†

ConnectX®-6 Lx Ethernet SmartNIC
25GbE Performance at the Speed of Lite

ConnectX-6 Lx SmartNICs deliver scalability, high-performance, advanced security capabilities and accelerated networking with the best total cost of ownership for 25GbE deployments in cloud, telco, and enterprise data centers.

Providing up to two ports of 25GbE or a single-port of 50GbE connectivity, and PCIe Gen 3.0/4.0 x8 host connectivity, ConnectX-6 Lx is a member of Mellanox's world-class, award-winning, ConnectX family of network adapters. Continuing Mellanox's consistent innovation in networking, ConnectX-6 Lx provides agility and efficiency at every scale. ConnectX-6 Lx delivers cutting edge 25GbE performance and security for uncompromising data centers.

Wide Selection of SmartNICs

ConnectX-6 Lx SmartNICs are available in several form factors including low-profile PCIe and OCP 3.0 cards with SFP28 connectors for 10/25GbE applications, or QSFP28 for 50GbE applications. Low-profile PCIe cards are available with tall and short brackets, while OCP3.0 cards are available with either a pull tab or an internal lock bracket.

Best-in-Class SDN Acceleration

Mellanox's ASAP² - Accelerated Switch and Packet Processing® technology offloads the SDN data plane to the SmartNIC, accelerating performance and offloading the CPU in virtualized or containerized cloud data centers. Customers can accelerate their data centers with an SR-IOV or VirtIO interface while continuing to enjoy their SDN of choice.

ConnectX-6 Lx ASAP² rich feature set accelerates public and on-premises enterprise clouds, and boosts communication service providers (CSP) transition to NFV. ASAP² supports these communication service providers by enabling packet encapsulations, such as MPLS and GTP, along side cloud encapsulations, such as VXLAN, GENEVE and others.

Industry-leading RoCE

Following the Mellanox ConnectX tradition of industry-leading RoCE capabilities, ConnectX-6 Lx enables more scalable, resilient, and easy-to-deploy RoCE solutions – Zero Touch RoCE. ConnectX-6 Lx allows RoCE payloads to run seamlessly on existing networks without requiring network configuration (no PFC, no ECN) for simplified RoCE deployments. ConnectX-6 Lx ensures RoCE resiliency and efficiency at scale.

Secure Your Infrastructure

In an era where privacy of information is key and zero trust is the rule, ConnectX-6 Lx adapters offer a range of advanced built-in capabilities that bring infrastructure security down to every endpoint with unprecedented performance and scalability. ConnectX-6 Lx offers IPsec inline encryption/decryption acceleration. ASAP² connection-tracking hardware offload accelerates L4 firewall performance.

ConnectX-6 Lx also delivers supply chain protection with hardware Root-of-Trust (RoT) for Secure Boot as well as Secure Firmware Update using RSA cryptography and cloning-protection, via a device-unique key, to guarantee firmware authenticity.

HIGHLIGHTS

SmartNIC Portfolio

– 10/25/50 Gb/s Ethernet
– Various form factors:
  • PCIe low-profile
  • OCP 3.0 Small Form Factor (SFF)
– Connectivity options:
  • SFP28, QSFP28
– PCIe Gen 3.0/4.0 x8
– Crypto and non-crypto versions

Features & Applications

– Line speed message rate of 75Mpps
– Advanced RoCE
– ASAP² - Accelerated Switching and Packet Processing
– IPsec in-line crypto acceleration
– Overlay tunneling accelerations
– Stateful rule checking for connection tracking
– Hardware Root-of-Trust and secure firmware update
– Best-in-class PTP performance
– ODCC compatible

Solutions

– Enterprise data centers
– Cloud-native, Web 2.0, hyperscale
– Secured infrastructure
– Telco and Network Function Virtualization (NFV)

†For illustration only. Actual products may vary.

Features*

Network Interface
– 2 x 10/25GbE / 1 x 50GbE

Host Interface
– PCIe Gen 4.0, 3.0, 2.0, 1.1
– 16.0, 8.0, 5.0, 2.5 GT/s link rate
– 8 lanes of PCIe
– MSI/MSI-X mechanisms
– Advanced PCIe capabilities

Virtualization / Cloud Native
– Single Root IOV (SR-IOV) and VirtIO acceleration
  • Up to 512 VFs per port
  • 8 PFs
– Support for tunneling
  • Encap/decap of VXLAN, NVGRE, GENEVE, and more
  • Stateless offloads for overlay tunnels

Mellanox ASAP²
– SDN acceleration for:
  • Bare metal
  • Virtualization
  • Containers
– Full hardware offload for OVS data plane

– Flow update through RTE_Flow or TC_Flower
– OpenStack support
– Kubernetes support
– Rich classification engine (L2 to L4)
– Flex-Parser: user defined classification
– Hardware offload for:
  • Connection tracking (L4 firewall)
  • NAT
  • Header rewrite
  • Mirroring
  • Sampling
  • Flow aging
  • Hierarchal QoS
  • Flow-based statistics

Cyber Security
– Inline hardware IPsec encryption & decryption
  • AES-GCM 128/256 bit key
  • IPsec over RoCE
– Platform security
  • Hardware root-of-trust
  • Secure firmware update

Stateless Offloads
– TCP/UDP/IP stateless offload
– LSO, LRO, checksum offload
– Received Side Scaling (RSS) also on encapsulated packet
– Transmit Side Scaling (TSS)
– VLAN and MPLS tag insertion/stripping
– Receive flow steering

Advanced Timing & Synchronization
– Advanced PTP
  • IEEE 1588v2 (any profile)
  • PTP Hardware Clock (PHC) (UTC format)
  • Line rate hardware timestamp (UTC format)
– Time triggered scheduling
– PTP based packet pacing
– Time based SDN acceleration (ASAP²)

