

# ConnectX<sup>®</sup>-6 EN Card

## 200GbE Ethernet Adapter Card



World's first 200GbE Ethernet network interface card, enabling industry-leading performance smart offloads and in-network computing for Cloud, Web 2.0, Big Data, Storage and Machine Learning applications

ConnectX-6 EN provides up to two ports of 200GbE connectivity, sub 0.8usec latency and 215 million messages per second, enabling the highest performance and most flexible solution for the most demanding data center applications.

ConnectX-6 is a groundbreaking addition to the Mellanox ConnectX series of industry-leading adapter cards. In addition to all the existing innovative features of past versions, ConnectX-6 offers a number of enhancements to further improve performance and scalability, such as support for 200/100/50/40/25/10/1 GbE Ethernet speeds and PCIe Gen 4.0. Moreover, ConnectX-6 Ethernet cards can connect up to 32-lanes of PCIe to achieve 200Gb/s of bandwidth, even on Gen 3.0 PCIe systems.

### Cloud and Web 2.0 Environments

Telco, Cloud and Web 2.0 customers developing their platforms on Software Defined Network (SDN) environments are leveraging the Virtual Switching capabilities of the Operating Systems on their servers to enable maximum flexibility in the management and routing protocols of their networks.

Open vSwitch (OVS) is an example of a virtual switch that allows Virtual Machines to communicate among themselves and with the outside world. Software-based virtual switches, traditionally residing in the hypervisor, are CPU intensive, affecting system performance and preventing full utilization of available CPU for compute functions.

To address this, ConnectX-6 offers ASAP<sup>2</sup> - Mellanox Accelerated Switch and Packet Processing<sup>®</sup> technology to offload the vSwitch/vRouter by handling the data plane in the NIC hardware while maintaining the control plane unmodified. As a result, significantly higher vSwitch/vRouter performance is achieved without the associated CPU load.

The vSwitch/vRouter offload functions supported by ConnectX-5 and ConnectX-6 include encapsulation and de-capsulation of overlay network headers, as well as stateless offloads of inner packets, packet headers re-write (enabling NAT functionality), hairpin, and more.

In addition, ConnectX-6 offers intelligent flexible pipeline capabilities, including programmable flexible parser and flexible match-action tables, which enable hardware offloads for future protocols.

## HIGHLIGHTS

### FEATURES

- Up to 200GbE connectivity per port
- Maximum bandwidth of 200Gb/s
- Up to 215 million messages/sec
- Sub 0.8usec latency
- Block-level XTS-AES mode hardware encryption
- Optional FIPS-compliant adapter card
- Support both 50G SerDes (PAM4) and 25G SerDes (NRZ) based ports
- Best-in-class packing with sub-nanosecond accuracy
- PCIe Gen4/Gen3 with up to x32 lanes
- RoHS compliant
- ODCC compatible

### BENEFITS

- Most intelligent, highest performance fabric for compute and storage infrastructures
- Cutting-edge performance in virtualized HPC networks including Network Function Virtualization (NFV)
- Advanced storage capabilities including block-level encryption and checksum offloads
- Host Chaining technology for economical rack design
- Smart interconnect for x86, Power, Arm, GPU and FPGA-based platforms
- Flexible programmable pipeline for new network flows
- Enabler for efficient service chaining
- Efficient I/O consolidation, lowering data center costs and complexity

## Storage Environments

NVMe storage devices are gaining momentum, offering very fast access to storage media. The evolving NVMe over Fabric (NVMe-oF) protocol leverages RDMA connectivity to remotely access NVMe storage devices efficiently, while keeping the end-to-end NVMe model at lowest latency. With its NVMe-oF target and initiator offloads, ConnectX-6 brings further optimization to NVMe-oF, enhancing CPU utilization and scalability.

## Security

ConnectX-6 block-level encryption offers a critical innovation to network security. As data in transit is stored or retrieved, it undergoes encryption and decryption. The ConnectX-6 hardware offloads the IEEE AES-XTS encryption/decryption from the CPU, saving latency and CPU utilization. It also guarantees protection for users sharing the same resources through the use of dedicated encryption keys.

By performing block-storage encryption in the adapter, ConnectX-6 excludes the need for self-encrypted disks. This gives customers the freedom to choose their preferred storage device, including byte-addressable and NVDIMM devices that traditionally do not provide encryption. Moreover, ConnectX-6 can support Federal Information Processing Standards (FIPS) compliance.

## Machine Learning and Big Data Environments

Data analytics has become an essential function within many enterprise data centers, clouds and hyperscale platforms. Machine learning relies on especially high throughput and low latency to train deep neural networks and to improve recognition and classification accuracy. As the first adapter card to deliver 200GbE throughput, ConnectX-6 is the perfect solution to provide machine learning applications with the levels of performance and scalability that they require. ConnectX-6 utilizes the RDMA technology to deliver low-latency and high performance. ConnectX-6 enhances RDMA network capabilities even further by delivering end-to-end packet level flow control.

