Linux Kernel Upstream Release Notes v6.19

Table of contents

Overview

Supported Uplinks to Servers

These release notes provide information on new features and supported NICs in Linux Kernel Upstream.

Uplink/Adapter Card	Driver Name	Uplink Speed
ConnectX-8	mlx5	• InfiniBand: EDR, HDR100 [1], HDR [1], NDR [1], NDR200 [1], XDR [1] • Ethernet: 50GbE [2], 100GbE [2], 200GbE [1], 400GbE [1]
BlueField-3		• InfiniBand: EDR, HDR100, HDR, NDR200 [1], NDR [1] • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE [2], 100GbE [2], 200GbE [1], 400GbE [1]
BlueField-2		• InfiniBand: SDR, FDR, EDR, HDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE [2], 100GbE [2]
BlueField		• InfiniBand: SDR, QDR, FDR, FDR10, EDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 100GbE
ConnectX-7		• InfiniBand: EDR, HDR100, HDR, NDR200, NDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE [2], 100GbE [2], 200GbE [1], 400GbE
ConnectX-6 Lx		• Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE [2]
ConnectX-6 Dx		• Ethernet: 10GbE, 25GbE, 40GbE, 50GbE [2], 100GbE [2], 200GbE [2]
ConnectX-6		• InfiniBand: SDR, FDR, EDR, HDR • Ethernet: 10GbE, 25GbE, 40GbE, 50GbE [2], 100GbE [2], 200GbE [2]
ConnectX-5/ConnectX-5 Ex		• InfiniBand: SDR, QDR, FDR, FDR10, EDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 100GbE
ConnectX-4 Lx		• Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE

Uplink/Adapter Card	Driver Name	Uplink Speed
ConnectX-4		• InfiniBand: SDR, QDR, FDR, FDR10, EDR • Ethernet: 1GbE, 10GbE, 25GbE, 40GbE, 50GbE, 56GbE [3], 100GbE

1. Speed that supports PAM4 mode only. 🔲 🔲 🔲 🔲 🔲 🔲 🔲 🔲 🔲 🔲 🔲

2. Speed that supports both NRZ and PAM4 modes in Force mode and Auto-Negotiation mode. 🔲 🔲 🔲 🔲 🔲 🔲 🔲 🔲 🔲 🔲 🔲 🔲 🔲 🔲

3. 56GbE is an NVIDIA proprietary link speed and can be achieved while connecting an NVIDIA adapter card to NVIDIA SX10XX switch series or when connecting an NVIDIA adapter card to another NVIDIA adapter card. 🔲

Supported HCAs Firmware Versions

Linux Kernel Upstream supports the following NVIDIA network adapter cards firmware versions:

Adapter Card	Recommended Firmware Version
ConnectX-8	40.44.0208
ConnectX-7	28.42.1000
ConnectX-6 Lx	26.42.1000
ConnectX-6 Dx	22.42.1000
ConnectX-6	20.42.1000
BlueField-3	32.42.1000
BlueField-2	24.42.1000
ConnectX-5	16.35.1012
ConnectX-4 Lx	14.32.1900
ConnectX-4	12.28.2302

Changes and New Features

Upstream Kernel 6.19				
Internal Ref.	Feature	Description	Support Added in User Space Version	Support Added in Firmware Version
RDMA				
4624966	Increase Number of Concurrent STs	**[ConnectX-8]** Added support for direct ST mode, where ST Table Location equals `PCI_TPH_LOC_NONE`. In that case, no steering table exists, the steering tag itself will be used directly by the software, firmware and hardware from the mkey. In that mode of work, the driver is no longer limited to the 64 max entries of the capability config space table, and 256 concurrent STs can be used.	N/A	N/A
3770283	Support for PCI Peer-to-peer Functionality over VFIO	**[All HCAs]** Added support for PCI peer-to-peer functionality over VFIO. Extended: • DMA API - to work with physical addresses instead of struct page • Block layer - to reuse newly introduced infrastructure architectures Updated multiple platform architectures in the Kernel to be based on the physical addresses and rewrote NVMe driver.	N/A	N/A

Changes and New Features History

Upstream Kernel 6.18		
Internal Ref.	Feature	Description
Core		
4314703	Support LAG Net Namespace	**[All HCAs]** Added LAG support for net namespace isolation. hardware LAG forms on between mlx5 interfaces that share the same network namespace. This avoids cross-namespace interference and matches user expectations when devices are placed in different netns. For example: creating a LAG when a user is passing an entire NIC PFs to a container.
4279288	NIC Availability Monitoring Counters in Devlink Interface	**[ConnectX-8 and above]** Added support for three new counters: • `bar_uar_access` • `odp_local_triggered_page_fault` • `odp_remote_triggered_page_faul` The counters are exposed to the user through the `devlink vnic_reporter` interface. This change allows users to monitor per function usage of certain resources that may be involved in a DoS attack.
Net Dev		

Upstream Kernel 6.18		
4533359	RS-FEC Histogram Statics	**[ConnectX-7]** Added support for reporting R! FEC histogram statistics through ethtool. These histograms are defined by IEEE 802.3 and present the distribution of corrected codewords based on the number of symbol errors corrected per block, helping diagnose link quality and detect early signal degradation.
4373543	Doorbells Number	**[All HCAs]** Increased the number of doorbells from 1 to 8, assigned in a round-robin fashior to channel send and receive queues. This increases scalability on systems with many cores and high packet rates by reducing MMIO-contention for doorbell addresses.
4322159	PTP	**[ConnectX-5 and above]** Added support in th PTP subsystem and driver to expose the raw free-running cycle counter of PTP hardware clocks. This allows user space to retrieve (cycl system-time) pairs through new ioctls and correlate hardware cycle timestamps with ho! time without requiring PHC synchronization.
4318230	Support PCIe Congestion Events	**[ConnectX-7 and ConnectX-8]** Added suppor for receiving and exposing PCIe congestion events as ethtool counters. PCIe congestion events are events generated by the firmware when the device side has sustained PCIe inbound or outbound traffic above certain thresholds.
RDMA		
4444316	Flow Steering Rules	**[All HCAs]** Added support for creating flow steering rules using bulk counter objects.
3671725	IB Service Record Resolution	**[All HCAs]** Added support for APIs to resolve and query destination IB address information from IB service name or ID.
Security		

Upstream Kernel 6.18		
4088978	Support RSS for IPSec Offload	**[ConnectX-6Dx and above]** Added support fo identifying the inner protocol of decrypted packets in firmware. Added new flow groups and steering rules to redirect them for proper L4-based RSS. This ensures traffic is spread across multiple CPU cores.
4034941	PSP Security Protocol	**[ConnectX-7 and above]** Added PSP Security Protocol support to the core, including implementation in the mlx5 driver.

Upstream Kernel 6.17

Internal Ref.	Feature	Description
Net Dev		
4370178	ETS Rate Limit for VFs Group	**[ConnectX-6 Dx and above]** Added support for specifying bandwidth proportions betwee traffic classes (TC) in the devlink-rate API. Thi option allows users to allocate bandwidth across multiple traffic classes in a single command. This feature provides a more granular control over traffic management, especially for scenarios requiring Enhanced Transmission Selection. Users can now define a specific relative bandwidth share for each traffic class, such as allocating 20% for TC0 (TCP/UDP) and 80% for TC5 (RoCE).
3358561	TCP Zero Copy	**[ConnectX-7 and above]** Exposed Header-Data Split to the TCP layer to support TCP Zero Copy. This allows headers up to and including TCP to be received by the software, while packet data to be received directly by th GPU via DMA from the NIC, without the CPU's intervention, achieving true zero-copy.
RDMA		

Upstream Kernel 6.18		
4372311	Network Namespace	**[ConnectX-6 and above]** Added support for setting a device's network namespace upon allocation. The name space is inherited from its devlink instance following a devlink reload command. The IPoIB driver is now namespace-aware, creating its interfaces in the same namespace as the parent RDMA device.
4220319	TPH and Steering Tags Support	**[ConnectX-8]** Added support for TPH and steering tags, which allows the device to provide hints about where an MR TLPs should be processed in system memory.
4063646	DMABUF Optimization	**[ConnectX-4 and above]** Added support for optimizing the page size used for DMABUF memory keys (mkeys) in the mlx5 driver. The goal is to enhance performance by selecting the most appropriate page size based on the underlying system configuration and workload requirements, which can lead to improved memory management and BW.
4032404	Flow Control for Solicited MADs	**[All HCAs]** Added support for flow control mechanism for non-user solicited MADs to prevent congestion and excessive retries in the MAD QP layer. This limits the number of outstanding solicited MAD requests an agent can issue, holding excess requests in a backlog until the MAD QP is ready to process more.

Upstream Kernel 6.16		
Internal Ref.	Feature	Description

Net Dev

Upstream Kernel 6.18		
3383458	Hardware-Managed Flow Steering	**[ConnectX-6 Dx and above]** Added support for Hardware-Managed Flow Steering (HMFS which is a new approach for managing steering rules where STEs are written to ICM by hardware (as opposed to software in software-managed steering), which allows higher rate of rule insertion.

Upstream Kernel 6.15

Internal Ref.	Feature	Description
Core		
4017792	VFIO-VIRTIO Live-Migration Support for Block Devices	**[BlueField-3]** Added live-migration support fc VIRTIO block devices over VFIO.
3820686	Trust Lockdown Health Syndrome	**[ConnectX-8]** Added a new health syndrome for trust lockdown and exposed crr bit in the health buffer. When crr bit is set, the syndrome requires cold reset to recover.
3400839	User-Space Firmware Access via FWctl	**[ConnectX-5 and onward]** Added support for FWctl, which is a new subsystem designed to establish a common framework for securely exposing firmware interfaces to user space. It aims to provide consistent access for debugging, configuration, and provisioning of devices. This also adds support for mlx5, enabling user-space access to debug and configuration registers of the ConnectX hardware family, including NICs, PCI switches, and SmartNIC SoCs.
Net Dev		

Upstream Kernel 6.18		
4122077	Group of groups	**[ConnectX-6 Dx and above]** Added support for a group of groups, which enables multi-tie resource control, where parent groups enforc aggregate limits while child groups dynamical share bandwidth. This structure is crucial for QoS enforcement burst management, and fair bandwidth distribution, ensuring that workloads receive adequate resources.
3894595	Support PTP Device	**[ConnectX-6 Dx]** Added support for one PTP device per hardware clock. The PTP device represents the shared PHC when it is running in real time mode.
3358661	Lanes number	**[All HCAs]** Added support for specifying the number of lanes (protocol width) when auto-negotiation is off. As the combination of spee and number of lanes corresponds to one link mode, this enables the user to select exactly one link node when auto negotiation is off. Th feature follows a firmware update that expec the driver to specify a single link mode when auto-negotiation is off, rather than a bitmap corresponding to the user-defined speed.
RDMA		
3955559	Log IB State Transitions	**[All HCAs]** Added a feature that enhances th visibility for IB device state transitions by adding log messages to the Kernel log (dmesg). Every time an IB device changes state, a relevant print will be printed. Example: "mlx5_0: Port: 1 Link DOWN"

Upstream Kernel 6.18		
3806138	Optional Counters Binding Support	**[All HCAs]** Add support for a new optional filter named optional-counter to the iproute2 command: `rdma stat qp set link [link_name] auto` . It indicates that when binding counters to a QP we also want the currently enabled optional-counters on the link to be bound as well. In addition manual QP binding in iproute now also binds the currently configured optional-counters for the said counter. This feature allows tracking optional counters on a per (QP)/(QP group) granularity.
Upstream Kernel 6.14		
Internal Ref.	Feature	Description
Core		
4034927	ConnectX-7 Multihost+LAG Support for Tencent New Server Platform	**[ConnectX-7]** Added a feature that enables each host to get partial PFs and to own its unique LAG u nder multi-host deployment. All the lags are independent.
Upstream Kernel 6.13		
Internal Ref.	Feature	Description
Core		
3150054	Support Live Migration for Virtio-net VFs Devices	**[BlueField-3]** Added a feature that enhances the vfio-virtio driver to support live migration for virtio-net Virtual Functions (VFs) that are migration-capable. It follows the Virtio 1.4 specification to implement the necessary device parts commands, enabling a device to participate in the live migration process.
RDMA		

Upstream Kernel 6.18

Internal Ref.	Feature	Description
4069604	Devx Qps Parallel Cleanup	**[All HCAs]** Added a feature that considerably reduces cleanup (kill) time of userspace application that use DevX QPs without the user having to change anything on their end. When using a kernel that includes this patch automatically, DevX QPs destruction upon cleanup will be sent to the firmware in parallel instead of in serial (how it was before), which result in a faster cleanup time. The more DevX QPs the application has the more noticeable it would be.
3816574	Out-Of-Order RX WQE Consumption Support	**[ConnectX-8 and above]** Added a feature that allows WRs on the receiver side of the QP to be consumed OOO. It permits the sender side to transmit messages without guaranteeing arrival order on the receiver side.

Upstream Kernel 6.12

Internal Ref.	Feature	Description

NetDev

3612463	Precision Time Measurement Support on ConnectX	**[ConnectX-7]** Added support for Precision Time Measurement (PTM), that enhances clock synchronization accuracy between the NIC's hardware clock and the system's host clock. PTM resolves inaccuracies caused by asymmetric PCIe request and response times which affect the PTP. PTM ensures more accurate synchronization, which is important for time sensitive networking.