Storage Accelerations
– NVMe over Fabric offloads for target
– Storage protocols: iSER, NFSoRDMA, SMB Direct, NVMe-oF, and more

RDMA over Converged Ethernet
– RoCE v1/v2
– Zero-Touch RoCE: no ECN, no PFC
– RoCE over overlay networks
– IPsec over RoCE
– Selective repeat
– GPUDirect®
– Dynamically Connected Transport (DCT)
– Burst buffer offload

Management and Control
– SMBus 2.0
– Network Controller Sideband Interface (NC-SI)
– NC-SI, MCTP over SMBus and MCTP over PCIe - Baseboard Management Controller interface
– PLDM for Monitor and Control DSP0248
– PLDM for Firmware Update DSP026

Remote Boot
– Remote boot over Ethernet
– Remote boot over iSCSI
– UEFI support for x86 and Arm servers
– PXE boot

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

Standards*

– IEEE 802.3ae 10 Gigabit Ethernet
– 25/50 Ethernet Consortium 25G and 50G supporting all FEC modes
– IEEE 802.3by 25G supporting all FEC modes
– IEEE 802.3ad, 802.1AX Link Aggregation

– IEEE 802.3az Energy Efficient Ethernet (supports only "Fast-Wake" mode)
– IEEE 802.3ap based auto-negotiation and KR startup
– IEEE 802.1Q, 802.1P VLAN tags and priority
– IEEE 802.1Qaz (ETS)

– IEEE 802.1Qbb (PFC)
– IEEE 802.1Qbg
– IEEE 1588v2
– IEEE 1149.1 and IEEE 1149.6 JTAG
– PCI Express Gen 3.0 and 4.0

SmartNIC Portfolio & Ordering Information

Table 1 - PCIe HHHL Form Factor

Max Network Speed	Interface Type	Supported Ethernet Speeds [GbE]	Host Interface [PCIe]	OPNs Without Crypto	OPNs With Crypto(1)
1 x 25GbE	SFP28	25, 10, 1	Gen 4.0 x8	Contact Mellanox	Contact Mellanox
2 x 25GbE	SFP28	25, 10, 1	Gen 4.0 x8	MCX631102AN-ADAT	MCX631102AE-ADAT
1 x 50GbE	QSFP28	50, 25, 10, 1	Gen 4.0 x8	MCX631105AN-GDAT	MCX631105AE-GDAT

1. The above SmartNICs do not enforce Secure Boot by default, please contact Mellanox for models with Secure Boot enabled.
2. By default, the above products are shipped with a tall bracket mounted, and a short bracket is included as an accessory.
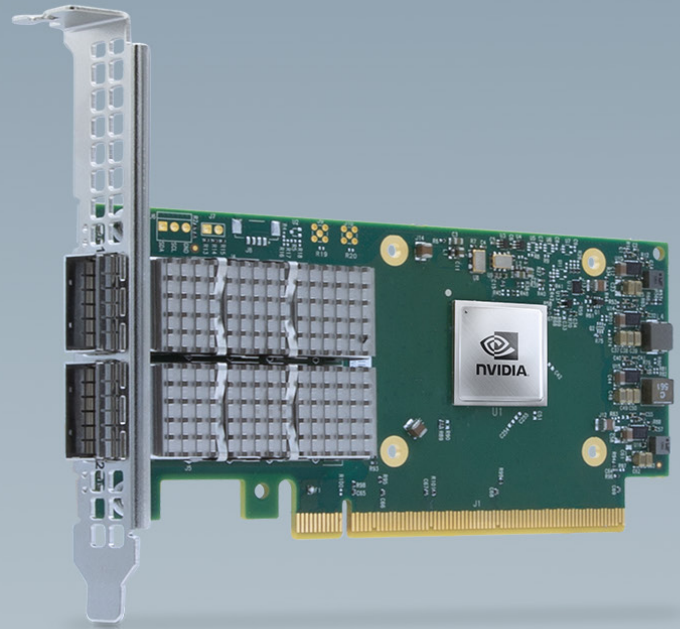
Table 2 - OCP 3.0 Small Form Factor

Max Network Speed	Interface Type	Supported Ethernet Speeds [GbE]	Host Interface [PCIe]	OPNs Without Crypto	OPNs With Crypto(1)
1 x 25GbE	SFP28	25, 10, 1	Gen 4.0 x8	Contact Mellanox	Contact Mellanox
2 x 25GbE	SFP28	25, 10, 1	Gen 4.0 x8	MCX631432AN-ADAB	MCX631432AE-ADAB
1 x 50GbE	QSFP28	50, 25, 10, 1	Gen 4.0 x8	MCX631435AN-GDAB	MCX631435AE-GDAB

1. The above SmartNICs do not enforce Secure Boot by default, please contact Mellanox for models with Secure Boot enabled.
2. The above OCP 3.0 OPNs are shipped with Thumbscrew (pull tab) brackets; contact Mellanox for additional bracket types, such as Internal Lock or Ejector latch.

† For illustration only. Actual products may vary.

350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com

60383PB
Ver 1.0

NVIDIA MELLANOX CONNECTX-6 DX
ETHERNET SMARTNIC

NVIDIA® Mellanox® ConnectX®-6 Dx SmartNIC is the industry's most secure and advanced cloud network interface card to accelerate mission-critical data-center applications, such as security, virtualization, SDN/NFV, big data, machine learning, and storage. The SmartNIC provides up to two ports of 100 Gb/s or a single-port of 200 Gb/s Ethernet connectivity and delivers the highest return on investment (ROI) of any smart network interface card.

ConnectX-6 Dx is a member of NVIDIA Mellanox's world-class, award-winning ConnectX series of network adapters powered by leading 50 Gb/s (PAM4) and 25/10 Gb/s (NRZ) SerDes technology and novel capabilities that accelerate cloud and data-center payloads.