## Mellanox Socket Direct®

Mellanox Socket Direct technology improves the performance of dual-socket servers, such as by enabling each of their CPUs to access the network through a dedicated PCIe interface. As the connection from each CPU to the network bypasses the QPI (UPI) and the second CPU, Socket Direct reduces latency and CPU utilization. Moreover, each CPU handles only its own traffic (and not that of the second CPU), thus optimizing CPU utilization even further.

Mellanox Socket Direct also enables GPUDirect® RDMA for all CPU/GPU pairs by ensuring that GPUs are linked to the CPUs closest to the adapter card. Mellanox Socket Direct enables Intel® DDIO optimization on both sockets by creating a direct connection between the sockets and the adapter card.

Mellanox Socket Direct technology is enabled by a main card that houses the ConnectX-6 adapter card and an auxiliary PCIe card bringing in the remaining PCIe lanes. The ConnectX-6 Socket Direct card is installed into two PCIe x16 slots and connected using a 350mm long harness. The two PCIe x16 slots may also be connected to the same CPU. In this case the main advantage of the technology lies in delivering 200GbE to servers with PCIe Gen3-only support.

Please note that when using Mellanox Socket Direct in virtualization or dual-port use cases, some restrictions may apply. For further details, Contact Mellanox Customer Support.

## Host Management

Mellanox host management and control capabilities include NC-SI over MCTP over SMBus, and MCTP over PCIe - Baseboard Management Controller (BMC) interface, as well as PLDM for Monitor and Control DSP0248 and PLDM for Firmware Update DSP0267.

## Compatibility

### PCI Express Interface

- PCIe Gen 4.0, 3.0, 2.0, 1.1 compatible
- 2.5, 5.0, 8, 16 GT/s link rate
- 32 lanes as 2x 16-lanes of PCIe
- Support for PCIe x1, x2, x4, x8, and x16 configurations
- PCIe Atomic
- TLP (Transaction Layer Packet) Processing Hints (TPH)

- PCIe switch Downstream Port Containment (DPC) enablement for PCIe hot-plug
- Advanced Error Reporting (AER)
- Access Control Service (ACS) for peer-to-peer secure communication
- Process Address Space ID (PASID) Address Translation Services (ATS)
- IBM CAPIv2 (Coherent Accelerator Processor Interface)
- Support for MSI/MSI-X mechanisms

### Operating Systems/Distributions\*

- RHEL, SLES, Ubuntu and other major Linux distributions
- Windows
- FreeBSD
- VMware
- OpenFabrics Enterprise Distribution (OFED)
- OpenFabrics Windows Distribution (WinOF-2)

### Connectivity

- Up to two network ports
- Interoperability with Ethernet switches (up to 200GbE, as 4 lanes of 50GbE data rate)
- Passive copper cable with ESD protection
- Powered connectors for optical and active cable support

## Features\*

### Ethernet

- 200GbE / 100GbE / 50GbE / 40GbE / 25GbE / 10GbE / 1GbE
- IEEE 802.3bj, 802.3bm 100 Gigabit Ethernet
- IEEE 802.3by, Ethernet Consortium 25, 50 Gigabit Ethernet, supporting all FEC modes
- IEEE 802.3ba 40 Gigabit Ethernet
- IEEE 802.3ae 10 Gigabit Ethernet
- IEEE 802.3az Energy Efficient Ethernet
- IEEE 802.3ap based auto-negotiation and KR startup
- IEEE 802.3ad, 802.1AX Link Aggregation
- IEEE 802.1Q, 802.1P VLAN tags and priority
- IEEE 802.1Qau (QCN) – Congestion Notification
- IEEE 802.1Qaz (ETS)
- IEEE 802.1Qbb (PFC)
- IEEE 802.1Qbg
- IEEE 1588v2
- Jumbo frame support (9.6KB)

### Enhanced Features

- Hardware-based reliable transport
- Collective operations offloads
- Vector collective operations offloads
- Mellanox PeerDirect® RDMA (aka GPUDirect®) communication acceleration
- 64/66 encoding
- Enhanced Atomic operations

- Advanced memory mapping support, allowing user mode registration and remapping of memory (UMR)
- Extended Reliable Connected transport (XRC)
- Dynamically Connected transport (DCT)
- On demand paging (ODP)
- MPI Tag Matching
- Rendezvous protocol offload
- Out-of-order RDMA supporting Adaptive Routing
- Burst buffer offload
- In-Network Memory registration-free RDMA memory access