RDMA

3554675	ODP Memory Scheme	**[ConnectX-7 and above]** Added support for memory scheme ODP. This allows the usage of ODP on DevX QPs by passing the responsibility of parsing the WQEs and MRs of the page fault to the hardware.

VDPA

Upstream Kernel 6.18		
3799752	Optimize mkey Creation Flow	**[ConnectX-6DX, ConnectX-7]** Added support for optimization for VDPA memory mapping operations. This feature reduces live migration downtime.
3799751	Suspend and Resume Hardware VQs in Parallel	**[ConnectX-6DX, ConnectX-7]** Added support for parallelize VQ suspend and resume operations. This feature reduces live migration downtime.

Upstream Kernel 6.11		
Internal Ref.	Feature	Description
Core		
3832246	Show SF Device IRQs	**[ConnectX-5 and above that support SFs, AFAIR]** Added IRQs directory inside SF SYSFS directory. This allows users to know the mapping of SFs and their IRQs.
3807493	Configurable SF EQs	**[BlueField-2 and above and ConnectX-6 and above]** Added support for setting a maximum number of completion EQs for SFs. It allows the user to set the number of queues for an SF.
NetDev		
3639446	Support Lightweight EQ	**[BlueField-3]** Page EQ is not created, and no BOOT/INIT pages are allocated for non-page supplier functions. Optimize RDMA driver initialization by delaying QPs and the corresponding CQs creation. Re-implement write combining test with SQ, and optimize for SF by reading the result from its parent PF.
3612469	Support Hardware GRO	**[ConnectX-7]** Added support for Hardware GRO. Hardware GRO reduces CPU overhead and improves throughput.
RDMA		

Upstream Kernel 6.18		
3910418	Expose req_transport_retries_exceeded Counter	**[All HCAs]** Added support for exposing req_transport_retries_exceeded counter (per QP counter). It can be seen using the RDMA statistics and it shows the number of times the requester detected transport retries exceed error.
3845636	Support Up to 23 Bits for uar_page_index in create_cq Flow	**[ConnectX-7]** Added support for allowing users to perform create_cq with a uar_page_index larger than 2^16 as supporte by the HW/FW.
VDPA		
3861923	Pre-create Hardware VQs at Device Creation Time	**[ConnectX-6 Dx, ConnectX-7, BlueField-3]** Added support for pre-creating the VQs at device creation time. This improves the downtime for the Live Migration process.

Upstream Kernel 6.10

Internal Ref.	Feature	Description
NetDev		
3787035	Support Matching on l4_type for ttc_table	**[BlueField3]** Added a feature that enables sending CC IFA2.0 prob packets on the same LAG ports as functional packets.
3602814	Monitor VFs/SFs	**[ConnectX-6 DX and above]** Added a feature that exposes the out of buffer counters on th VF/SF to the VF/SF representor, to improve monitoring on the host side.
RDMA		
3693622	Limit UMAD Receive List	**[All HCAs]** Added a feature that limits the amount of MAD packets the kernel holds for the UMAD clients. The driver would silently drop newly received packets when the list is full, thus, limiting the effects of a faulty or malicious user overflowing the network with MADs.

Upstream Kernel 6.18

Upstream Kernel 6.9

Internal Ref.	Feature	Description

NetDev

Internal Ref.	Feature	Description
3762046	SyncE Userspace Support through Linux Kernel DPLL Subsystem	**[All HCAs that support HCA.synce_registers** Added support to use Linux kernel DPLL subsystem as a mechanism for working with clock signals in NVIDIA's proprietary synchronous ethernet protocol daemon. This new mechanism enables the use of VFs and SFs.
3412918	Socket-Direct multi-PF Netdev	**[ConnectX-6 DX and above with socket direc support]** Added support to create single netdev abstraction for multiple socket-direct PFs of the same port that run on multiple NUMA nodes. This single netdev handles traffic locally on each of the NUMA nodes, without upper layer complexities. This improves the applications' locality and performance in a multi-NUMA system.

Upstream Kernel 6.8

Internal Ref.	Feature	Description

Core Features

Internal Ref.	Feature	Description
3403931	Transitional Virtio-net Support Over vfio	**[BlueField3]** Introduced a vfio driver over virti devices to support the legacy interface functionality for the VFs.

VDPA Features

Upstream Kernel 6.18		
3016493	Improve SW Assisted Live Migration Downtime	**[ConnectX-6 DX and above]** Added a feature that improves the downtime during live migration with qemu and kernel vdpa (mlx5_vdpa). Downtime is the period when traffic cannot work due to queue reconfiguration and migration. It is improved through the following changes: • During live migration, allowed re-mapping only of the virtqueue descriptor area instead of the whole VM memory which takes a long time • Introduced a virtqueue resume option to avoid virtqueue destruction and re-creation cycle which takes a long time

Upstream Kernel 6.7

Internal Ref.	Feature	Description

ASAP Features

| 3562445 | Aggregated Port Speed During Rate Configuration | **[ConnectX-5 Dx and above]** Added support for functions (PF,VF,SF) to utilize the maximum aggregated link speed for transmission when LAG is configured.
In this case, the aggregated link speed of both ports, rather than the max link speed of single one, is considered when using devlink rate and tc-matchall to limit the transmission speed. |

Core Features

| 3414893 | SF Creation Time Optimization | **[ConnectX-4 and above]** Added a feature that reduces SF creation time from 2.4s to < 500ms. SF creation time per device was 2.4s before, which is unacceptable for large scale applications (1k ~ 2k SFs). |

Upstream Kernel 6.18		
3375076	Synchronous SF activation	**[ConnectX-5 and above]** Added devlink-nested to show relationships between devlink instances through devlink dev show/devlink port show commands. Devlink SF activate command returns before finished to enable partial parallelization of SFs activation and reduce time. Devlink port show can be used to see when SFs are activated. In SFs case, it shows the relationship between SF auxiliary device devlink instance and the E-switch representor devlink port.
3283517	Image Live Migration	**[ConnectX-7]** Added a feature that enables live migration support for image size above 4GB. This is achieved by reading/writing the device state from/to the FW in a chunk mode (i.e. small part per read/write) instead of reading/writing the full image at once.
NetDev Features		
3358686	Increase netdev Interface max num Channels	**[ConnectX-4 and above]** Added a feature that bumps the software upper bound of the netdev max channels from 128 to 256. This is relevant for systems with a high number of cores.
Steering Features		
3416340	Multicast bridges offload	**[ConnectX-6 DX and above]** Added the support for multicast bridges' offload. Multicast packets will be offloaded by the HW and will not go up to the kernel stuck. The feature is enabled according to FW support.

Upstream Kernel 6.6				
Internal Ref.	Feature	Description	**Support Added in User Space Version**	**Support Added in Firmware Version**
Core Features				

Upstream Kernel 6.6				
3415953	Dynamic EQ allocation	Added a feature that modifies driver completion EQ allocation method to be dynamic (on-demand), instead of a single static allocation of all completion EQs on driver load. Meaning, only when a completion EQ is needed, the driver will allocate it, reducing the memory overhead of driver load.	N/A	N/A
3233734	VFIO P2P migration support in QEMU	**[ConnectX-6 Dx and above]** Added a mechanism that safely starts and stops multiple VFIO devices that might do P2P transactions between themselves. Such mechanism didn't exist before, and thus, VFIO migration was allowed only for VMs with a single VFIO device. With this mechanism, this restriction is lifted and VMs with multiple VFIO devices can be migrated, given that all devices support P2P VFIO migration.	N/A	N/A
2982377	Live migration support with IPsec full offload	**[ConnectX-7 and above]** Added the option to do a live migration on machine with IPsec enabled, meaning we can have a VF on VM that has a valid IPsec state and migrate it without having to remove the IPsec.	N/A	28.38.1002
2461105	Expose max SF configure via devlink resource	**[ConnectX-4 and above]** Added a feature to allow user applications to view the maximum number of SFs configured on the system via general linux API, devlink resource.	N/A	N/A
NetDev Features				

Upstream Kernel 6.6				
3037972	Expose NIC temperature via hardware monitoring kernel API	**[ConnectX-5 and above]** Added a feature that exposes NIC temperature by implementing hwmon kernel API (which are the standards that the drivers follow, so that an application program can scan for entries and access thermal sensors data in a simple and consistent way via libsensors userspace library). For each one of the supported and exposed thermal diode sensors, this feature exposes the following attributes:1) Input temperature: the current measured temperature.2) Highest temperature: historically, the highest temperature that have been measured.3) Temperature label (Or name): depends on the firmware capability, if firmware doesn't support sensors naming, the fallback naming convention would be: "sensorX", where X is the HW spec (MTMP register) sensor index.4) Temperature critical max value: refers to the high threshold of Warning Event. Will be exposed as `tempY_crit` hwmon attribute (RO attribute). For example, for ConnectX5 HCA's, this temperature value will be 105 Celsius, 10 degrees lower than the HW shutdown temperature.5) Temperature reset history: resets the highest temperature value.	N/A	xx.39.1002 GA Rel (rel-12_39_0054)
RDMA Features				
2844159	Baremetal MACsec offload	**[ConnectX-7 and above]** Added support for MACsec operations' offload to the HW, which yields greater performance than software counterpart, and most importantly enables sending MACsec encrypted RoCE traffic, which isn't possible in software MACsec.	N/A	28.35.0328+

Upstream Kernel 6.6				
Security Features				
3183351	Support IPSec packet offload in switchdev mode	**[ConnectX-6 Dx and above]** Added support for IPSec packet offload in switchdev mode.	N/A	N/A
3000605	Configuration from VM/VF for full offload transport mode	**[ConnectX-6 Dx and above]** Introduced two new Boolean attributes of a port function: ipsec_crypto and ipsec_packet. The goal is to provide a level of granularity for controlling VF IPsec offload capabilities, similar to what is currentlyoffered in the software mode. This allows users to decide if they want bothtypes of IPsec offload enabled for a VF, just one of them, or none at all(which is the default). A hypervisor-level control parameters to set the IPsec functionality of PCI VF devices passed through to guests. The administrator of a hypervisor host maychoose to change the settings of a port function from the defaults configuredby the device firmware.	iproute2 v6.6	20.37.1014
Steering Features				
3375973	Forwarding INT packets	**[Connect-6 DX, Connect-7, BlueField-2, BlueField-3]** Added support for matching over BTH acknowledge bit using the mlx5dv_dr API.	N/A	N/A

Upstream Kernel 6.5				
Internal Ref.	Feature	Description	**Support Added in User Space Version**	**Support Added in Firmware Version**
ASAP Features				
3376168	Bridge Debuggability Extensions: Expose FDB Data via debugfs	[ConnectX-6 Dx and above] This new capability improves the bridge's offload debuggability.	N/A	N/A
Core Features				
3427627	Full Chip Reset on BlueField-2 and BlueField-3 DPUs	[BlueField-2 and BlueField-3] Added support for full chip reset in DPU mode on BlueField-2 and BlueField-3 DPUs, using mlxfwreset the "--sync 1" option. During this flow, Arm is going through reboot and firmware is reloaded.	N/A	xx.38.1002
3405789	Embedded CPU Virtual Functions	[BlueField-2 and BlueField-3] Enabled the creation of Virtual Functions within the Arm. **Note:** This capability requires setting the following parameters: • In mlxconfig: PF_NUM_OF_VF_VALID=True • For each Arm and Host PF: PF_NUM_OF_VF	N/A	xx.36.1010
3307360	QEMU VFIO Migration pre-copy	[ConnectX-6 Dx and above] VFIO migration pre-copy support extends basic VFIO migration by allowing the device state's migration while the VM and the device are running. This new capability reduces migration downtime, especially for devices that use a lot of resources.	QEMU 8.1	xx.37.1014

Upstream Kernel 6.5				
30011500	Light Weight Local SFs	[ConnectX-5 and above] Probing local SFs with devlink instance only. The local SFs-SFs spawns over the device (PF/ECPF) which is the eSwitch manager. This new capability decreases the amount of time needed to probe and configure SFs by saving the time required for the devlink to reload the SF.	N/A	N/A
2831943	4 Ports VF LAG	[ConnectX-7] Enabled VF LAG over 4 ports HCAs. **Note:** This capability is supported ONLY in LAGs that included all HCA ports, e.g.: with 4 port HCA, only 4 port LAG is supported. 2 ports or 3 ports LAG is not supported.	N/A	xx.36.1010
NetDev Features				
3320236	Indication for Packet Drops due to Severe Steering Errors	[All HCAs] This new capability exposes the generated_pkt_steering_fail and handled_pkt_steering_fail counters through the devlink health reporter to provide the user an indication on any severe Steering errors	N/A	N/A
RDMA Features				
3376634	RDMA, Static Rate	This new capability reduces the time consumption of "resolve route" during a large number of rdma_cm connections establishment by setting the rdma_cm RoCE static rate to 0.	N/A	N/A
3355352	QKEY Mitigation in Kernel	[All HCAs] Non-privileged users are now blocked by default from setting controlled/privileged QKEYs (QKEY with MSB set).	N/A	xx.38.1002