SECURITY FROM ZERO TRUST TO HERO TRUST

In an era where privacy of information is key and zero trust is the rule, ConnectX-6 Dx adapters offer a range of advanced built-in capabilities that bring security down to the endpoints with unprecedented performance and scalability, including:

> Crypto – IPsec and TLS data-in-motion inline encryption and decryption offload, and AES-XTS block-level data-at-rest encryption and decryption offload.

> Probes & DoS Attack Protection – ConnectX-6 Dx enables a hardware-based L4 firewall by offloading stateful connection tracking through Mellanox ASAP[2] - Accelerated Switch and Packet Processing® .

> NIC Security – Hardware Root-of-Trust (RoT) Secure Boot and secure firmware update using RSA cryptography, and cloning-protection, via a device-unique secret key.

ADVANCED VIRTUALIZATION

ConnectX-6 Dx delivers another level of innovation to enable building highly efficient virtualized cloud data centers:

> Virtualization – Mellanox ASAP[2] technology for vSwitch/vRouter hardware offload delivers orders of magnitude higher performance vs. software-based solutions. ConnectX-6 Dx ASAP[2] offers both SR-IOV and VirtIO in-hardware offload capabilities, and supports up to 8 million rules.

> Advanced Quality of Service – Includes traffic shaping and classification-based data policing.

SmartNIC Portfolio

> 1/10/25/40/50/100/200 Gb/s Ethernet, PAM4/NRZ
> Various form factors:
> > PCIe low-profile
> > OCP 3.0 Small Form Factor (SFF)
> > OCP 2.0
> Connectivity options:
> > SFP28, SFP56, QSFP28, QSFP56, DSFP
> PCIe Gen 3.0/4.0 x16 host interface
> Multi-host and single-host flavors
> Crypto and non-crypto versions

Key Features

> Up to 200 Gb/s bandwidth
> Message rate of up to 215 Mpps
> Sub 0.8 usec latency
> Flexible programmable pipeline for new network flows
> Mellanox Multi-Host with advanced QoS
> ASAP[2] - Accelerated Switching and Packet Processing for virtual switches/routers
> Overlay tunneling technologies
> IPsec and TLS in-line crypto acceleration
> Block crypto acceleration for data-at-rest
> Hardware Root-of-Trust and secure firmware update
> Connection Tracking offload
> Advanced RoCE capabilities
> Best in class PTP for TSN applications
> GPUDirect® for GPU-to-GPU communication
> Host chaining technology for economical rack design
> Platform agnostic: x86, Power, Arm
> ODCC compatible

INDUSTRY-LEADING ROCE

Following the Mellanox ConnectX tradition of industry-leading RoCE capabilities, ConnectX-6 Dx adds another layer of innovation to enable more scalable, resilient and easy-to-deploy RoCE solutions.

> Zero Touch RoCE – Simplifying RoCE deployments, ConnectX-6 Dx allows RoCE payloads to run seamlessly on existing networks without requiring special configuration on the network (no PFC, no ECN). New features in ConnectX-6 Dx ensure resiliency and efficiency at scale of such deployments.

> Configurable Congestion Control – API to build user-defined congestion control algorithms, best serving various environments and RoCE and TCP/IP traffic patterns.

BEST-IN-CLASS PTP FOR TIME SENSITIVE APPLICATIONS

Mellanox offers a full IEEE 1588v2 PTP software solution as well as time sensitive related features called 5T45G. Mellanox PTP and 5T45G software solutions are designed to meet the most demanding PTP profiles. ConnectX-6 Dx incorporates an integrated Hardware Clock (PHC) that allows the device to achieve sub-20 usec accuracy while offering various timing related functions, including time-triggered scheduling or time-based SND accelerations (time based ASAP^2). Furthermore, 5T45G technology enables software applications to transmit front-haul (ORAN) compatible in high bandwidth. The PTP solution supports slave clock, master clock, and boundary clock.

Selected ConectX-6 Dx SmartNICs provide PPS-Out or PPS-In signals from designated SMA connectors.

EFFICIENT STORAGE SOLUTIONS

With its NVMe-oF target and initiator offloads, ConnectX-6 Dx brings further optimization to NVMe-oF, enhancing CPU utilization and scalability. Additionally, ConnectX-6 Dx supports hardware offload for ingress/egress of T10-DIF/PI/CRC32/CRC64 signatures, as well as AES-XTS encryption/decryption offload enabling user-based key management and a one-time-FIPS-certification approach.

WIDE SELECTION OF NICS

ConnectX-6 Dx SmartNICs are available in several form factors including low-profile PCIe, OCP2.0 and OCP3.0 cards, with various network connector types (SFP28/56, QSFP28/56, or DSFP). The ConnectX-6 Dx portfolio also provides options for Mellanox Multi-Host® and Mellanox Socket Direct® configurations.

Mellanox Multi-Host® connects multiple compute or storage hosts to a single interconnect adapter and enables designing and building new scale-out compute and storage racks. This enables better power and performance management, while reducing capital and operational expenses.

Mellanox Socket Direct® technology brings improved performance to multi-socket servers, by enabling each CPU in a multi-socket server to directly connect to the network through its dedicated PCIe interface. This enables data to bypass the QPI (UPI) and the other CPU, improving latency, performance and CPU utilization.

