[BlueField] BlueField is now fully supported as part of the NVIDIA OFED mainstream version sharing the same code baseline with all the adapters product line.
Representor Name Change	[BlueField] In SwitchDev mode: • Uplink representors are now called p0/p1 • Host PF representors are now called pf0hpf/pf1hpf • VF representors are now called pf0vfN/pf1vfN
ECPF Net Devices	[BlueField] In SwitchDev mode, net devices enp3s0f0 and enp3s0f1 are no longer created.
Setting Host MAC and Tx Rate Limit from ECPF	[BlueField] Expanded to support VFs as well as the host PFs.
RDMA-CM Application Managed QP	[All HCAs] Added support for the RDMA application to manage its own QPs and use RDMA-CM only for exchanging Address information.
RDMA-CM QP Timeout Control	[All HCAs] Added a new option to rdma_set_option that allows applications to override the RDMA-CM's QP ACK timeout value.

Category	Description
MLNX_OFED Verbs API	[All HCAs] As of MLNX_OFED v5.0 release (Q1 2020) onwards, MLNX_OFED Verbs API will be migrated from the legacy version of the user space verbs libraries (libibervs, libmlx5 ..) to the upstream version rdma-core. More details are available in MLNX_OFED user manual under Installing Upstream rdma-core Libraries.

Category	Description
4.5-1.0.1.0	
VFs per PF	[ConnectX-5] Increased the amount of maximum virtual functions (VF) that can be allocated to a physical function (PF) to 127 VF.
SW-Defined UDP Source Port for RoCE v2	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] UDP source port for RoCE v2 packets is now calculated by the driver rather than the firmware, achieving better distribution and less congestion. This mechanism works for RDMA- CM QPs only, and ensures that RDMA connection messages and data messages have the same UDP source port value.
Local Loopback Disable	[mlx5 Driver] Added the ability to manually disable Local Loopback regardless of the number of open user-space transport domains.
Adapter Cards	[ConnectX-6] Added support for ConnectX-6 Ready. For further information, please contact NVIDIA Support.
Bug Fixes	See Bug Fixes.
4.4-2.0.7.0	
Operating Systems	[All HCAs] Added support for additional OSs. See "General Support" section.
4.4-1.0.1.0	
Adaptive Interrupt Moderation	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for adaptive Tx, which optimizes the moderation values of the Tx CQs on runtime for maximum throughput with minimum CPU overhead. This mode is enabled by default.
	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Updated Adaptive Rx to ignore ACK packets so that queues that only handle ACK packets remain with the default moderation.

Category	Description
Docker Containers [Beta]	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for Docker containers to run over Virtual RoCE and InfiniBand devices using SR-IOV mode.
Firmware Tracer	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added a new mechanism for the device's FW/HW to log important events into the event tracing system (/sys/kernel/debug/tracing) without requiring any NVIDIA-specific tool. Note: This feature is enabled by default.
CR-Dump	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Accelerated the original cr-dump by optimizing the reading process of the device's CR-Space snapshot.
VST Q-in-Q	[ConnectX-4/ConnectX-4 Lx] Added support for C-tag (0x8100) VLAN insertion to tagged packets in VST mode.
OVS Offload using ASAP2	[ConnectX-4 Lx/ConnectX-5] Added support for NVIDIA Accelerated Switching And Packet Processing (ASAP2) technology, which allows OVS offloading by handling OVS data-plane, while maintaining OVS control-plane unmodified. OVS Offload using ASAP2 technology provides significantly higher OVS performance without the associated CPU load. For further information, refer to ASAP2 Release Notes under nvidia.com/en-us/networking/.com Products Software ASAP^2.
4.3-1.0.1.0	
Adaptive Interrupt Moderation	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for adaptive Tx, which optimizes the moderation values of the Tx CQs on runtime for maximum throughput with minimum CPU overhead. This mode is enabled by default.
	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Updated Adaptive Rx to ignore ACK packets so that queues that only handle ACK packets remain with the default moderation.
Docker Containers [Beta]	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for Docker containers to run over Virtual RoCE and InfiniBand devices using SR-IOV mode.
Firmware Tracer	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added a new mechanism for the device's FW/HW to log important events into the event tracing