Upstream Kernel 6.5

3113659	Expose DC MRA Caps	[All HCAs] The current system behavior reports only the value for the RC QPs, thus the users running DC need to query this value and negotiate server-client with this. Up until now, the value for both RC and DC was set to 16 which might have resulted in performance issues. To avoid performance issues, we changed the max outstanding DC atomic reads to 32, and have different RC and DC values.	N/A	N/A

Upstream Kernel 6.4

Internal Ref.	Feature	Description	Support Added in User Space Version	Support Added in Firmware Version
ASAP^2 Features				
3412747	Offloaded Connections	[ConnectX-6 Dx and above] Added support for deleting offloaded connections by running `conntrack -D`.	N/A	N/A
3309368	VxLAN Group Based Policy (GBP) Offload	[ConnectX-6 Dx and above] Added support for hardware offload of VxLAN with GBP configured.	N/A	N/A
3204401	Representor Inside Namespace	[All HCAs] Added support for using switchdev mode for a devlink device within a net namespace. This includes full usage of uplink and VF representors within the net namespace of the devlink device, enabling isolated routing.	N/A	N/A

Upstream Kernel 6.4				
3180556	OVS Offload with MACVLAN Interface Above Bond	[ConnectX-5 and above] Added support for OVS offload when MACVLAN interface above bond is attached to OVS bridge.	N/A	N/A
3180548	CT with Header Rewrite	[BlueField-2] Added support for offloading CT rules with header rewrite of L3.	N/A	N/A
3158449	CT with NAT and Mirroring	[BlueField-2] Added support for offloading CT rules with NAT and mirroring.	N/A	N/A
3119303	mlx5_vdpa: Mergeable Buffer Support	[ConnectX-6 Dx and above] Added support for Enabled Mergeable Buffer feature on vdpa interfaces using vdpa tool to achieve better performance with large MTUs.	iproute v6.3	xx.37.1014
2211199	mlx5_vdpa: Posted Interrupts	[ConnectX-6 Dx and above] Added support for posted interrupts in mlx5_vdpa, allowing direct IRQ propagation from the NIC to vCPU within the guest.	iproute v6.3	xx.37.1014
2700002	Linux Bridge Multicast Offload	[ConnectX-6 Dx and above] Added support for multicast control path packet snooping (IGMP/MLD) and offload of MDB notifications in order to replicate multicast traffic to multiple destinations in the hardware.	N/A	xx.37.1014
Core Features				
3196184	Relaxed Ordering in VFs	[ConnectX-6 and above] Added support for using Relaxed Ordering in VFs directly and in VFs assigned to QEMU. Relaxed Ordering can significantly improve performance on certain setups and, until now, it could be used only in PFs.	N/A	xx.37.1014

Upstream Kernel 6.4				
2982389	QEMU Support for Device Dirty Page Tracking	[ConnectX-7 and above] Added support for QEMU to do device dirty page tracking during migration. This allows QEMU to use this capability to mark only the pages that were actually dirtied by the device as dirty instead of the entire memory. This significantly improves migration downtime and makes migration feasible. Note that device dirty page tracking with vIOMMU is not yet supported.3196184	QEMU v8.0	xx.36.1010
NetDev Features				
3193125	Hairpin Queues Configuration	[ConnectX-5 and above] Added the ability to configure the number and size of hairpin queues through devlink param command. For example: <pre>devlink dev param set pci/0000:08:00.0 name hairpin_queue_size value 512 cmode driverinit</pre><pre>devlink dev param set pci/0000:08:00.0 name hairpin_num_queues value 1 cmode driverinit</pre>	N/A	N/A
3015462	mlx5 driver Page Management via Kernel's Page Pool	[ConnectX-4 and above] Removed the internal Rx path from the internal page cache to always use the kernel's page_pool for improved performance.	N/A	N/A

Upstream Kernel 6.4				
2938348	XDP Enhancements	[ConnectX-5 and above] Implemented the following modifications in XDP in order to enhance performance: 1. Added XDP multi-buffer support to the default RQ type (Striding RQ). 2. Added support for XDP multi-buffer redirect-in. 3. Allowed non-linear single-segment frames in XDP TX MPWQE.	N/A	N/A
RDMA Features				
3259484	Monitoring VF RoCE Statistics on the Host Side	[All HCAs] Added support for the host system to expose and utilize a VF representor to monitor and track the same counters as the VF itself. This allows the host system to keep track of various statistics associated with the VF, even after it has been allocated to a VM.	N/A	xx.37.1014
2320718	Selective Repeat Protocol	[ConnectX-7 and above] Added support for Selective Repeat (SR) protocol to be enabled by default. The SR protocol retransmits only the specific damaged frame, rather than resending all the previously sent frames. This selective approach results in more efficient utilization of network bandwidth.	N/A	xx.37.1014
Security				
3004432	MACsec Support for VLAN with Inner and Outer Headers	[ConnectX-7 and above] Added support for MACsec configuration on top of a VLAN with both an inner and outer headers.	N/A	xx.34.1002
Steering				

Upstream Kernel 6.4

3325721	IB BTH Matching in RoCE Packets	[ConnectX-6 Dx and above] Added support for matching the IB BTH in RoCE packets. This enables users to perform monitoring based on specific values within the BTH field, such as monitoring R oCEv2 CNP by matching BTH opcode 0x81.	rdma-core	N/A
2205563	Enhanced modify_header Firewall Object Design	[ConnectX-6 Dx and above] Updated the design of the modify_header firewall object to incorporate patterns and arguments. This enhancement allows for the handling of modify_header actions in a larger number of flows by utilizing a single pattern with different arguments across various flows.	N/A	N/A

Upstream Kernel 6.3

Internal Ref.	Feature	Description	Suppo Added in Use Space Versio
ASAP^2 Features			
3206721	Multiport E-Switch	[ConnectX-6 Dx] Added support for Multiport E-Switch, a mode where a single E-Switch connects all VPorts and physical ports on the NIC. This allows for scenarios such as sending traffic from a VF created on PF0 to an uplink that is natively associated with the uplink of PF1.	N/A
3179772	CT UDP Unidirectional Traffic Offload	[All HCAs] Added support for offloading long-running unidirectional UDP connection with conntrack.	N/A
2825303	TC Police MTU Match	[ConnectX-6 Dx and above] Added support for offloading TC police action MTU attribute, allowing users to compare the pkt len to the MTU value. In addition, support was added for offloading TC police jump conform-exceed control action. Note: supported only in software steering mode	N/A

Upstream Kernel 6.3			
2681185	Offloading of GRE and Geneve Tunnels for VF Source IPs	[ConnectX-6 Dx and above] Added support for offloading of GRE and Geneve tunnels when the source IP address for the tunnel resides on a VF.	N/A
Core Features			
3149902	PCC fwtrace	[ConnectX-6 Dx and BlueField-2] Added support for installing a special user image into the firmware, which can be burned into either of two available slots for such applications, which enables monitoring the image's activities using the fw_tracer located inside the mlx5 driver. To view the output of the tracer, the user can access trace point, but it is important to note that they can only view traces that are generated after enabling the trace point by using the following command: 'echo 1 > sys/kernel/debug/tracing/events/mlx5/mlx5_fw/enable'."	N/A
2863362	QEMU Support for Live Migration	[ConnectX-6 Dx and above] We have added support for enabling migration of a QEMU VM with an assigned VF from one source host to another destination host. This is achieved as part of the general QEMU migration flow, which involves suspending the VF on the source host, transferring all its data to the destination host, and resuming the VF on the destination. It's important to note that this migration feature includes only basic functionality, and does not yet support advanced features such as dirty page tracking or pre-copy.	QEMU 8.0
NetDev Features			
3015464	Improve Affinity Hints According to NUMA Distances	[All HCAs] Updated the binary NUMA preference for our system to consider actual distances, so that remote NUMA nodes with shorter distances are preferred over farther ones, rather than relying solely on local/remote distinctions.	N/A
3015086	PTP Assistant SyncE Holdover	[ConnectX-6 Dx and above] Enhanced the PTP holdover performance by integrating the nullf servo with the SyncE daemon. Furthermore, we ensured that the PTP holdover is fully supported through SyncE alone, in order to meet our customers' expectations.	linuxpt
RDMA Features			

Upstream Kernel 6.3

3260484	CC - RTT response SL (CNP SL)	[ConnectX-6 Dx and above] Users can now customize the DSCP value of RTT Response packets (in Ethernet, using debugfs). This feature allows for prioritization of RTT Response packets, preventing any delay that might lead to incorrect congestion assumptions on the RTT Requester side.	N/A
3019875	IPsec Full Offload RoCE Support (also via an SR-IOV VF)	[ConnectX-7 and above] Added support for configuring IPsec even when all IPsec operations are completely offloaded to hardware. This feature not only provides a significant performance improvement but also enables the use of IPsec over RoCE packets, which are outside the network stack and cannot be used without full hardware offload. As a result, users can now leverage the benefits of IPsec protocol with RoCE V2, even when using SR-IOV VFs.	iproute
3005605	Fatal QP Error Logging	[All HCAs] W ith this feature, a kernel error log is now generated when certain fatal QP errors occur.	N/A
General			
2850256	Support Fast Update Encryption Key	[ConnectX-6 Dx and above] Enhanced the performance of DEK operations by introducing a DEK pool that serves the same key purpose and utilizes bulk allocation, destruction, and invalidation provided by firmware. Users can now retrieve a DEK object from the pool and update it with a key using the modify_DEK command.	N/A

Upstream Kernel 6.2

Internal Ref.	Feature	Description	Support Added in User Space Version	Support Added in Firmware Version
ASAP[2] Features				
3258291	TC Rules: Additional Actions	[ConnectX-5 and above] Added support for adding TC rules with trap action with additional actions (mirror and pedit).	N/A	N/A
Core Features				

Upstream Kernel 6.2				
3048267	Migratable Bit	[ConnectX-6 Dx and above] Added support for Migratable Bit in Live Migration. Becasue some features cannot be migrated, such as IPsec, for example, when VF is marked as migratable, those features are disabled. This feature allows the user to configure whether VF can be migrated.	iproute2-6.2.0	xx.34.1002
3012923	PRE_COPY Migration Support	[ConnectX-6 Dx and above] Added support for PRE_COPY migration. The optional PRE_COPY state opens the saving data transfer FD before reaching STOP_COPY and allows the device to dirty track the internal state changes with the general idea to reduce the volume of data transferred in the STOP_COPY stage. In short, PRE_COPY helps reduce the downtime of the VM.	QEMU	xx.35.2000
NetDev Features				
3253579	C-TAG	[ConnectX-4 and above] Fixed SR-IOV legacy to enable VST C-TAG insertion when a guest tag is already present on the packet. This change enables sending packets with two cVLAN tags, where the outer cVLAN tag is added by the eSwitch.	N/A	N/A
2645259	Zerocopy Sendfile	[ConnectX-6 Dx and above] Added support for a new optional mode of TLS sendfile(), in which the extra copy is skipped. Removing this copy improves performance significantly, as TLS and TCP sendfile perform the same operations, and the only overhead is TLS header/trailer insertion.	N/A	N/A
RDMA Features				

Upstream Kernel 6.2

3195715	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.	N/A	N/A
3020161	Open Multiple Blocks AES-XTS Support	[BlueField-3 and above] Added support for open multiple blocks of AES-XTS. Prior to BlueField-3, encrypting only a single block at a time is supported using NIC AES-XTS hardware engine. Starting BlueField-3, AES-XTS multiblock encryption is supported.	rdma-core	N/A
General				
3042126	Enhanced CQE Compression	[ConnectX-6 and above] Added support for the enhanced version of the RX CQE compression hardware feature. By compressing RX CQEs, the PCI bandwidth utilization is imporved and the load on it are reduced. The enhanced version of this device feature has improved latency and CPU utilization.	N/A	xx.30.1004
2724901	IPsec Packet Offload	[ConnectX-7 and above] IPsec packet offload is an improved version of IPsec crypto mode. In packet offload mode, the hardware is responsible to trim/add headers in addition to decrypt/encrypt. In this mode, the packet arrives to the stack as already decrypted, and vice versa, for TX (exits to hardware as not-encrypted).	N/A	N/A

Upstream Kernel 6.1

Internal Ref.	Feature	Description	Support Added in User Space Version	Support Added in Firmware Version
Core Features				

Upstream Kernel 6.1				
2982381	Live Migration Dirty Page Tracking Support in Linux	[ConnectX-7] The dirty pages tracking support enables reducing downtime upon live migration. Once it is used, only the pages that were really dirtied by the device will be marked in QEMU as dirty and will be sent to the target upon stop. Without dirty tracking, all RAM is marked dirty so all RAM is resent upon stop and the downtime is increased.	N/A	xx.35.1012
2701349	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/A	N/A
NetDev Features				
3020168	Support RSS over XSK Queues	Use default RSS functionality to spread traffic across different XSK queues instead of having to provide explicit steering rules.	N/A	N/A
3018567	kTLS Offload	Added support of kTLS offload with key size of 256 bits.	N/A	N/A
2633397	Decouple Flow Steering	[All HCAs] Added support for Decouple Flow steering into a separate module, making it loosely coupled and thus easier to read, maintain, and debug.	N/A	N/A
2329724	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.	N/A	xx.35.2000
RDMA Features				