Solutions

> Cloud-native, Web 2.0, hyperscale
> Enterprise data-centers
> Cyber security
> Big data analytics
> Scale-out compute and storage infrastructure
> Telco and Network Function Virtualization (NFV)
> Cloud storage
> Machine Learning (ML) & Artificial Intelligence (AI)
> Media and Entertainment

FEATURES*

Network Interface
> 2 x 25/50/100 GbE; 1 x 200 GbE

Host Interface
> PCIe Gen 4.0, 3.0, 2.0, 1.1
> 16.0, 8.0, 5.0, 2.5 GT/s link rate
> 16 lanes of PCIe
> MSI/MSI-X mechanisms
> Advanced PCIe capabilities

Virtualization/Cloud Native
> Single Root IOV (SR-IOV) and VirtIO acceleration
  > Up to 1 K VFs per port
  > 8 PFs
> Support for tunneling
  > Encap/decap of VXLAN, NVGRE, Geneve, and more
  > Stateless offloads for Overlay tunnels

Mellanox ASAP²
> SDN acceleration for:
  > Bare metal
  > Virtualization
  > Containers
> Full hardware offload for OVS data plane
> Flow update through RTE_Flow or TC_Flower
> OpenStack support
> Kubernetes support
> Rich classification engine (L2 to L4)
> Flex-Parser: user defined classification
> Hardware offload for:
  > Connection tracking (L4 firewall)
  > NAT
  > Header rewrite
  > Mirroring
  > Sampling
  > Flow aging
  > Hierarchial QoS
  > Flow-based statistics

Cyber Security
> Inline hardware IPsec encryption & decryption
  > AES-GCM 128/256 bit key
  > IPsec over RoCE
> Inline hardware TLS encryption & decryption
  > AES-GCM 128/256 bit key
> Data-at-rest AES-XTS encryption & decryption
  > AES-GCM 128/256 bit key
> Platform security
  > Hardware root-of-trust
  > Secure firmware update

Stateless Offloads
> TCP/UDP/IP stateless offload
> LSO, LRO, checksum offload
> Receive Side Scaling (RSS) also on encapsulated packet
> Transmit Side Scaling (TSS)
> VLAN and MPLS tag insertion/stripping
> Receive flow steering

Advanced Timing & Synchronization
> Advanced PTP
  > IEEE 1588v2 (any profile)
  > PTP Hardware Clock (PHC) (UTC format)
  > 16 nsec accuracy
  > Line rate hardware timestamp (UTC format)
  > PPS In and configurable PPS Out
> Time triggered scheduling
> PTP based packet pacing
> Time based SDN acceleration (ASAP²)
> Time Sensitive Networking (TSN)

Storage Accelerations
> NVMe over Fabric offloads for target
> Storage protocols: iSER, NFSoRDMA, SMB Direct, NVMe-oF, and more
> T-10 Dif/Signature Handover

RDMA over Converged Ethernet (RoCE)
> RoCE v1/v2
> Zero Touch RoCE: no ECN, no PFC
> RoCE over overlay networks
> IPsec over RoCE
> Selective repeat
> Programmable congestion control interface
> GPUDirect®
> Dynamically connected transport (DCT)
> Burst buffer offload

Management and Control
> NC-SI, MCTP over SMBus and MCTP over PCIe - Baseboard Management Controller interface, NCSI over RBT in OCP 2.0/3.0 cards
> PLDM for Monitor and Control DSP0248
> PLDM for Firmware Update DSP026
> I²C interface for device control and configuration

Remote Boot
> Remote boot over Ethernet
> Remote boot over iSCSI
> UEFI support for x86 and Arm servers
> PXE boot

STANDARDS*

> IEEE 802.3bs, 200 Gigabit Ethernet
> IEEE 802.3cd, 50, 100 and 200 Gigabit Ethernet
> IEEE 802.3bj, 802.3bm 100 Gigabit Ethernet
> IEEE 802.3by, 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 (supports only "Fast-Wake" mode)
> 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 802.1Qaz (ETS)
> IEEE 802.1Qbb (PFC)
> IEEE 802.1Qbg
> 25/50 Ethernet Consortium "Low Latency FEC" for 50GE/100GE/200GE PAM4 links
> PCI Express Gen 3.0 and 4.0

SMARTNIC PORTFOLIO & ORDERING INFORMATION

PCIE HHHL FORM FACTOR

Max Network Speed	Interface Type	Supported Ethernet Speeds [GbE]	Host Interface [PCIe]	Ordering Part Number (OPN)		
				No Crypto, No Secure Boot	With Crypto[1], No Secure Boot	With Crypto[1], With Secure Boot
2 x 25 GbE	SFP28	1/10/25	Gen 4.0 x8	MCX621102AN-ADAT	MCX621102AE-ADAT	MCX621102AC-ADAT
			Gen 4.0 x16	MCX623102AN-ADAT	Contact Mellanox	MCX623102AC-ADAT
2 x 50 GbE	SFP56	1/10/25/50	Gen 4.0 x16	MCX623102AN-GDAT	MCX623102AE-GDAT	MCX623102AC-GDAT
1 x 100 GbE	QSFP56	1/10/25/40/50[2]/100[3]	Gen 4.0 x16	MCX623105AN-CDAT	Contact Mellanox	Contact Mellanox
2 x 100 GbE	QSFP56	1/10/25/40/50[2]/100[3]	Gen 4.0 x16	MCX623106AN-CDAT	MCX623106AE-CDAT	MCX623106AC-CDAT
	QSFP56 + PPS In/Out SMAs	1/10/25/40/50[2]/100[3]	Gen 4.0 x16	MCX623106PN-CDAT	MCX623106PE-CDAT	MCX623106PC-CDAT
	DSFP	1/10/25/50[2]/100	Gen 4.0 x16	Contact Mellanox	Contact Mellanox	Contact Mellanox
1 x 200 GbE	QSFP56	10/25/40/50[2]/100[3]/200	Gen 4.0 x16	MCX623105AN-VDAT	MCX623105AE-VDAT	MCX623105AC-VDAT
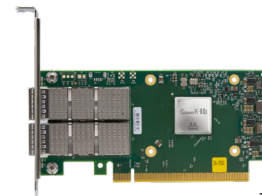
1 Use Crypto enabled cards to utilize IPsec/TLS/AES-XTS encryption/decryption hardware offload.
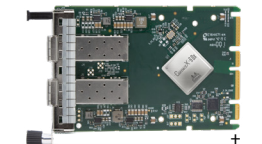2 50G can be supported as either 2x25G NRZ or 1x50G PAM4 when using QSFP56.
3 100G can be supported as either 4x25G NRZ or 2x50G PAM4 when using QSFP56.