Category	Description
	system (/sys/kernel/debug/tracing) without requiring any NVIDIA-specific tool. Note: This feature is enabled by default.
CR-Dump	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Accelerated the original cr-dump by optimizing the reading process of the device's CR-Space snapshot.
VST Q-in-Q	[ConnectX-4/ConnectX-4 Lx] Added support for C-tag (0x8100) VLAN insertion to tagged packets in VST mode.
OVS Offload using ASAP2	[ConnectX-4 Lx/ConnectX-5] Added support for NVIDIA Accelerated Switching And Packet Processing (ASAP2) technology, which allows OVS offloading by handling OVS data-plane, while maintaining OVS control-plane unmodified. OVS Offload using ASAP2 technology provides significantly higher OVS performance without the associated CPU load. For further information, refer to ASAP2 Release Notes under nvidia.com/en-us/networking/.com Products Software ASAP^2.
4.3-1.0.1.0	
Erasure Coding Offload verbs	[ConnectX-5] Added support for erasure coding offload software verbs (encode/decode/update API) supporting a number of redundancy blocks (m) greater than 4.
Virtual MAC	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Removed support for Virtual MAC feature.
RoCE LAG	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added out of box RoCE LAG support for RHEL 7.2 and RHEL 6.9.
Dropped Counters	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added a new counter *rx_steer_missed_packets* which provides the number of packets that were received by the NIC, yet were discarded/dropped since they did not match any flow in the NIC steering flow table.
	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added the ability for SR-IOV counter *rx_dropped* to count the number of packets that were dropped while vport was down.
Reset Flow	[mlx5 Driver] Added support for triggering software reset for firmware/driver recovery. When fatal errors occur, firmware can be reset and driver reloaded.

Category	Description
Striding RQ with HW Time-Stamping	[ConnectX-4 Lx/ConnectX-5] Added the option to retrieve the HW timestamp when polling for completions from a completion queue that is attached to a multi-packet RQ (Striding RQ).
4.2-1.0.1.0	
Physical Address Memory Allocation	[mlx5 Driver] Added support to register a specific physical address range.
Innova IPsec Adapter Cards	[Innova IPsec EN] Added support for NVIDIA Innova IPsec EN adapter card, that provides security acceleration for IPsec-enabled networks.
Precision Time Protocol (PTP)	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for PTP feature over PKEY interfaces. This feature allows for accurate synchronization between the distributed entities over the network. The synchronization is based on symmetric Round Trip Time (RTT) between the master and slave devices, and is enabled by default.
Virtual MAC	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for Virtual MAC feature, which allows users to add up to 4 virtual MACs (VMACs) per VF. All traffic that is destined to the VMAC will be forwarded to the relevant VF instead of PF. All traffic going out from the VF with source MAC equal to VMAC will go to the wire also when Spoof Check is enabled. For further information, please refer to “Virtual MAC” section in MLNX_EN User Manual.
Receive Buffer	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added the option to change receive buffer size and cable length. Changing cable length will adjust the receive buffer's xon and xoff thresholds. For further information, please refer to “Receive Buffer” section in MLNX_EN User Manual.
GRE Tunnel Offloads	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for the following GRE tunnel offloads: • TSO over GRE tunnels • Checksum offloads over GRE tunnels • RSS spread for GRE packets

Category	Description
NVMEoF	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for the host side (RDMA initiator) in RedHat 7.2 and above.
Dropless Receive Queue (RQ)	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for the driver to notify the FW when SW receive queues are overloaded.
PFC Storm Prevention	[ConnectX-4/ConnectX-4 Lx/ConnectX-5] Added support for configuring PFC stall prevention in cases where the device unexpectedly becomes unresponsive for a long period of time. PFC stall prevention disables flow control mechanisms when the device is stalled for a period longer than the default pre-configured timeout. Users now have the ability to change the default timeout by moving to auto mode. For further information, please refer to “PFC Stall Prevention” section in MLNX_EN User Manual.
Q-in-Q	[ConnectX-5] Added support for Q-in-Q VST feature in ConnectX-5 adapter cards family.
Virtual Guest Tagging (VGT+)	[ConnectX-5] Added support for VGT+ in ConnectX-4/ConnectX-5 HCAs. This feature is s 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. The policy determines which VLAN IDs are allowed to be transmitted or received. The policy does not determine the user priority, which is left unchanged. For further information, please refer to “Virtual Guest Tagging (VGT+)” section in MLNX_EN User Manual.
Tag Matching Offload	[ConnectX-5] Added support for hardware Tag Matching offload with Dynamically Connected Transport (DCT).
CR-DUMP	[All HCAs] Added support for the driver to take an automatic snapshot of the device’s CR-Space in cases of critical failures. For further information, please refer to “CRDUMP” section in MLNX_EN User Manual.
4.1-1.0.2.0	
RoCE Diagnostics and ECN Counters	[mlx5 Driver] 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.