Upstream Kernel 6.1				
3173656	RDMA Security	[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.	N/A	N/A
3019755	Floating LID (FLID)	[ConnectX-7] Added support for Floating LID (FLID) which can be used to identify a group of InfiniBand routers that allow communication with another subnet's entity. With this feature, multiple routers can be used per destination so that adaptive routing is supported. The FLID feature needs support from components such as the host, the subnet manager, the router, and more. This feature is only supported on the host portion of the sysrem.	N/A	xx.35.2000
3008739	Firmware Port Selection Based on Steering	[ConnectX-6 Dx and above] Added firmware support to allow explicit port selection based on steering and not QP affinity. Use LAG Hash Mode for the HCA with two ports, if supported. Make sure to keep port affinity function in LAG Hash Mode if it supports bypass select flow table in non-SwitchDev mode.	N/A	xx.34.1000
General				
3004428	Offloading MACsec with Replay Protection	[ConnectX-7 and above] Added driver support to allow offloading MACsec with replay protection enabled. This provides protection against replay attacks while using hardware offload.	IProute2 6.1.0	xx.34.1002
3004426	Offloading MACsec with EPN	[ConnectX-7 and above] Added driver support to allow offloading MACsec with extended packet number (EPN) enabled. This feature extends the packet number field to a 64 bit field allowing a longer lifetime of the MACsec session. FW version: xx.34.0364User space:-IProute2 version 6.1.0	IProute2 6.1.0	xx.34.1002

Upstream Kernel 6.1				
2821726	Offloading MACsec	[ConnectX-7 and above] Added driver support to allow offloading MACsec. This feature adds support for hardware offloading for MACsec protocol with encryption. Adding/removing/updating SAs/SecYs is supported. Note: Hardware offloading for MACsec protocol with no encryption/authentication is a separate feature and is not part of this feature.	IProute2 6.1.0	xx.34.1002

Unsupported Features

Internal Ref.	Keyword	Feature
2942773	Multi-Block Encryption	[ConnectX-6 Dx, BlueField-2, and above] Multi-block encryption is currently unsupported, due to a hardware limitation.

Upstream Kernel 6.0				
Internal Ref.	Feature	Description	Support Added in User Space Version	Support Added in Firmware Version
ASAP^2 Features				
3034240	MC Offload for OVS	[All HCAs] Added support for MC offload for OVS using ovs mcast_snooping_enable configuration.	N/A	N/A
2685267	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). And 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.	N/A	N/A

Upstream Kernel 6.0				
2666892	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.	N/A	N/A
2641218	Metering Offload	[ConnectX-6 Lx, ConnectX-6 Dx and above] Added support for per flow metering using OVS or TC.	Open vSwitch v3.0.0	xx.34.100
Core Features				
3015003	VNIC Diagnostics Counters	[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//vnic_diag/	N/A	N/A
3005449	DevLink Parallel Command	[ConnectX-4 and above] In this feature, a global DevLink mutex was removed, so running DevLink commands in parallel on different DevLink devices is possible. For example, burning firmware on afew cards on the same host in parallel using DevLink API is now possible.	iproute2 code	N/A
NetDev Features				
3015063	XDP in Switchdev Mode	[ConnectX-5 and ConnectX-6 Dx] Added XDP support for uplink representors in SwitchDev mode.	N/A	N/A
2914151	tx_port_ts Resiliency	[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, leading to losing the synchronization altogether. This feature allows for fast recovery and allows to quickly regain synchronization.	N/A	xx.34.1002

Upstream Kernel 6.0

2619241	Protection from Flood Attacks	[BlueField and above?] Upstream XDP program to accelerate part of synproxy that generates and checks SYN cookies, protecting from TCP SYN flood attacks. New upstream BPF helpers required for this program.	N/A	N/A
2092924	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.	N/A	N/A
RDMA Features				
3015279	Forwarding Packets Into a Level 0 Table	[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.	rdma-core v42.0	xx.34.100
2816261	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.	N/A	xx.34.100

Upstream Kernel 5.19

Internal Ref.	Feature	Description	Support Added in User Space Version	Support Added in Firmware Version
ASAP^2 Features				

Upstream Kernel 5.19				
2346181	OVS Failover Improvement	[ConnectX-6 Dx] Improved OVS failover through support for OVS groups in fast-failover mode + VF_LAG configuration with OVS.	N/A	xx.32.1010
3019886	Filtering Packets per Packet VLAN ID	[ConnectX-6 Dx and above] Added support for filtering packets per packet VLAN ID.	N/A	N/A
2211207	VDPA Vendor Statistics	[ConnectX-6 Dx and above] Added support for reading vendor statistics for a given virtqueue.	iproute2	N/A
Core Features				
2967769	Firmware Pre-Initialization Timeout Improvement	[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.	N/A	N/A
2783777	Firmware Reset in DPU NIC Mode	[BlueField-2] Added support of Firmware Reset in DPU NIC mode.	N/A	N/A
Software Steering Features				
2885073	IP Explicit Congestion Notification (ECN)	[ConnectX-5 and above] ECN field is 2 bits in the traffic control IP header. This feature allows reliable communication by notifying when congestion occurs by modify ECN field to CE(b'11). Adds the ECN field modification support for IPv4 and IPv6 packets.	DPDK 22.07	xx.33.1048

Upstream Kernel 5.18				
Internal Ref.	Feature	Description	**Support Added in User Space Version**	**Support Added in Firmware Version**
ASAP^2 Features				
2883524	Multiple VF/SF Interfaces in the Shared Block in TC Offload	[ConnectX-5 and above] Added support for creation of TC flows on shared block of VF representors.	N/A	N/A
2429749	sFlow Sampling Rules Offload	[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 the visibility into virtual servers, ensuring data center visibility and control.	Not supported	xx.29.1016
2204161	VLAN Push on RX and Pop on TX	[ConnectX-6 Dx and above] Added support in SwitchDev mode to push VLAN on RX (rules on uplink representor) and pop on TX (rules on VF representor). This feature is only supported in Software Steering mode.	N/A	N/A
Core Features				
2626743	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.	N/A	N/A
RDMA Features				

Upstream Kernel 5.18				
2634958	mlx5 VFIO CQ Interrupt Mode	[ConnectX-5 and above] Added support for mlx5 VFIO CQ interrupt mode. This feature allows applications to listen on and capture CQ complection events via the Event Queue mechanism with mlx5 VFIO driver.	rdma-core v41	N/A
2600493	mlx5 VFIO Asynchronous Event	[ConnectX-5 and above] Added support for mlx5 VFIO asynchronous event. This feature supports applications to listen on and capture device asynchronous events via the Event Queue mechanism with mlx5 VFIO driver.	rdma-core v41	N/A
Software Steering Features				
2895043	Matching Granularity Change	[ConnectX-5 and above] Added support for matching granularity change. Currently when creating FDB flow with destination of VPORT, a src_port matching must be added. This feature removes that limitation so a FDB flow can match all VPorts and go to a VPORT destination. The new behavior is the same as as that on firmware steering.	rdma-core	N/A
2855804	Full Tunnel Header Matching	[ConnectX-6 Dx and above] Added support for full tunnel header matching. Currently, the full-tunnel-header matching cannot be used along with many other criteria within one matcher. This feature is to add the support for that by using the new definer index, defined in the firmware, to build matcher so that the full tunnel header matching can be used along with all other criteria.	rdma-core	xx.32.1010

Upstream Kernel 5.17

Internal Ref.	Feature	Description	Support Added in User Space Version	Support Added in Firmware Version
ASAP² Features				
2668112	Setting VF Group Rate Limit	[ConnectX-5 and above] Added support for setting VF group rate limit using Devlink command.	iproute2 (devlink) 5.16	N/A
Core Features				
2658197	Steering of IPoIB Packets via Ethtool	[All HCAs] Enabled steering of IPoIB packets via Ethtool, in the same way it is done today for Ethernet packets.	N/A	N/A
2633768	Memory Improvements	[ConnectX-4 and above] Added support for providing knobs which enable users to minimize memory consumption of mlx5 functions (PF/VF/SF).	N/A	xx.32.1010
2339680	512 vDPA Interfaces	[ConnectX-6 Dx and BlueField-2] Added support in multiple flow steering priorities for FDB rules.	N/A	N/A
NetDev Features				
2801496	Expose FEC Counters via Standard Ethtool Interface	[ConnectX-4 and above] Exposed Ethtool counters which report FEC (forward error correction) statistics via the 'ethtool -I --show-fec ethX' command. • fc_fec_corrected_blocks_laneX • rs_fec_corrected_blocks • fc_fec_uncorrectable_blocks_laneX • rs_fec_uncorrectable_blocks • phy_corrected_bitz	ethtool v5.13	N/A

Upstream Kernel 5.17				
2800578	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.	ethtool v5.15	N/A
2482544	Reduce Memory Consumption	[ConnectX-4 and above] Reduced memory consumption. Previously, several hardware/memory resources were pre-allocated per NetDev to serve a prefixed maximum allowed value. Now, the actual maximum is used instead of the prefixed theoretical maximum and resources are allocated on demand, improving memory consumption.	N/A	N/A
Software Steering Features				
2338344	Dumping Software Steering Information	[ConnectX-5 and above] Extended mlx5 debugfs support to present Software Steering resources: dr_domain including its tables, matchers, and rules. The interface is read-only. While dump is being presented, new steering rules cannot be inserted/deleted.This allows to see the existing steering configuration and to be able to debug steering issues.	N/A	N/A

Upstream Kernel 5.16			
Internal Ref.	Feature	Description	Support Added in User Space Version
ASAP² Features			

Upstream Kernel 5.16			
2754419	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.	N/A
2666887	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.	N/A
2293534	Tunnel Rules Offload	[ConnectX-6 Dx and above] Added support for offloading tunnel rules when the source interface is OpenvSwitch bridge (internal port).	N/A
Core Features			
2713656	Single IRQ for PCI Function	[ConnectX-4 and above] Added support for single IRQ for PCI function. To use a high number of VFs, a large amount of IRQs is required which the device cannot always support. This feature enables VFs to function with a minimum of a single IRQ instead of two. This is done via dynamic MSIX feature. In the case that dynamic MSIX feature is not supported (old kernels), the following configuration will probe all VFs with single IRQ: $ mlxconfig -d s NUM_VF_MSIX=0 STRICT_VF_MSIX_NUM=1.	N/A
2655918	Steering-Based LAG	[ConnectX-6 Dx and BlueField-2] Added support to have physical port selection based on the hash function defined by the bond so that different packets of the same flow will be egress from the same physical port.	N/A
NetDev Features			
2816603	Firmware Health Asserts Enhancement	[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.	N/A

Upstream Kernel 5.16			
2300278	SHAMPO (GRO Offload)	[HCA?] Added support for SHAMPO (Split Header And Merge Payload Offload), a new functionality that enables stitching of packets data to unified skb representation in the network stack to reduce per-skb overhead.	N/A
RDMA Features			
2438531	Lossy RoCE	[ConnectX-5 and above] Extended 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. The 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-countersNote: Need ECN/PFC counter statistics based on VF	iprout2 v5.15.0
Steering UserSpace Features			

Upstream Kernel 5.16

2680924	Flow Direct to Port	[ConnectX-6 Dx] When using Socket-Direct NICs, each interface may have different PCIe IDs and even different domains where the NIC resides on multiple NUMA nodes. The current logic is that interfaces which have the same "system_image_guid" reside in the same physical NICs. "native_port_num" represents interface index. When using firmware older than xx.32.1xxx, the original PCIe BDF approach is used and can be queried at sysfs "phys_switch_id" and "phys_port_name".	mlnx_dpdk_20
2441633	MLX5DR SF	[ConnectX-5 and above] Added support for up to 512 SFs with the mlx5dv_dr API.	rdma-core v38

Upstream Kernel 5.15

Internal Ref.	Feature	Description	Support Added in User Space Version	Support Added in Firmware Version
ASAP[2] Features				
2250409	Kernel VDPA Network Devices	[ConnectX-6 Dx] Added support for virtio network drivers utilizing VDPA interfaces. This requires recent libvirtd and qemu.	quemu and libvirt >= 6.0	N/A
2193092	VF Group Throttling	[ConnectX-5 and above] Added support to allow user to apply QoS configuration of single function (VF, SF) or group of functions using iproute2 devlink command.	iproute2 v5.13.0 korg kernel v5.15-rc1	N/A
NetDev Features				

Upstream Kernel 5.15

2192365	Application Device Queues (ADQ)	[ConnectX-4 and above] Added driver-level support for Application Device Queues. It allows partition defining over the RX/TX queues into groups and isolates traffic of different applications. This mainly improves predictability and tail latency.	N/A	N/A
RDMA Features				
2565548	DV API for DMA GGA memcpy	[BlueField-2 and above] DMA memcpy is one of several Memory-to-Memory Offloads (MMO) available from BlueField-2 onwards. It utilizes the GGA modules on the DPU to perform DMA memcpy, thus improving performance. The memcpy can be done locally, on the same host, or between the host and the Arm. To use this feature, expose DV API.	rdma-core only	N/A
2452045	mlx5 over VFIO	[ConnectX-4 and above] Added support for mlx5 user space driver over VFIO. This feature enables an application to take full ownership on the opened device and run any firmware command (e.g., port up/down) safely. The application look and feel is like regular RDMA application over DEVX. It uses verbs API to open/close a device and then mostly uses DEVX APIs to interact with the device. New mlx5 DV APIs were added to get ibv_device for a given mlx5 PCI name and to manage device-specific events. For description of the relevant APIs and expected usage of those APIs, look up the following: mlx5dv_get_vfio_device_list() mlx5dv_vfio_get_events_fd() mlx5dv_vfio_process_events()	rdma-core v37	N/A
2297914	DV API for AES-XTS	[ConnectX-6 and above] Added DV API that allows configuration of MKey with AES-XTS crypto offloads. The MKey can be configured for both crypto and signature offloads.	rdma-core only	xx.31.1014