By default, the above products are shipped with a tall bracket mounted; a short bracket is included as an accessory.

†

OCP 3.0 SMALL FORM FACTOR

Max Network Speed	Interface Type	Supported Ethernet Speeds [GbE]	Host Interface [PCIe]	Ordering Part Number (OPN)		
				No Crypto, No Secure Boot	With Crypto[1], No Secure Boot	With Crypto[1], With Secure Boot
2 x 25 GbE	SFP28	1/10/25	Gen 4.0 x16	MCX623432AN-ADAB	MCX623432AE-ADAB	MCX623432AC-ADAB
2 x 50 GbE	SFP56	1/10/25/50	Gen 4.0 x16	MCX623432AN-GDAB	Contact Mellanox	MCX623432AC-GDAB
2 x 100 GbE	QSFP56	1/10/25/40/50[2]/100[3]	Gen 4.0 x16	MCX623436AN-CDAB	MCX623436AE-CDAB	MCX623436AC-CDAB
1 x 200 GbE	QSFP56	1/10/25/40/50[2]/100[3]/200	Gen 4.0 x16	MCX623435AN-VDAB	Contact Mellanox	MCX623435AC-VDAB

1 Use Crypto enabled cards to utilize IPsec/TLS/AES-XTS encryption/decryption hardware offload.
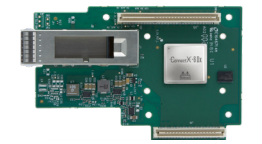2 50G can be supported as either 2x25G NRZ or 1x50G PAM4 when using QS·FP56.
3 100G can be supported as either 4x25G NRZ or 2x50G PAM4 when using QSFP 56.

These OPNs are Single Host; contact Mellanox for OCP OPNs with Mellanox Multi-Host support.
The above OCP3.0 OPNs come with Thumbscrew (pull tab) brackets; contact Mellanox for additional bracket types, such as Internal Lock or Ejector latch.

†

OCP 2.0 FORM FACTOR

Max Network Speed	Interface Type	Supported Ethernet Speeds [GbE]	Host Interface [PCIe]	Ordering Part Number (OPN)		
				No Crypto, No Secure Boot	With Crypto[1], No Secure Boot	With Crypto[1], With Secure Boot
1 x 100 GbE	QSFP56	1/10/25/40/50[2]/100[3]	Gen 4.0 x16	MCX623405AN-CDAN	Contact Mellanox	MCX623405AC-CDAN
1 x 200 GbE	QSFP56	1/10/25/40/50[2]/100[3]/200	Gen 4.0 x16	Contact Mellanox	Contact Mellanox	Contact Mellanox

1 Use Crypto enabled cards to utilize IPsec/TLS/AES-XTS encryption/decryption hardware offload.
2 50G can be supported as either 2x25G NRZ or 1x50G PAM4 when using QSFP56.
3 100G can be supported as either 4x25G NRZ or 2x50G PAM4 when using QSFP56.

These OPNs are Single Host; contact Mellanox for OCP OPNs with Mellanox Multi-Host or Mellanox Socket Direct support.
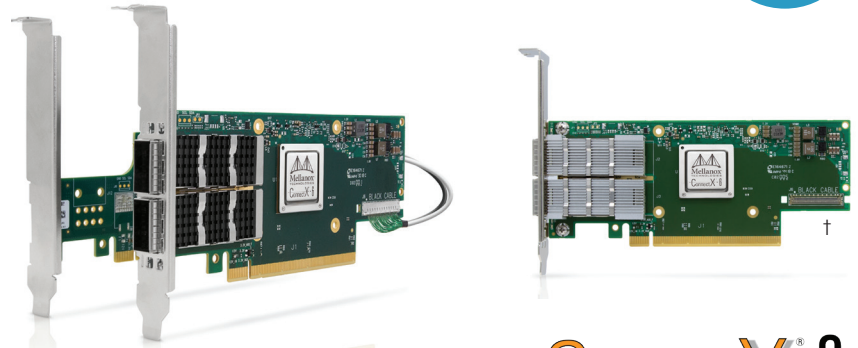ConnectX-6 Dx PCIe 4.0 cards are backward compatible.

†

† For illustration only. Actual products may vary.

Learn more at **www.mellanox.com/products/ethernet/connectx-smartnic**

ConnectX®-6 EN Card

200GbE Ethernet Adapter Card

ConnectX·6

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 ASAP² - Mellanox Accelerated Switch and Packet Processing® 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 packet pacing 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

†For illustration only. Actual products may vary.

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[2]
– 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[2]

– 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[2]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	100^2/50/40/25/10/1	Gen 3.0 2x16 Socket Direct	MCX614106A-CCAT
			Gen 4.0 x16	Contact Mellanox
	SFP-DD			Contact Mellanox
1x 200 GbE	QSFP56	200/100^2/50/40/25/10/1	Gen 3.0 2x16 Socket Direct	MCX614105A-VCAT
			Gen 4.0 x16	Contact Mellanox
2x 200 GbE	QSFP56	200/100^2/50/40/25/10/1	Gen 3.0 2x16 Socket Direct	MCX614106A-VCAT
			Gen 4.0 x16	MCX613106A-VDAT
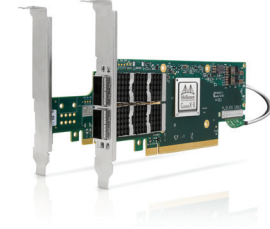
1. By default, the above products are shipped with a tall bracket mounted; a short bracket is included as an accessory.