Category	Description
rx-fcs Offload (ethtool)	[mlx5 Driver] 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	[mlx5 Driver] 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 NVIDIA Adapters* Community post.
RoCE ECN Parameters	[mlx5 Driver] 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	[mlx5 Driver] Added support for mlx_fs_dump, which is a python tool that prints the steering rules in a readable manner.
Secure Firmware Updates	[mlx5 Driver] 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 NVIDIA Firmware Tools (MFT) User Manual.
PeerDirect	[mlx5 Driver] 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.
Probed VFs	[mlx5 Driver] 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	[mlx5 Driver] 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

Category	Description
	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)	[mlx5 Driver] 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 NVIDIA Adapters* Community post.
Fast Driver Unload	[mlx5 Driver] Added support for fast driver teardown in shutdown and kexec flows.
NVMEoF Target Offload	[ConnectX-5/ConnectX-5 Ex] 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* .
RDMA CM	[All HCAs] 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.
Lustre	[All HCAs] Added support for Lustre file system open-source project.
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.

Category	Description
NFS over RDMA (NFSoRDMA)	Removed support for NFSoRDMA drivers. These drivers are no longer provided along with the MLNX_EN package.

Customer Affecting Change	Description
Customer Affecting Changes 5.6-1.0.3.5	
Interface Renaming, PF/VF, Udev	The OFED driver no longer performs Ethernet NetDev interface renaming for PFs and VFs. The udev rules file which implemented renaming (82-net-setup-link.rules) and its supporting script vf-net-link-name.sh are no longer installed by default. Renaming is thus performed by underlying mechanisms -- in udev, in the kernel, and in the BIOS. Users who wish to continue using the OFED driver renaming mechanism must add option —copy-ifnames-udev to the OFED install command. To install these files at a later time, copy them from one of the following directories: • /usr/share/doc/mlnx-ofa_kernel (RHEL8 and newer) • /usr/share/doc/mlnx-ofa_kernel-[1-9]* (RHEL 7.X) • /usr/share/doc/packages/mlnx-ofa_kernel (SLES) • /usr/share/doc/mlnx-ofed-kernel-utils/examples (Debian-based releases) **Note** • File 82-net-setup-link.rules should be copied to directory /etc/udev/rules.d • File vf-net-link-name.sh should be copied to directory /etc/infiniband (make sure that it has both read and execute permission) • After copying over the files, the driver should be restarted for the copied files to take effect

Customer Affecting Change	Description
	• Customers who wish prevent renaming of NetDev names should add "net.ifnames=0 biosdevname=0" to the kernel boot command line, and then reboot the host
Community Operating Systems	Starting OFED 5.6, NVIDIA is introducing a new support model for OFED used on open source community operating systems. The goal of this new support model is to enable customers to use community-maintained variants of the Linux operating system, without being limited to major distributions that NVIDIA provides primary support for. For more information, see "Installation on Community Operating Systems" section in the user manual. For a list of supported Community OSs, please see "Supported Community Operating Systems" section in the release notes.
OVS-DPDK—Partial Offload	Starting OFED 5.6, OVS-DPDK does not support partial offload.

Customer Affecting Change	Description
5.5-1.0.3.2	
Disabling RoCE While Using sysfs	When using sysfs to enable/disable roce in kernel 5.5 and up, the "devlink reload" command (using iproute2 with devlink tool) will need to be used to activate the RoCE status change. Disable RoCE example: 1. echo 0 > /sys/bus/pci/devices/0000:08:00.0/roce_enable 2. devlink dev reload pci/0000:08:00.0
mlnx-ofa_kernel	The source code for mlnx-ofa_kernel is no longer installed by default on RPM-based distributions (e.g., RHEL and SLES). **Notes:**

Customer Affecting Change	Description
Installation	• mlnx-ofa_kernel is included in the < > in the MLNX_OFED distributions under RPMS/ and may be manually installed from there. • There is no change for deb-based distributions (Debian and Ubuntu). The full source is included, as before, in the package mlnx-ofed-kernel-dkms.
Software Encapsulation Compatibility	There is an encapL2 compatibility issue with accelerated reformat action creation using mlx5dv_dr API. Using OFED 5.4 with firmware xx.32.1xxx and above or using OFED 5.5 with firmware lower than xx.32.1xxx will not allow accelerated reformat action. (Using OFED 5.4 and 5.5 with bundle firmware works properly.)
xpmem in RHEL8	Added xpmem packages in RHEL8 builds.
Python3	Starting OVS DPDK 2.15, the Python minimum required version is 3 and OVS-DPDK will not be compiled using Python 2.