Upstream Kernel 5.15				
1711066	Single FDB	[All HCAs] Added support for bonding the uplink representors under the same bond device so that a single InfiniBand device is created which can be used to manage VPorts which are on both E-Switch managers.	N/A	N/A
Steering UserSpace Features				
2669516	Extended Flex Parser Ability	[ConnectX-5 and ConnectX-6 Dx] Added userspace ability to expose flex parsers 4-7 provided by misc4 to extend matching ability of flex parsers. Now all flex parsers can be matched at the same time.	rdma-core v38	xx.32.1xxx
2669419	Setting DR Matcher Layout	[ConnectX-6 Dx] Added support for a new RDMA CORE DR API to set the DR matcher layout by calling mlx5dv_dr_matcher_set_layout. Setting the matcher layout allows presetting the matcher size and increasing matcher rule capacity as well as other performance improvements in case matcher size is known.	rdma-core v37	N/A
2616897	Software Steering Stability	[ConnectX-6 Dx] Improved Software Steering rule creation stability so that all rule's insertion occur in a defined time using defined (export) size of Htble and decreased use of dynamic allocation.	rdma-core v37	N/A
2436999	Software Steering Rule Optimization	[ConnectX-6 Dx] Added Software Steering Rule Optimization which improved rate of updating steering rules, insertion, and deletion. The feature includes definers, multi-qp approach, and better memory usage.	rdma-core v37	N/A
2350115	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.	rdma-core v37	N/A

Upstream Kernel 5.15

Internal Ref.	Feature	Description		
2248827	mlx5dv_dr Support of Scalable Function	[All HCAs] Added mlx5dv_dr SF support. This feature allows directing traffic using mlx5dv_dr, dest_ib_port action to a SF VPort.	rdma core v37	N/A
2239853	mlx5dv_dr Support for VF LAG	[All HCAs] Feature Description: This feature allows directing traffic using mlx5dv_dr, dest_ib_port action to any port on the e-switch, including the other PF.	rdma core v37	N/A

Upstream Kernel 5.14

Internal Ref.	Feature	Description
ASAP^2 Features		
2198071	Enlarge Switchdev Tables	[ConnectX-5 and above] Added support for allov kernel to support up to 128 matches (groups) pε 16M entries per group.
2047159	Bridge Offloads with VLAN Support	[ConnectX-6 Dx and above] Added support for b offloads with VLAN support that works on top o representors in switchdev mode.
ASAP^2 and Steering Features		
2200978	sFlow Sampling Rules Offload	[ConnectX-5 and above] Added support for offlc sFlow sampling rules.
RDMA Features		
2431025	Raw WQE	[ConnectX-5 and above] Added support for Raw (mlx5dv_wr_raw_wqe). This feature allows applic; build a new custom work request (WQE) that is r supported by the verbs or driver and post it on r It is an extension for IBV work request (ibv_wr_*) specific features for sending a work request.
2356957	ah_to_qp Mapping	[ConnectX-6 Dx] Added support for mapping a (over DEVX API, which enables DC/UD QPs to usε CC algorithms in the same data center.

Upstream Kernel 5.14

2352179	Relaxed Ordering for Kernel ULPs	[ConnectX-4 and above] Added support for enal Relaxed Ordering for Kernel ULPs. Using relaxed can improve performance in some setups. Since ULPs are expected to support RO, it is enabled fι default so they can benefit from it.
2339814	Timestamp Format in QP Database	When the device works in real-time mode, and tι created with IBV_WC_EX_WITH_COMPLETION_TIMESTAMP_W the CQEs timestamps will be in a real-time form
2163186	ibv_query_qp_data_in_order() verb	[All HCAs] Added support for ibv_query_qp_data API. This API enables an application to check if tι data is guaranteed to be in order, enabling poll fι instead of poll for completion.
1748169	DV "Signature API"	[ConnectX-5 and above] Added support for "Sigι which allows application-level data-integrity che signature handover mechanism. Various signatu including CRC32 and T10-DIF, can be automaticɑ calculated and checked, stripped, or appended c transfer at full wire speed.
Steering UserSpace Features		
2446723	Duplicate Rules Insertion	[ConnectX-5 and above] Added support for abiliı or prevent insertion of duplicate rules. This enak choose between the following behaviors: 1. Prevent duplicate rules, so that already-exi and fail can be detected. 2. Allow duplicate rules, to enable updating tι action (this will only take effect once the pι is deleted). By default, duplicate rules are a
2368372	MLX5DR Packet OK and Checksum Checks	[ConnectX-6 Dx] Added support for new matchi ipv4_checksum_ok and l4_checksum_ok. l3_ok,
2350116	MLX5DR Match Definer	[ConnectX-6 Dx] Added support for match defiı are used internally in the mlx5dv_dr API. Definer filtering on more packet fields, improving the pɑ and accelerates mlx5dv_dr API.

Upstream Kernel 5.14		
2334773	Matching on RAW Tunnel Headers	[ConnectX-5 and above] Added DR support for ʳ on RAW tunnel headers using the misc5 parame feature allows matching on each bit of the head reserved fields.

Upstream Kernel 5.13		
Internal Ref.	Feature	Description
ASAP^2 Features		
2576790	Offloading Extended ct_state Flags	**[ConnectX-5 and above]** Added support to offload ct_state rpl, inv, and rel. For rpl, support was added for both set and set matching offload (i.e., +rpl and -rpl). For inv and rel, supp was added only for the not set option (i.e., -rel and -inv).
2392316	vDPA Instances on Scalable Functions	**[ConnectX-5 and above]** Allow to create vDPA instances on Scalable Functions (subfunctions) using devlink commands
2055404	Uplink Representor Modes	**[ConnectX-5, ConnectX-6, ConnectX-6 Dx and above]** Sup for new_netdev mode in SwitchDev mode. The new default behaviour is to always keep the NIC netdev.
ASAP^2 and Steering Features		
1886181	Offloading sFlow Sampling Rules	**[ConnectX-6 Dx]** Added support for offloading sFlow samp rules.
2330570	Simultaneous Activation of CQE Compression and PTP	**[ConnectX-4 and above]** Added support for the activation PTP and CQE compression simultaneously. Since CQE compression might harm the accuracy of the PTP, the featu enables PTP packets to be moved to a dedicated queue wh they are not subjected to compression. However, this configuration conflicts with setting aRFS.Turn off CQE compression, causes a hiccup in traffic which may a loss of synchronization. To overcome this, restart the synchronization.Note: this combination is supported only fo Ethernet driver. Other driver profiles, like IPoIB and represer do not support this combination.
NetDev Features		

Upstream Kernel 5.13

1821120	TLS RX Hardware Offload	[**ConnectX-6 Dx and above**] Added GA-level support for hardware offload decryption of TLS traffic over crypto-enal ConnectX-6 Dx NICs and above.
2257993	TLS RX Resynchronization Resiliency	[**ConnectX-6 Dx and above**] Added support to have driver resiliency against high load of RX resync operations.
2225882	ethtool Extended Link State in mlx5e	[**ConnectX-4 and above**] Added ethtool Extended Link Stat mlx5e. ethtool can be used to get more information to help user troubleshoot the state. For example, if there is no link a missing cable ($ ethtool eth1 … Link detected: no (No cab Besides the general extended state, drivers can pass additi information about the link state using the sub-state field (f example: $ ethtool eth1 … Link detected: no (Autoneg, No p detected)). The extended state is available only for some ca no link; in other cases, ethtool will print only "Link detected: as before.
Core Features		
2041295	Scalable Functions (Subfunctions)	[**ConnectX-5 and above**] Added support for scalable functi (also called subfunctions). The feature enables the user to c configure, and deploy a scalable functions (e.g., RDMA and networking applications) and to assign them to a container a container is started via mlxdevm tool. A scalable function can also be deployed in an untrusted guest/host system from the NIC/DPU. This enables full configuration of the function and its representors from the NIC/DPU before giving the function for a container to run ir host system. For more information, see https://github.com/Mellanox/scalablefunctions/wiki/MLNX_ step-by-step-guide.
2401360	Reading DSFP Module Information	[**ConnectX-4 and above**] Added support for reading DSFP module information. The change included adding new options to ethtool netlink EEPROM module read API, to read specific page and bank.
RDMA Features		
2404639	Export Object IDs to Users	[**ConnectX-4 and above**] Extended support for the "rdma r show" command to SRQ and context resources.

Upstream Kernel 5.13		
2360568	Huge Page Support for devx umems	**[ConnectX-4 and above]** Added support to allow DEVX UM be created with larger page sizes than 4K. This is a must for some device objects (e..g. RegEx) and, in a case, page size larger than 4K may need less MTTs which m improve performance. A new API mlx5dv_devx_umem_reg_ex() was added which requests a specific page sizes. It enables better application control on the required UMEM page size.
2335586	Dynamically Assign MSI-X Vectors Count	**[ConnectX-5, ConnectX-6, ConnectX-6 Dx and above]** The number of MSI-X vectors is PCI property visible through lsp that field is read-only and configured by the device. The static assignment of an amount of MSI-X vectors does allow utilize the newly created VF because it is not known to device the future load and configuration where that VF will used.The VFs are created on the hypervisor and forwarded VMs that have different properties (for example number of CPUs).To overcome the inefficiency in the spread of such M vectors, we allow the kernel to instruct the device with the needed number of such vectors, before VF is initialized and bounded to the driver.
Steering Features		
2239829	Match on Dynamic Configured Flex Parser	**[ConnectX-5 and ConnectX-6 Dx]** Added support to allow match on dynamic configured flex parser. The PRM allows f configuring a flex parser by creating a PARSE_GRAPH_NOC general object. This version contains the support to match parsed value using this configured flex parser using misc4 struct containing an index and value for the configured flex parser.
Steering UserSpace Features		
2380582	Pop VLAN on VF/SF Tx Direction	**[ConnectX-6 Dx and above]** Added support to pop VLAN or VF/SF Tx direction.
2363360	Dump Single Flow	**[ConnectX-5, ConnectX-6 Dx and above]** Added support to dump single flow/rule with flow-id.

Upstream Kernel 5.12			
Internal Ref.	Feature	Description	**Suppor Added i User Sp Version**
ASAP² Features			
2402429	Connection Tracking Replay State	**[BlueField-2 with CentOS kernels]** Added support for matching on CT state replay.	iproute2 V5.12, C 2.14.1
2198083	Connection Tracking Mirroring Offload	**[ConnectX-5 and above]** Added support for using Mirroring Offload with Connection Tracking.	N/A
2103546	Tunnel Rules Offload	**[ConnectX-6 Dx]** Added support for offloading tunnel rules when the source interface is VF (in addition to uplink) in the Hypervisor.	N/A
2089370	Kernel Software Steering ConnectX-6 Dx Support	**[ConnectX-6 Dx]** Added support for kernel software steering on ConnectX-6 Dx adapter cards.	N/A
NetDev Features			
2252726	kTLS RX and TX Hardware Offload Support for BOND	**[ConnectX-6 Dx and above]** Added support in bond driver to benefit from kTLS offload on capable lower devices.	N/A
1871320	Traffic Engineering— Hierarchical QoS	**[ConnectX-5 and above]** This feature allows the user to offload the HTB qdisc to the NIC, allowing it to scale better by eliminating a single locking point. The configuration is done with the:) similar tc commands.	iprroute V5.12

Upstream Kernel 5.12			
1869050	WHAT-JUST-HAPPENED-support for NICs	**[ConnectX-4 and above]** Added support for WJH on NICs. WJH allows visibility of dropped packets and their contents to increase ability for analysis and debugging. 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. • 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 Devlink provides an <u>infrastructure</u> 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, WJH support for NICs will not implement Policy.	Devlink infrastrt
1839182	Synchronize Hardware Clock to PTP Signal on ConnectX-6 DX	**[ConnectX-6 Dx]** Added support for the hardware clock device to be adjusted and provide timestamps which are translated into real-time nanoseconds. This can be used by the driver for PTP protocol.	rdma-cc v34
Core Features			