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	200/100^2/50/40/25/10/1	Gen 4.0 x16	MCX613436A-VDAI
1x 200 GbE				Contact Mellanox
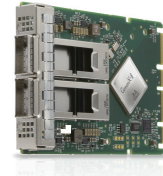
1. Above OPNs support a single host; contact Mellanox for OCP OPNs with Mellanox Multi-Host support.

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.

350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com

53724PB
Rev 2.1

ConnectX®-5 EN Card

Up to 100Gb/s Ethernet Adapter Cards

Intelligent RDMA-enabled, single and dual-port network adapter with advanced application offload capabilities for Web 2.0, Cloud, Storage, and Telco platforms

ConnectX-5 Ethernet adapter cards provide high performance and flexible solutions with up to two ports of 100GbE connectivity, 750ns latency, up to 200 million messages per second (Mpps), and a record-setting 197Mpps when running an open source Data Path Development Kit (DPDK) PCIe (Gen 4.0). For storage workloads, ConnectX-5 delivers a range of innovative accelerations, such as Signature Handover (T10-DIF) in hardware, an embedded PCIe Switch, and NVMe over Fabric target offloads. ConnectX-5 adapter cards also bring advanced Open vSwitch offloads to telecommunications and cloud data centers to drive extremely high packet rates and throughput with reduced CPU resource consumption, thus boosting data center infrastructure efficiency.
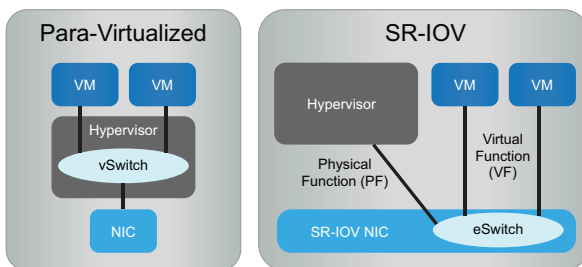
ConnectX-5 adapter cards are available for PCIe Gen 3.0 and Gen 4.0 servers and provide support for 1, 10, 25, 40, 50 and 100 GbE speeds in stand-up PCIe cards, OCP 2.0, and OCP 3.0 form factors. ConnectX-5 cards also offer advanced Mellanox Multi-Host® and Mellanox Socket Direct® technologies.

Cloud and Web 2.0 Environments

ConnectX-5 adapter cards enable data center administrators to benefit from better server utilization and reduced costs, power usage, and cable complexity, allowing for more virtual appliances, virtual machines (VMs) and tenants to co-exist on the same hardware.

Supported vSwitch/vRouter offload functions include:

- Overlay Networks (e.g., VXLAN, NVGRE, MPLS, GENEVE, and NSH) header encapsulation & decapsulation.

- Stateless offloads of inner packets and packet headers' re-write, enabling NAT functionality and more.

- Flexible and programmable parser and match-action tables, which enable hardware offloads for future protocols.

- SR-IOV technology, providing dedicated adapter resources, guaranteed isolation and protection for virtual machines (VMs) within the server.

- Network Function Virtualization (NFV), enabling a VM to be used as a virtual appliance. The full data-path operation offloads, hairpin hardware capability and service chaining enables data to be handled by the virtual appliance, with minimum CPU utilization.

Para-Virtualized

VM | VM
Hypervisor
vSwitch
NIC

SR-IOV

Hypervisor
Physical Function (PF)
VM | VM
Virtual Function (VF)
SR-IOV NIC | eSwitch

HIGHLIGHTS

FEATURES

- Tag matching and rendezvous offloads
- Adaptive routing on reliable transport
- Burst buffer offloads for background checkpointing
- NVMe over Fabric offloads
- Backend switch elimination by host chaining
- Embedded PCIe switch
- Enhanced vSwitch/vRouter offloads
- Flexible pipeline
- RoCE for overlay networks
- PCIe Gen 4.0 support
- RoHS compliant
- ODCC compatible
- Various form factors available

BENEFITS

- Up to 100Gb/s connectivity per port
- Industry-leading throughput, low latency, low CPU utilization and high message rate
- Innovative rack design for storage and Machine Learning based on Host Chaining technology
- Smart interconnect for x86, Power, Arm, and GPU-based compute & storage platforms
- Advanced storage capabilities including NVMe over Fabric offloads
- Intelligent network adapter supporting flexible pipeline programmability
- Cutting-edge performance in virtualized networks including Network Function Virtualization (NFV)
- Enabler for efficient service chaining capabilities
- Efficient I/O consolidation, lowering data center costs and complexity

†For illustration only. Actual products may vary.

Cloud and Web 2.0 customers developing platforms on Software Defined Network (SDN) environments are leveraging their servers' Operating System Virtual-Switching capabilities to achieve maximum flexibility. Open vSwitch (OvS) is an example of a virtual switch that allows Virtual Machines to communicate with each other and with the outside world. Traditionally residing in the hypervisor where switching is based on twelve-tuple matching onflows, the virtual switch, or virtual router software-based solution, is CPU-intensive. This can negatively affect system performance and prevent the full utilization of available bandwidth.

Mellanox ASAP² - Accelerated Switching and Packet Processing® technology enables offloading the vSwitch/vRouter by handling the data plane in the NIC hardware, without modifying the control plane. This results in significantly higher vSwitch/vRouter performance without the associated CPU load.

Additionally, intelligent ConnectX-5's flexible pipeline capabilities, including flexible parser and flexible match-action tables, are programmable, enabling hardware offloads for future protocols.

Storage Environments

NVMe storage devices are gaining popularity by offering very fast storage access. The evolving NVMe over Fabric (NVMe-oF) protocol leverages the RDMA connectivity for remote access. ConnectX-5 offers further enhancements by providing NVMe-oF target offloads, enabling very efficient NVMe storage access with no CPU intervention, and thus improving performance and reducing latency.