### CPU Offloads

- RDMA over Converged Ethernet (RoCE)
- TCP/UDP/IP stateless offload
- LSO, LRO, checksum offload
- RSS (also on encapsulated packet), TSS, HDS, VLAN and MPLS tag insertion/stripping, Receive flow steering
- Data Plane Development Kit (DPDK) for kernel bypass application
- Open vSwitch (OVS) offload using ASAP<sup>2</sup>
- Flexible match-action flow tables
- Tunneling encapsulation / decapsulation
- Intelligent interrupt coalescence
- Header rewrite supporting hardware offload of NAT router

### Hardware-Based I/O Virtualization - Mellanox ASAP<sup>2</sup>

- Single Root IOV
- Address translation and protection
- VMware NetQueue support
  - SR-IOV: Up to 1K Virtual Functions
  - SR-IOV: Up to 8 Physical Functions per host
- Virtualization hierarchies (e.g., NPAR)
- Virtualizing Physical Functions on a physical port
- SR-IOV on every Physical Function
- Configurable and user-programmable QoS
- Guaranteed QoS for VMs

### Storage Offloads

- Block-level encryption: XTS-AES 256/512 bit key
- NVMe over Fabric offloads for target machine
- T10 DIF - signature handover operation at wire speed, for ingress and egress traffic
- Storage Protocols: SRP, iSER, NFS RDMA, SMB Direct, NVMe-oF

### Overlay Networks

- RoCE over overlay networks
- Stateless offloads for overlay network tunneling protocols
- Hardware offload of encapsulation and decapsulation of VXLAN, NVGRE, and GENEVE overlay networks

### HPC Software Libraries

- HPC-X, OpenMPI, MVAPICH, MPICH, OpenSHMEM, PGAS and varied commercial packages

### Management and Control

- NC-SI, MCTP over SMBus and MCTP over PCIe - Baseboard Management Controller interface
- PLDM for Monitor and Control DSP0248
- PLDM for Firmware Update DSP0267
- SDN management interface for managing the eSwitch
- I<sup>2</sup>C interface for device control and configuration
- General Purpose I/O pins
- SPI interface to Flash
- JTAG IEEE 1149.1 and IEEE 1149.6

### Remote Boot

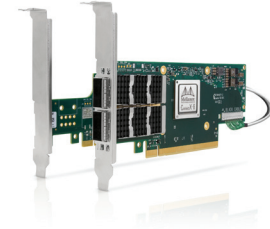
- Remote boot over Ethernet
- Remote boot over iSCSI
- Unified Extensible Firmware Interface (UEFI)
- Pre-execution Environment (PXE)

(\*) This section describes hardware features and capabilities. Please refer to the driver and firmware release notes for feature availability.

## Adapter Card Portfolio & Ordering Information

Table 1 - PCIe HHHL Form Factor

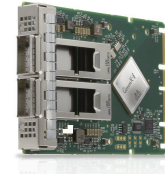
Max. Network Speed	Interface Type	Supported Ethernet Speed [GbE]	Host Interface [PCIe]	OPN
2x 100 GbE	QSFP56	100 <sup>2</sup> /50/40/25/10/1	Gen 3.0 2x16 Socket Direct	MCX614106A-CCAT
	SFP-DD		Gen 4.0 x16	Contact Mellanox
1x 200 GbE	QSFP56	200/100 <sup>2</sup> /50/40/25/10/1	Gen 3.0 2x16 Socket Direct	MCX614105A-VCAT
			Gen 4.0 x16	Contact Mellanox
2x 200 GbE	QSFP56	200/100 <sup>2</sup> /50/40/25/10/1	Gen 3.0 2x16 Socket Direct	MCX614106A-VCAT
			Gen 4.0 x16	MCX613106A-VDAT



1. By default, the above products are shipped with a tall bracket mounted; a short bracket is included as an accessory.
2. 100GbE can be supported as either 4x25G NRZ or 2x50G PAM4 when using QSFP56.
3. Contact Mellanox for other supported options.

Table 2 - OCP 3.0 Small Form Factor

Max. Network Speed	Interface Type	Supported Ethernet Speed [GbE]	Host Interface [PCIe]	OPN
2x 200 GbE	QSFP56	200/100 <sup>2</sup> /50/40/25/10/1	Gen 4.0 x16	MCX613436A-VDAI
1x 200 GbE				Contact Mellanox



1. Above OPNs support a single host; contact Mellanox for OCP OPNs with Mellanox Multi-Host support.
2. 100GbE can be supported as either 4x25G NRZ or 2x50G PAM4 when using QSFP56.
3. Above OCP3.0 OPNs come with Internal Lock Brackets; Contact Mellanox for additional bracket types, e.g., Pull Tab or Ejector latch.