Upstream Kernel 5.12			
2479696	Flex IO	**[ConnectX-7 and above]** Added support for auxiliary processor on NICs called Flex IO. It is possible to open a dedicated CQ for this processor in order to trigger it. This requires support for opening a CQ without a valid EQ number when the following conditions are met: (1) HCA_CAP.apu = 1, (2) apu_thread_cq = 1 in CQC. In such a case, EQ number should be set to 0xFFs on CQ creation.	N/A
RDMA Features			
2392058	ODP Locking Optimization	**[ConnectX-4 and above]** Added support for cleanup of the synchronize_srcu() from the ODP flow because it was a time consuming part of dereg_mr. **Note:** This only affects the driver and not the firmware.	N/A
2359452	NDR Speed	**[ConnectX-7]** Added support for NDR Speed (400Gb/s InfiniBand rate) in mlx5 driver.	N/A
2320662	DCT Support for Connection Establishment with RDMA_CM	**[ConnectX-5 and above]** Added support for the dv APIs to allocate/deallocate a unique QP number that can be used as DC QPN in RDMA_CM connection establishment.	rdma-c(v34
Steering UserSpace Features			
2330345	Connection Tracking Window Validation	**[ConnectX-6 Dx]** Added support for ASO connection tracking of action creation and modification. This action allows performing TCP connection tracking using hardware offloads. Using this offload, the validity of the connection state of the incoming or outgoing packets on this TCP connection can be examined. Also added the ability for an ASO CT action created on one GVMI to be used on different GVMI.	rdma-c(v34
2275829	mlx5dv_dr API Support for GTP-U Extension Header	**[ConnectX-5 and ConnectX-6 Dx]** A dded 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.	rdma-c(v34

Upstream Kernel 5.12			
2264190	Modifying GTPU TEID	**[ConnectX-6 Dx and BlueField-2]** Added support to modify GTPU TEID. This support is experimental. Modifying TEID requires configuring flex parser.	rdma-c⊏ v34

Upstream Kernel 5.11				
Internal Ref.	Feature	Description	**Support Added in User Space Version**	**Support Added in Firmware Version**
Core Features				
2329342	Auxiliary Bus in mlx5 Driver	**[ConnectX-4 and above]** Updated mlx5 driver to use auxiliary bus in order to integrate different driver components into driver core and optimize module load/unload sequences.	N/A	N/A
NetDev Features				
2263652	TX CQE Error Resiliency	**[ConnectX-3 and ConnectX-3 Pro]** Added TX CQE error resiliency on mlx4 driver. When error CQE was found while polling TX CQ, the QP is in error state and all posted WQEs generated error CQEs without any data transmitted. Added resiliency flow to handle it by reopening the channels.	N/A	N/A

Upstream Kernel 5.11				
2143621	TX Port Timestamping for PTP Packets	**[ConnectX-6 Dx]** Added support for TX Port Timestamping for PTP Packets.Transmitted packet timestamping accuracy can be improved when using timestamp generated at the port level instead of timestamp generated when CQE is generated. TX port timestamping better reflects the actual time of a packet's transmit and has less jitter. As part of the feature, the driver is to open a special SQ per TC (PTP SQ). The driver is to steer the following streams to these special SQs:1. SKBTX_HW_TSTAMP was set at tx_flag (SO_TIMESTAMPING was set via setsockopt() or similar)2. Packet type is (a) non-IP, with EtherType of PTP over IEEE 802.3 (0x88f7) or (b) UDP over IPv4/IPv6.By default, the feature is disabled (to avoid the extra memory allocated for these SQs).The feature can be enabled/disabled by the following command:ethtool --set-priv-flags <ifs-name> tx_port_ts on / off.	N/A	xx.29.1016
RDMA Features				
2256043	Multi-Application QoS	**[ConnectX-5 and above]** Added support for configuring QoS on a single QP or on a group of QPs.	N/A	N/A
RDMA UserSpace (RDMA-Core) Features				
2315048	mlx5dv API	**[ConnectX-4 and above]** 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.	rdma-core v33	xx.29.1016
Steering UserSpace Features				

Upstream Kernel 5.11				
1884522	Kernel Software Managed Flow Steering (SMFS) Performance	**[ConnectX-5, ConnectX-6 Dx & BlueField, BlueField-2]** Improved the performance of Kernel software steering by reducing its memory consumption.	rdma-core v33	N/A
2254030	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 the devx command, after which a rule can be created to match on parser ID and data.	rdma-core v33	N/A
2239811	Pop/Push VLAN	**[ConnectX-6 Dx]** Enabled pop VLAN action in TX and push VLAN action in RX. This will allow the host to send and receive packets with a VLAN, but, in steering, this VLAN can be removed or replaced.	rdma-core v33	N/A
2238570	mlx5dv_dr API ASO First Hit	**[ConnectX-6 Dx & BlueField-2]** 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.	rdma-core v33	xx.29.1016

Upstream Kernel 5.11				
2228100	mlx5dv_dr API ASO Flow Meter	**[ConnectX-6 Dx & BlueField-2]** 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.. It is marked according to the current rate compared to the reference rate.	rdma-core v33	xx.31.10xx
2154380	Software Steering Parallel Rules Insertion	**[ConnectX-6 Dx & 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.	rdma-core v33	N/A

Upstream Kernel 5.10				
Internal Reference Number	**Feature/Change**	**Description**	**Support Added in User Space Version**	**Support Added in Firmware Version**
ASAP Features				
2282002	IPv6 TOS Rewrite	[**ConnectX-5 and above**] Added support for offloading TOS value rewrite for IPv6 packets.	N/A	N/A
2275891	Connection Tracking with Hairpin	[**ConnectX-5 and above**] Added support for adding connection-tracking rules on virtual functions (VF) in order to forward traffic from one VF to the other.	N/A	xx.28.4000

Upstream Kernel 5.10				
1890766	Connection Tracking with Hairpin	[**ConnectX-5 and above**] Added support for adding connection-tracking rules on uplinks in order to forward traffic from one uplink to the other.	N/A	N/A
2234654	TC Trap	[**ConnectX-5 and above**] Added support for offloading TC trap actions, allowing for redirection of packets to slow path (software).	N/A	N/A
1750259	SR-IOV Live Migration	[**ConnectX-5 and above**] 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.	N/A	N/A
RDMA Core Features				

Upstream Kernel 5.10				
1995519	Command Interface Resiliency From A Lost Interrupt	[**ConnectX-4 and above**] Added support for command interface resiliency from a lost interrupt, which manually polls the command EQ in case of a command timeout. If the resiliency mechanism finds non-handled EQE (due to a lost interrupt), it will consume it and return the actual status to the caller. Without this feature, when the driver detects a command interface command timeout, it warns the user and returns a timeout error to the caller and the entry of the command is not evacuated (because only real event interrupt can clear a command interface entry). If the hardware event interrupt never arrives, this entry is left unused forever. Because command interface entries are limited, the ability to post a new command can eventually be lost. In addition, if the driver does not consume the EQE of the lost interrupt and rearm of EQ, no new interrupts will arrive for other completed commands.	N/A	N/A
1959838	Firmware Live Patching	[**ConnectX-4 and above**] Added support for firmware live patching, which allows to update firmware without reset and applies to minor changes in firmware. The firmware checks if the firmware diff can be applied by live patching.	iproute2 v5.10	xx.28.2006
1889839	Firmware Reset on Devlink	[**ConnectX-4 and above**] Added support for initiating NIC firmware reset for a firmware upgrade through the devlink tool.	iproute2 v5.10	xx.28.2006

Upstream Kernel 5.10				
1869155	Multi-Host Reset	[**ConnectX-4 and above**] Added support for multi-host firmware reset for firmware upgrades. When a new firmware was burnt on flash and is pending activation, the multi-host reset loads the new firmware. If no new firmware is pending, the current firmware image will be reloaded from flash.	iproute2 v5.10	xx.28.2006
NetDev Features				
2176729	CQE Compression Multi-Strides	[**ConnectX-5 and above**] Added support for RX CQE compression for multi-strides packets.	N/A	N/A
2145391	IPsec Crypto Offload	[**ConnectX-6 Lx and above**] Added support for IPsec crypto offload. IPsec offload is a set of features aimed at reducing the CPU overhead of using IPsec. It targets the heavy crypto (encryption/decryption) operation, offloading them to the hardware. Connectx6-Lx is the first ASIC to support IPsec crypto offload and handles all encryption, decryption, and authentication, leaving the rest (replay, encapsulation, and decapsulation) to the software IPsec implementation. This approach benefits from increased flexibility in software, at the cost of some performance overhead.	N/A	xx.28.1002
1885835	Enhanced TX MPWQE	[**ConnectX-5 and above**] Added support for Enhanced TX MPWQE which allows for acceleration of TX datapath by saving PCI bandwidth and CPU utilization. The savings are achieved by aggregating multiple packets into a single WQE. The feature is driven by xmit_more for certain traffic types, such as UDP.	N/A	N/A

Upstream Kernel 5.10				
RDMA Features				
2149631	Memory Consumption of Memory Registration	[**ALL**] Added support for optimization of memory consumption of memory registration. Without this change, a huge pages system with 2MB page size would have a memory consumption of x512).	N/A	N/A
2048075	Query GID API	[**ALL**] Added support for a new query GID API which works over ioctl and allows querying a single GID entry by its port and GID index or querying all the GID tables of a specific device. This is much faster than querying it over sysfs.	rdma-core v32	N/A
1974041	ODP On Demand Synchronization	[**ConnectX-4 and above**] Added support to expose an option to prefetch ODP MR without faulting. This enables updating the device page table with the presenting CPU pages and reducing page faults in the system.	rdma-core v32	N/A
Steering UserSpace Features				
2228217	Remote Mirroring	[**ConnectX-5 and above**] Added support for remote mirroring, which allows for duplication of packets to multiple different destinations as well as performing a variety of different actions per destination.	rdma-core v32	N/A
2068596	Software Steering	[**ConnectX-6 Dx and above**] Added support for software steering on ConnectX-6 Dx adapter cards in the user-space RDMA-Core library through the mlx5dv_dr API.	rdma-core v32	N/A

Upstream Kernel 5.10				
2066943	Sampling Action on Direct Rule API	[**ConnectX-5 and above**] Added support for Direct Rule (DR) API, enabling sampling action to allow duplication and sampling of a ratio of the traffic. The API also allows to perform a set of different DR actions on the sampled traffic.	rdma-core v33	xx.28.1002

Upstream Kernel 5.9				
Internal Reference Number	**Feature/Change**	**Description**	**Support Added in User Space Version**	**Support Added in Firmware Version**
NetDev Features				
2067103	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 supported. pci_ro_write [on/off]. 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: "RlxdOrd-" : setpci -s82:00.0 CAP_EXP+8.w=294e	N/A	N/A

Upstream Kernel 5.9

1888097	Devlink Port Health Reporters	[**ConnectX-4 and above**] 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.	iproute2 v5.9	N/A
1821120	TLS RX Hardware Offload	[**ConnectX-6 Dx and above**] Added beta-level support for hardware offload decryption of TLS traffic over crypto-enabled ConnectX-6 Dx NICs and above.	N/A	xx.28.2006
RDMA Features				
2067098	CQ and QP Context Exposure	[**ConnectX-4 and above**] Exposed QP, CQ and key context in raw format via RDMA tool.	N/A	xx.28.0214
2056081	IPoIB QP Number Creation	[**ConnectX-4 and above**] Enabled setting the QP number of an IPoIB PKey interface in Enhanced mode. This is done using the standard `ip link add` command while padding the hardware address of the newly created interface. The QP number is the 2nd-4th bytes. To enable the feature, the `MKEY_BY_NAME` configuration should firstly be enabled in NVCONFIG.	N/A	N/A

Upstream Kernel 5.9				
2047060	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.	rdma-core v31	xx.27.1016
1955266	Shared PD	[**ConnectX-4 and above**] Added the ability to share a PD between processes for RDMA usage. This feature is particularly critical for business logic solutions that are based on multi-process designs, such as: • NGINX, with TCP load balancing, sharing the RSS indirection table with RQ per process. • HPC frameworks with multi-rank (process) solution on single hosts. UCX can share IB resources using the shared PD and can help dispatch data to multiple processes/MR's in a single RDMA operation. • Primary processes registering a large shared memory range, where each worker process spawned can create a private QP on the shared PD, and use the shared MR to save the registration time per-process.	rdma-core v31	N/A

Upstream Kernel 5.8			
Internal Reference Number	Feature/Change	Description	Support Added in User Space Version
ASAP^2 Features			
1752536	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.	iproute2 v5.8
NVMe Features			
1915224	Metadata/T10-PI Offload Support for NVMe over Fabrics (RDMA) Host and Target	[**ConnectX-4 and above**] Added support for providing offloaded end-to-end data integrity and validation according to T10-PI standard, also known as Metadata in NVMe specification.	N/A
Steering Features			
2077839	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().	rdma-core v30
1899086	mlx5dv_dr Memory Allocation	[**ConnectX-5 and above**] Reduced memory allocation time when using the mlx5dv_dr API. This is particularly significant for the first inserted rules on which memory is allocated.	rdma-core v30
1894989	Default Miss Action Support	[**ConnectX-5 and above**] Added support for a new API for creating a Default Miss Action to achieve better performance. When a packet is matched with the default "miss action", the packet will be steered to the default "miss" of the steering domain.	rdma-core v30
RDMA Features			

Upstream Kernel 5.8

2053467	mlx5dv Helper APIs for Tx Affinity Port Selection	[**ConnectX-4 and above**] Added support for the following mlx5dv helper APIs which enable the user application to query or set a QP's Tx affinity port number in a LAG configuration. • mlx5dv_query_qp_lag_port • mlx5dv_modify_qp_lag_port	rdma-core v31
2051494	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.	rdma-core v30
1919657	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.	rdma-core v31
1888194	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.	N/A
1888159	RDMA-CM Disassociate Support	[**All**] Added support for connecting kernel and RDMA-CM in a reliable way based on device index.	rdma-core v30
1779659	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-core v30