Block Device / Native Application
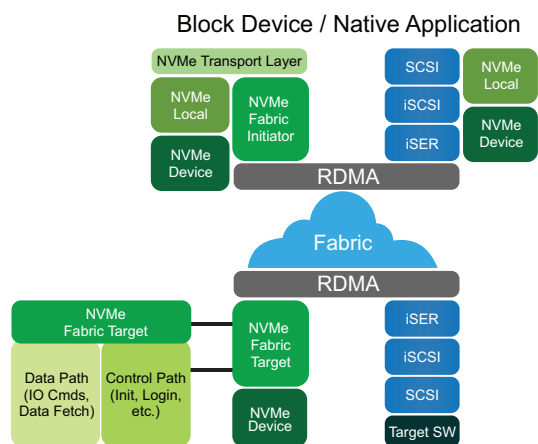
The embedded PCIe switch enables customers to build standalone storage or Machine Learning appliances. As with earlier generations of ConnectX adapters, standard block and file access protocols leverage RoCE for high-performance storage access. A consolidated compute and storage network achieves significant cost-performance advantages over multi-fabric networks.

ConnectX-5 enables an innovative storage rack design, Host Chaining, which enables different servers to interconnect without involving the Top of the Rack (ToR) switch. Leveraging Host Chaining, ConnectX-5 lowers the data center's total cost of ownership (TCO) by reducing CAPEX (cables, NICs, and switch port expenses). OPEX is also reduced by cutting down on switch port management and overall power usage.

Traditional Storage Connectivity → Eliminating Backend Switch → Host Chaining for Storage Backend
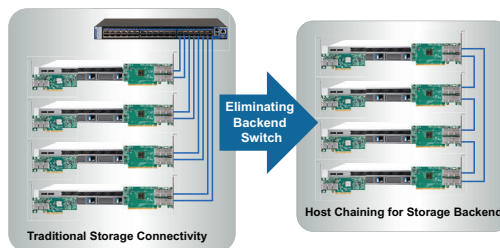
Telecommunications

Telecommunications service providers are moving towards disaggregation, server virtualization, and orchestration as key tenets to modernize their networks. Likewise, they're also moving towards Network Function Virtualization (NFV), which enables the rapid deployment of new network services. With this move, proprietary dedicated hardware and software, which tend to be static and difficult to scale, are being replaced with virtual machines running on commercial off-the-shelf (COTS) servers.

For telecom service providers, choosing the right networking hardware is critical to achieving a cloud-native NFV solution that is agile, reliable, fast and efficient. Telco service providers typically leverage virtualization and cloud technologies to better achieve agile service delivery and efficient scalability; these technologies require an advanced network infrastructure to support higher rates of packet processing. However, the resultant east-west traffic causes numerous interrupts as I/O traverses from kernel to user space, eats up CPU cycles and decreases packet performance. Particularly sensitive to delays are voice and video applications which often require less than 100ms of latency.

ConnectX-5 adapter cards drive extremely high packet rates, increased throughput and drive higher network efficiency through the following technologies; Open vSwitch Offloads (OvS), OvS over DPDK or ASAP², Network Overlay Virtualization, SR-IOV, and RDMA. This allows for secure data delivery through higher-performance offloads, reducing CPU resource utlization, and boosting data center infrastructure efficiency. The result is a much more responsive and agile network capable of rapidly deploying network services.

Wide Selection of Adapter Cards

ConnectX-5 Ethernet adapter cards are available in several form factors including: low-profile stand-up PCIe, OCP 2.0 Type 1 and Type 2, and OCP 3.0 Small Form Factor. (See the portfolio on the last page.)

Mellanox Multi-Host® technology allows multiple hosts to be connected into a single adapter by separating the PCIe interface into multiple and independent interfaces.

The portfolio also offers Mellanox Socket-Direct® configurations that enable servers without x16 PCIe slots to split the card's 16-lane PCIe bus into two 8-lane buses on dedicated cards connected by a harness.

Host Management

Host Management includes 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
– PCIe Gen 3.0, 1.1 and 2.0 compatible
– 2.5, 5.0, 8.0, 16.0 GT/s link rate
– Auto-negotiates to x16, x8, x4, x2, or x1 lane(s)
– PCIe Atomic
– TLP (Transaction Layer Packet) Processing Hints (TPH)
– Embedded PCIe Switch: Up to 8 bifurcations

– PCIe switch Downstream Port Containment (DPC) enablement for PCIe hot-plug
– Access Control Service (ACS) for peer-to-peer secure communication
– Advance Error Reporting (AER)
– Process Address Space ID (PASID) Address Translation Services (ATS)
– IBM CAPI v2 support (Coherent Accelerator Processor Interface)
– Support for MSI/MSI-X mechanisms

Operating Systems/Distributions*
– RHEL/CentOS
– Windows
– FreeBSD
– VMware
– OpenFabrics Enterprise Distribution (OFED)
– OpenFabrics Windows Distribution (WinOF-2)

Connectivity
– Interoperability with Ethernet switches (up to 100GbE)
– Passive copper cable with ESD protection
– Powered connectors for optical and active cable support

Features*

Ethernet
– 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
– Extended Reliable Connected transport (XRC)
– Dynamically Connected Transport (DCT)
– Enhanced Atomic operations
– Advanced memory mapping support, allowing user mode registration and remapping of memory (UMR)
– 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 applications
– Open VSwitch (OVS) offload using ASAP²
  • Flexible match-action flow tables
  • Tunneling encapsulation/ de-capsulation
– Intelligent interrupt coalescence
– Header rewrite supporting hardware offload of NAT router

Storage Offloads
– 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

Hardware-Based I/O Virtualization - Mellanox ASAP²
– Single Root IOV
– Address translation and protection
– VMware NetQueue support
  • SR-IOV: Up to 512 Virtual Functions
  • SR-IOV: Up to 8 Physical Functions per host
– Virtualization hierarchies (e.g., NPAR when enabled)
  • Virtualizing Physical Functions on a physical port
  • SR-IOV on every Physical Function
– Configurable and user-programmable QoS
– Guaranteed QoS for VMs

Management and Control
– NC-SI over MCTP over SMBus and NC-SI over 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
– I²C 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.