Upstream Kernel 5.7

Internal Reference Number	Feature/Change	Description	Support Added in User Space Version	Support Added in Firmware Version
ASAP^2 Features				

Upstream Kernel 5.7				
2074669	TC Query Stats and Flow Insertion Dependency Removal	[**ConnectX-5 and above**] Optimized the kernel's TC subsystem to use fine-grained per-action locks instead of global RTNL lock when constructing flow-action intermediate representation for the driver. This change allows for rules insertion to run in parallel with rule dump, which significantly improves performance of live Open vSwitch system.	N/A	N/A
1818081	Hairpin Support in Switch Mode	[**ConnectX-5 and above**] Added support for receiving tunneled traffic from the uplink port, after which it can be decapsulated and sent them back to the uplink port, possibly encapsulated with new tunnel information.	N/A	xx.28.1000
1765236	Connection Tracking Offload	[**ConnectX-5 and above**] Added support for offloading TC filters containing connection tracking matches and actions.	OVS v2.13 + iproute2 v5.4.0	xx.26.1040
Core Features				
1928787	Firmware Reactivation	[**ConnectX-4 and above**] Added support for safely inserting consecutive firmware images without the need to reset the NIC in between.	N/A	xx.28.1000
Steering Features				

Upstream Kernel 5.7				
1997168	GPRS Tunneling Protocol (GTP) Header	[**ConnectX-5 and above**] Added support for matching (filtering) GTP header-based packets using mlx5dv_dr API over user-space RDMA-Core library.	rdma-core v29	N/A
RDMA Features				
2067127	Custom Parent-Domain Allocators for CQ	[**ConnectX-4 and above**] Enabled specific custom allocations for CQs.	rdma-core v29	N/A
1890393	RDMA-CM DevX Support	[**ConnectX-4 and above**] Added support for DevX in RDMA-CM applications.	rdma-core v29	N/A
1830364	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().`	rdma-core v29	xx.27.1016
NetDev Features				

Upstream Kernel 5.7				
1888096	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.	N/A	N/A
1888020	Get FEC Status on PAM4/50G	[**ConnectX-6 & ConnectX-6-Dx**] Allowed configuration of Reed Solomon and Low Latency Reed Solomon over PAM4 link modes.	ethtool v5.5	N/A
1702814/1702706	Devlink Health Dump	[**ConnectX-4 and above**] Added support for the `devlink health dump show` command in the Rx and Tx reporters. The dump contains the content given by the firmware for wqump/rxdump/sxdump in binary format. The user can retrieve the binary data via devlink (v5.3 and on) and use a NVIDIA parsing tool to analyze the dump.	iproute2 v5.3	xx.27.1016

Upstream Kernel 5.6				
Internal Reference Number	**Feature/Change**	**Description**	**Support Added in User Space Version**	**Support Added in Firmware Version**
RDMA Features				

Upstream Kernel 5.6				
1978779	ODP Huge Pages Support	[**ConnectX-4 and above**] Enabled ODP MR to work with huge pages by exposing IBV_ACCESS_HUGETLB access flag to indicate MR range mapping by huge pages. This flag is applicable only with conjunction of IBV_ACCESS_ON_DEMAND.	rdma-core v28	N/A
1973016	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 • **roce_slow_restart_trans** - counts the number of times RoCE slow restart changed state to slow restart	N/A	N/A
1885435	vhost Data Path Acceleration (vDPA)	[**ConnectX-5 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().	rdma-core v28	xx.27.1016

Upstream Kernel 5.6

1882973	Memory Region	[**All**] Added support for the user to register memory regions with a relaxed ordering access flag. This can enhance performance, depending on architecture and scenario.	rdma-core v28	xx.27.1016
1769019	Kernel Recommendation on User Memory	[**ConnectX-4 and above**] Added support for the Kernel to advise on the user memory region that is registered in the HCA. In order to be able to use this functionality, virtual address and memory key must be known to the caller.	N/A	N/A
1711108	User Memory Kernel Registration	[**ConnectX-4 and above**] Added support for the Kernel to register user virtual memory in the HCA. This requires the user to pass the address and size of the memory to the kernel in some way and the kernel must register the memory from the contexts of the user process.	N/A	N/A
Steering Features				
1961704	Kernel Software Steering Remote Mirroring	[**ConnectX-5**] Added kernel support for OVS remote mirroring to allow hardware traffic mirroring to multiple ports.	N/A	N/A
1915174	Kernel Software Steering for Connection Tracking (CT)	[**ConnectX-5**] Added required kernel software steering support for CT.	N/A	N/A
1840759	Userspace Software Steering Debugging API	[**ConnectX-5**] Added support for software steering to dump flows for debugging purposes in the user-space RDMA-Core library through the mlx5dv_dr API.	rdma-core v28	N/A

Upstream Kernel 5.6				
1754441	Discards Counters[1]	[**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.	N/A	N/A
NetDev Features				
2042124	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.	iproute2 v5.6	N/A
1996062	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.	N/A	N/A

Upstream Kernel 5.5				
Internal Reference Number	**Feature/Change**	**Description**	**Support Added in User Space Version**	**Support Added in Firmware Version**
RDMA Features				

Upstream Kernel 5.5				
1895928	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.	rdma-core v27	N/A
1756530	ODP Diagnostic counters	[**ConnectX-4 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 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.	iproute v5.4	N/A
1747149	Performance Improvements	[**ConnectX-4 and above**] • Updated Blueflame capability reporting to prevent redundant use of Blueflame when Write-combining is not supported. • Added Blueflame capabilities over VFs.	N/A	xx.26.1040

Upstream Kernel 5.5				
1686052	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.	N/A	N/A
rdma-core Only Features[1]				
1871545	Hardware Clock Exposure	[**ConnectX-5 and above**] Added support for querying the adapter clock via mlx5dv_query_device.	rdma-core v27	N/A
Steering Features				
1915338	Direct Verbs Support for Batch Counters on Root Table	[**ConnectX-5**] Added support for mlx5dv_dr API to set batch counters for root tables.	rdma-core v27	N/A
1907530	Modify Header	[**ConnectX-5 & BlueField**] Added support for mlx5dv_dr_actions to support up to 32 modify actions.	rdma-core v27	N/A
1757878	GENEVE Encap/Decap Rules Offload	[**ConnectX-5**] Added support for GENEVE encapsulation/decapsulation rules offload using software steering.	rdma-core v27	N/A
ASAP^2 Features				
1920181	OVS-Kernel ToS Rewrite	[**ConnectX-5 and above**] Added support for Type of Service (ToS) rewrite in the OVS-Kernel.	N/A	N/A

Upstream Kernel 5.4				
Internal Reference Number	**Feature/Change**	**Description**	**Support Added in User Space Version**	**Support Added in Firmware Version**

Upstream Kernel 5.4

RDMA Features

1717501	RDMA_Rx RoCE Flow Steering Support	[**ConnectX-4 and above**] Extended support for mlx5 with RDMA_RX RoCE flow steering to DevX objects.	rdma-core v26	xx.26.1000
1747205		[**ConnectX-4 and above**] Extended support for mlx5 with RDMA_RX RoCE flow steering to QP objects.	rdma-core v26	xx.26.1000
1821108	ODP Support over DC Transport	[**ConnectX-4 and above**] Added support for On-Demand-Paging over DC transport.	rdma-core v26	N/A

Steering Features

1751025	Push/Pop VLAN over Software Steering	[**ConnectX-5 and above**] Added support to pop and push VLAN actions in software steering in SwitchDev mode.	N/A	N/A
1762181	Flow and Recycle Counters	[**ConnectX-4 and above**] Added support for a pool of flow counters, based on flow counter bulks, with the option to acquire and release flow counters from and to the pool. This shortens the time needed to allocate a flow counter using a FW command. The effective memory footprint of a flow counter is also increased by 10% thanks to this feature.	N/A	16.25.1000
1703811	Kernel Software Steering for eSwitch	[**ConnectX-5**] Added software steering capabilities to the SR-IOV eSwitch in SwitchDev mode. Software steering enables better rules update rate compared to the current firmware-based solution. This is achieved by performing calculations on the main CPU which allows for higher update rates.	N/A	N/A

NetDev Features

Upstream Kernel 5.4				
1700763	PF to Show VF Counters	[**ConnectX-4 and above**] Added support for the VF to send real-time rings statistics (Rx/Tx packets/bytes) to the PF over HyperV PCI configuration channel. These statistics can be monitored via perfmon tool.	N/A	N/A
1821239	Ethtool Counters	[**ConnectX-5 and above**] Added counter dev_out_of_buffer to ethtool statistics to count the number of times there was no enough buffer allocated to the device internal queue.	N/A	xx.26.1040
1821554	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.	N/A	xx.26.4012
1858107	AF_XDP need_wakeup Flag Support	[**ConnectX-4 and above**] Added driver support for the flag "need_wakeup" in AF_XDP Tx and FILL rings. When both the AF_XDP application and the driver support this flag, the driver is able to signal to the application when it needs to invoke a syscall to wake the driver explicitly, and when it needs not. This significantly improves single-core performance by reducing resource contention and avoiding wasting resources on busy polling and context switches.	N/A	N/A

Upstream Kernel 5.4				
1821229	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.	N/A	xx.26.1040
1760559	Virtual Output Queuing (VoQ) Counters	[**ConnectX-6 Dx**] 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.	N/A	See note below[1]
1821559	IP-in-IP RSS Offload	[**ConnectX-4 and above**] Added support for receive side scaling (RSS) offload in IP-in-IP (IPv4 and IPv6).	N/A	N/A
1782820	Devlink Health Reporters	[**ConnectX-4 and above**] Added support for monitoring and recovering from errors that occur on the Rx queue, such as CQE errors and timeout.	N/A	N/A
[1] ConnectX-6 Dx FW version supporting this feature is not available yet.				
ASAP^2 Features				
1682148	Parallel TC Rules Insertion	[**ConnectX-5 and above**] Removed global RTNL lock dependency in the kernel and in mlx5 TC layers to allow updating parallel TC rules.	N/A	N/A

Upstream Kernel 5.4

1767933	Offloading Ingress Rate Control Offload on E-Switch Ports	[**ConnectX-5 and above**] Added support for configuring the ingress rate of an E-Switch port. This is achieved by the ability to control the transmit rate of a virtual function whose data appears at the ingress of the Switch port. Thus, controlling the VF Tx rate effectively controls the ingress rate of an E-Switch Port.	N/A	N/A

Upstream Kernel 5.3

Internal Reference Number	Feature/Change	Description	Support Added in User Space Version	Support Added in Firmware Version
RDMA Features				
1755657	Device Emulation Infrastructure	[**BlueField**] Added support for Device Emulation in BlueField. This mechanism allows function-A to perform operations "on behalf" of function-B. The emulation manager creates a channel (named VHCA_TUNNEL general object) that acts as the direct command interface between the emulated function host and the HCA hardware. The emulation software creates this tunnel for every managed function and issues commands via the DevX general command interface.	N/A	xx.26.1040
1711102	DevX Async Events Support	[**ConnectX-4 and above**] Added support for reading device events over DevX API.	rdma-core v25	xx.25.1020

Upstream Kernel 5.3				
1686017	QP Counters and Firmware Errors per PID	[**ConnectX-5 and above**] QP counters and flow counters are now set per Process ID (PID) to allow better visibility of RDMA error states. Users will be able to manually tune the Q counter to monitor specific QPs, or automatically monitor QPs according to predefined criteria, such as the QP type.	iproute2 v5.3	xx.25.1020
1679990	Dynamic Interrupt Moderation for RDMA Kernel	[**ConnectX-5 and above**] Added dynamic interrupt moderation (DIM) library for RDMA to improve bandwidth and reduce CPU overhead.	iproute2 v5.3	N/A
1160889	Signature Offload Kernel Verbs Enhancements	[**ConnectX-4 and above**] Added a new API which enables posting a single WR that completes the Protection Information (PI) operation internally. This reduces CPU utilization for posting and processing multiple WRs and improves performance by choosing the optimal mkey for the hardware according to the buffer memory layout.	N/A	N/A
1704114	Message Signaled Interrupts-X (MSI-X) Vectors	[**ConnectX-4 and above**] 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.	N/A	xx.26.1040
NetDev Features				

Upstream Kernel 5.3				
1750279	Firmware Flash Update via Devlink	[**ConnectX-4 and above**] Added the option to update the firmware image in the flash using the devlink tool. **Usage**: ```devlink dev flash <dev> file <file_name>.mfa2```	iproute2 v5.3	xx.25.1020
1747672	Firmware Versions Query via Devlink	[**ConnectX-4 and above**] Added the option to query for running and stored firmware versions using devlink tool.	iproute2 v5.3	N/A
1706472	Generic Segmentation Offload (GSO) Optimization	[**ConnectX-4 and above**] Improved GSO workload performance by decreasing doorbells usage to the minimum required.	N/A	N/A
1704398	Devlink Firmware Fatal Errors Handling	[**ConnectX-4 and above**] Added support for the driver to detect firmware fatal errors and initiate CR-Dump and recovery flow.	iproute2 v5.3	N/A
1700834	AF_XDP Socket Zero Copy	[**ConnectX-4 and above**] Added zero-copy support for AF_XDP sockets to improve performance.	N/A	N/A
1700781	Devlink Health CR-Space Dump	[**ConnectX-4 and above**] Added the option to dump configuration space via the devlink tool to improve debuggability.	iproute2 v5.3	N/A
1610373	TLS TX Hardware Offload	[**ConnectX-6 Dx**] Added support for hardware offload encryption of kTLS traffic to improve performance. **Note**: This is an infrastructure feature only, to enable future support for ConnectX-6 Dx adapter cards.	N/A	xx.26.1040

Upstream Kernel 5.3

Internal Reference Number	Feature/Change	Description	Support Added in User Space Version	Support Added in Firmware Version
1186273	Devlink Health Firmware Core Dump	**[ConnectX-4 and above]** Added support for dumping firmware core data via the devlink tool to improve debuggability. Note that to allow firmware core dumping, FW Tracer should be enabled.	iproute2 v5.3	N/A

ASAP² Features

Internal Reference Number	Feature/Change	Description	Support Added in User Space Version	Support Added in Firmware Version
1767550	Dual-Port RoCE Support	**[ConnectX-5 and above]** Enabled simultaneous operation of dual-port RoCE and Ethernet in SwitchDev mode.	N/A	xx.25.6000
1765208	Geneve Encap/Decap Rules Offload	**[ConnectX-5 and above]** Added support for Geneve encapsulation/decapsulation rules offload.	N/A	xx.25.1020
1693782	VLAN PUSH/POP Offload (VGT)	**[ConnectX-5 and above]** Allowed offload of VLAN push/pop operations on both transmitted and received packets.	N/A	xx.25.1020

Upstream Kernel 5.2

Internal Reference Number	Feature/Change	Description	Support Added in User Space Version	Support Added in Firmware Version
RDMA Features				
1688852	DevX Representors	**[ConnectX-4 and above]** Changed InfiniBand representors into a single InfiniBand device with multiple ports, where each port represents a different virtual function (port 1 is for the uplink), to allow DevX control by exposing a further uniformed software model.	N/A	xx.25.1020

Upstream Kernel 5.2				
1706795	Userspace Software Steering for eSwitch	**[ConnectX-5]** Added software steering capabilities to the SR-IOV eSwitch. Software steering enables better rules insertion rate compared to the current firmware-based solution. This is achieved by performing calculations on the main CPU which allows for higher insertion rates.	rdma-core v24	xx.25.1020
1686065	Userspace Software Steering for NICs	**[ConnectX-5]** Added software steering capabilities to NIC Rx/Tx. Software steering enables better rules insertion rate compared to the current firmware-based solution. This is achieved by performing calculations on the main CPU which allows for higher insertion rates. This solution was designed to work with Virtio DPDK. **Note**: Support will be enabled by default once the support for GID change is added.	rdma-core v24	xx.25.1020
rdma-core Only Features[1]				
1677617	Dynamic Connection (DC) Data Path	**[ConnectX-4 and above]** Added DC QP data-path support using new Send APIs introduced in Direct Verbs (DV).	rdma-core v24	N/A
1686034	Send APIs	**[ConnectX-4 and above]** Introduced a new set of QP Send operations (APIs) that allow extensibility for new Send opcodes.	rdma-core v24	N/A
1705623	User-Mode Memory Registration (UMR)	**[ConnectX-4 and above]** Enabled registration of memory patterns that can be used for future RDMA operations.	rdma-core v24	xx.25.1020
[1] These features require rdma-core update only - no kernel update needed.				
NetDev Features				

Upstream Kernel 5.2				
1705845	XDP Inline Transmission of Small Packets	**[ConnectX-5 and above]** Added support for when forwarding packets with XDP, a packet smaller than 256 bytes would be sent inline within its WQE Tx descriptor for better performance. The number of packets that are transmitted inline depends on CPUs load, where lower load leads to a higher number of inline transmission.	N/A	N/A
1706862	EEPROM Device Thresholds via Ethtool	**[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 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`	N/A	N/A
1699196	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. 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.	N/A	xx.25.1020

Upstream Kernel 5.2				
1535442	Multiple Outstanding UMR Posts	**[ConnectX-4 Lx and above]** Added support for multiple outstanding UMR posts in order to allow faster gap closure between consuming Multi-Packet WQEs (MPWQEs) and re-posting them back into the WQ. This improves Rx packet rate on Striding RQ as it solves scaling and stability issues.	N/A	N/A
1703927	GENEVE Tunnel Stateless Offload	**[ConnectX-4 and above]** Added support for Generic Network Virtualization Encapsulation (GENEVE) tunneled hardware offload of TSO, CSUM and RSS.	N/A	N/A
ASAP² Features				
1699405	VLAN Rewrite	**[ConnectX-5 and above]** Added support for offloading VLAN ID modify operation, allowing the user to replace the VLAN tag of the incoming frame with a user-specified VLAN tag value.	N/A	xx.25.1020
1699042	VLAN Support over Uplink	**[ConnectX-5 and above]** Added support for offloading TC rules from VF representor and forwarding them to a VLAN device over the physical function (PF), and vice-versa.	N/A	N/A

Upstream Kernel 5.1				
Internal Reference Number	**Feature/Change**	**Description**	**Support Added in User Space Version**	**Support Added in Firmware Version**
RDMA Features				

Upstream Kernel 5.1				
1678107	SRQ and XRC Support on On Demand Paging (ODP) Memory Region (MR)	**[ConnectX-4 and above]** Added support for using ODP MR with SRQ WQEs and XRC transport.	rdma-core v23	xx.25.1000
1698300	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.	rdma-core v23	N/A
1685107	Indirect Mkey ODP	**[ConnectX-4 and above]** Added the ability to create indirect Mkeys with ODP support over DevX interface.	N/A	xx.25.1000
1698230	DevX Asynchronous Query Commands	**[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.	rdma-core v23	xx.25.1000
1686056	RDMA-CM QP Timeout Control	**[All]** Added a new option to rdma_set_option that allows applications to override the RDMA-CM's QP ACK timeout value.	rdma-core v23	N/A
1647527	Object IDs Exportation	**[All]** Added a unique ID for each verbs object to allow direct query over rdma-tool and rdma-netlink for enhanced debuggability.	iproute v5.1	N/A
rdma-core Only Features[1]				

Upstream Kernel 5.1				
1698278	Implicit ODP	**[ConnectX-4 and above]** Added support for reporting implicit ODP support to user applications in order to allow better granularity over ODP creation.	rdma-core v23	N/A
1698290	RDMA-CM Application Managed QP	**[All]** Added support for the RDMA application to manage its own QPs and use RDMA-CM only for exchanging Address information.	rdma-core v23	N/A
1680025	mlx4 Internal Resources Sharing	**[ConnectX-3/ConnectX-3 Pro]** Added support for a new API for sharing internal resources, such as UAR and memory allocation, which will enable higher layers, such as DPDK, to implement a secondary process.	rdma-core v23	N/A
NetDev Features				

Upstream Kernel 5.1				
1700770	Devlink Health Utility	**[ConnectX-4 and above]** Added support for real-time alerting of functionality issues that may be found in a system component (reporter). This utility helps detect and recover from a problem with a PCI device. It provides a centralize status of drivers' health activities in the generic Devlink instance and inter alia, supports the following: • Storing real-time error dumps • Performing automatic (configurable) real-time reporter recovery • Performing real-time reporter diagnosis • Indicating real-time reporter's health status • Providing admins with the ability to dump, diagnose and recover a reporter • Providing admins with the ability to configure a reporter	iproute v5.1	N/A

Upstream Kernel 5.1				
1698287	Ethtool 200Gbps	**[ConnectX-6 and above]** ConnectX-6 hardware introduces support for 200Gbps and 50Gbps-per-lane link mode. The driver 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. **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.	ethtool v5.1 (tentative)	xx.25.1000
1650372	Virtual Ethernet Port Aggregator (VEPA)	**[ConnectX-4 and above]** 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: bridge link dev <netdev> hwmode vepa	N/A	N/A

Upstream Kernel 5.1				
1656472	ConnectX Device IDs	Added support for the following new device IDs: • ConnectX-6 Dx (PF) • ConnectX Family mlx5Gen Virtual Function (VF) Note that every new device (adapter) VF will be identified with this device ID. Different VF models will be distinguished by their revision ID.	N/A	N/A
ASAP² Features				
1700775	Equal Cost Multi-Path (ECMP)	**[ConnectX-4 Lx and above]** Added support for offloading ECMP rules by tracking software multipath route and related next-hops, and reflecting this as port affinity to the hardware.	N/A	N/A
1600548	GRE Entropy Control	**[ConnectX-4 Lx and above]** Added support for controlling port tunnel entropy calculation in order to set the correct GRE key value.	N/A	xx.25.1000
BlueField Features				
1351321	Respsenetors for Host Physical Functions	**[BlueField]** When ECPF is the E-switch manager, ECPF driver will create a representor for the host PF in Offload mode in order to ensure secure control of the host PF traffic in a bare-metal environment.	N/A	18.24.1000

Upstream Kernel 5.1

1351323	ICM Pages Allocation on Host PFs in Non Secure Mode	**[BlueField]** Added the ability to use host memory in order to increase scalability in virtual cloud scenarios where the host PF is trusted. ICM page allocation is now managed in firmware as follows: 1. Function provides pages for itself 2. PF provides pages for its VF 3. ECPF provides pages to itself and for another function	N/A	18.24.1000

Upstream Kernel 5.0

Internal Reference Number	Feature/Change	Description	Support Added in User Space Version	Support Added in Firmware Version
RDMA Features				
1474608	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.	rdma-core v22	N/A
1470521	HDR Link Speed Exposure	[**ConnectX-6 and above**] Added support for HDR link speed in CapabilityMask2 field in port attributes.	rdma-core v22	xx.25.1000

Upstream Kernel 5.0				
1470528	QP Packet Based Credit Mode	[**ConnectX-6 and above**] Added support for an alternative end-to-end credit mode for QP creation. Credits transported from the responder to the requester are now issued per packet. This is particularly useful for sending large RDMA messages from HCA to switches that are short in memory.	rdma-core v22	xx.24.1000
1463974	Fragmented QPs Buffer	[**ConnectX-4 and above**] Added the ability to allocate a fragmented buffer to in-kernel QP creation requests, in cases of large QP size requests that used to fail due to low memory resources on the host.	N/A	N/A
1498575	Flow Counters Batch Query	[**ConnectX-4 and above**] Allowed flow counters created with the DevX interface to be attached to flows created with the raw flow creation API.	rdma-core v22	N/A
1550997	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.	N/A	xx.25.1000
1471289	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.	rdma-core v22	xx.25.1000

Upstream Kernel 5.0				
1519601	Restored Verbs Fork Support	Re-added support for libibverbs users to open ibv_context using ibv_open_device(), call fork and continue to use that context from the fork.	rdma-core v22	N/A
1551627	rdma-core Python Bindings (Pyverbs)	Added support for Python API over IB verbs. Pyverbs package shortens development time and provides a comfortable API to work with.	rdma-core v22	N/A
NetDev Features				
1558262	CQE Padding	[**ConnectX-5 and above**] Added support for padding 64B CQEs to 128B cache lines to improve performance on 128B cache line systems, such as PPC.	N/A	xx.21.1000
1464598	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.	N/A	xx.24.1000
1437218	Rx Hash Fields Configuration	[**ConnectX-4 and above**] Added the ability to configure Rx hash fields used for traffic spreading into Rx queues using ETHTOOL_SRXFH and ETHTOOL_GRXFH ethtool commands. Built-in Receive Side Scaling (RSS) profiles can now be changed on the following traffic types: UDP4, UDP6, TCP4 and TCP6. This configuration affects both outer and inner headers.	ethtool v2.6.33	N/A

Upstream Kernel 5.0				
1190960	XDP Multi-Packet Tx Work Queue Element (WQE)	[**ConnectX-5 and above**] Added support for Multi-Packet Tx WQEs in XDP transmit flows to work with a new and improved WQE layout that describes several packets.This saves PCI bandwidth and transactions, and improves transmit packet rate.	N/A	N/A
ASAP^2 Features				
1082182	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.	N/A	N/A
1441639	Remote Mirroring	[**ConnectX-5 and above**] Added support for additional mirroring output in SwitchDev mode. The mirroring port may either be a local or a remote VF, using VxLAN or GRE encapsulations.	OVS v2.10.90	xx.24.1000
	GRE HW Offloading	[**ConnectX-4 Lx and above**] Added the ability to offload TC filters set on GRE interfaces to NICs that support ASAP^2.	- OVS v2.10.90 - iproute v5.0	xx.24.1000
1592719	VxLAN/GRE Tunneling over VLAN	[**ConnectX-4 Lx and above**] Added support for VxLAN and GRE tunnel encap/decap offload over Ethernet tagged packets.	N/A	N/A
1306741	Uplink Representors	[**ConnectX-4 Lx and above**] Exposed PF (uplink) representors in SwitchDev mode, similarly to VF representors, as an infrastructure improvement for SmartNICs.	N/A	N/A