** When using Mellanox Socket Direct or Mellanox Multi-Host in virtualization or dual-port use cases, some restrictions may apply. For further details, contact Mellanox Customer Support.

Standards*

– IEEE 802.3cd, 50,100 and 200 Gigabit Ethernet
– 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 (supports only "Fast-Wake" mode)
– 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
– 25G/50G Ethernet Consortium "Low Latency FEC" for 50/100/200GE PAM4 links
– PCI Express Gen 3.0 and 4.0

Adapter Card Portfolio & Ordering Information

Table 1 - PCIe HHL Form Factor

Max Network Speed	Interface Type	Supported Ethernet Speeds (GbE)	Host Interface	Additional Features	OPN
2x **25**GbE	SFP28	25,10,1	PCIe 3.0 x8		MCX512A-ACAT
			PCIe 3.0 x8	UEFI Enabled (x86/Arm)	MCX512A-ACUT
			PCIe 4.0 x8	ConnectX-5 Ex	MCX512A-ADAT
			PCIe 3.0 x16		MCX512F-ACAT
			PCIe 3.0 x16	Enhanced Host Management	MCX512F-ACHT
2x **40**GbE	QSFP28	40,25,10,1	PCIe 4.0 x16	ConnectX-5 Ex	MCX516A-BDAT
1x **50**GbE	QSFP28	50,40,25,10,1	PCIe 3.0 x16		MCX515A-GCAT
2x **50**GbE	QSFP28	50,40,25,10,1	PCIe 3.0 x16		MCX516A-GCAT
1x **100**GbE	QSFP28	100,50,40,25,10,1	PCIe 3.0 x16		MCX515A-CCAT
			PCIe 3.0 x16	UEFI Enabled (x86/Arm)	MCX515A-CCUT
2x **100**GbE	QSFP28	100,50,40,25,10,1	PCIe 3.0 x16		MCX516A-CCAT
			PCIe 3.0 x16	Enhanced Host Management	MCX516A-CCHT
			PCIe 4.0 x16	ConnectX-5 Ex	MCX516A-CDAT

Notes: All tall-bracket adapters are shipped with the tall bracket mounted and a short bracket as an accessory.
For Mellanox Socket-Direct models, refer to the *ConnectX-5 VPI Mellanox Socket Direct Product Brief.*

Table 2 - OCP 2.0 Type 1 Form Factor

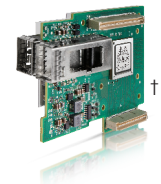| Max Network Speed	Interface Type	Supported Ethernet Speeds (GbE)	Host Interface	Additional Features	OPN
2x **25**GbE	SFP28	25,10,1	PCIe 3.0 x16		MCX542A-ACAN
			PCIe 3.0 x8		MCX542B-ACAN
			PCIe 3.0 x8	UEFI Enabled	MCX542B-ACUN
1x **100**GbE	QSFP28	100,50,40,25,10,1	PCIe 3.0 x16	UEFI Enabled	MCX545B-CCUN

Table 3 - OCP 2.0 Type 2 Form Factor

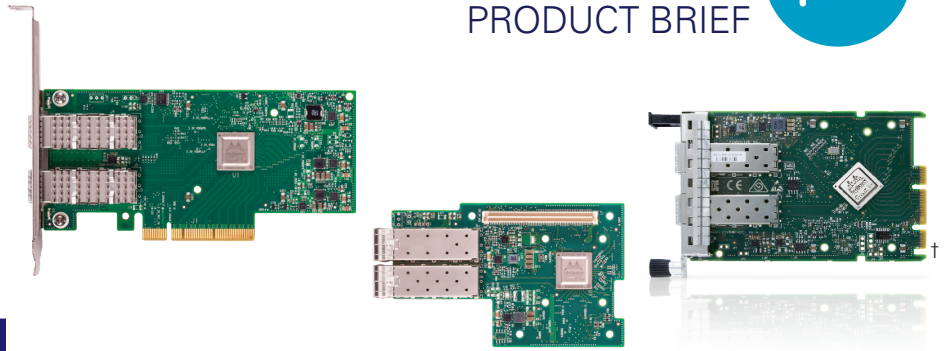| Max Network Speed	Interface Type	Supported Ethernet Speeds (GbE)	Host Interface	Additional Features	OPN
1x **100**GbE	QSFP28	100,50,40,25,10,1	PCIe 3.0 x16		MCX545A-CCAN
			PCIe 3.0 x16	UEFI Enabled	MCX545A-CCUN
2x **100**GbE	QSFP28	100,50,40,25,10,1	PCIe 4.0 x16	ConnectX-5 Ex	MCX546A-CDAN
			PCIe 4.0 x16	ConnectX-5 Ex, Mellanox Multi-Host	MCX546M-CDAN

Table 4 - OCP 3.0 Small Form Factor

Max Network Speed	Interface Type	Supported Ethernet Speeds (GbE)	Host Interface	Additional Features	OPN
2x **25**GbE	SFP28	25,10,1	PCIe 3.0 x16	Thumbscrew (pull tab) Bracket	MCX562A-ACAB
			PCIe 3.0 x16	Internal Lock Bracket	MCX562A-ACAI
2x **50**GbE	QSFP28	50,40,25,10,1	PCIe 4.0 x16	ConnectX-5 Ex, Mellanox Multi-Host, Internal Lock Bracket	MCX566M-GDAI
1x **100**GbE	QSFP28	100,50,40,25,10,1	PCIe 3.0 x16	Thumbscrew (pull tab) Bracket	MCX565A-CCAB
			PCIe 4.0 x16	ConnectX-5 Ex, Internal Lock Bracket	MCX565M-CDAI
			PCIe 4.0 x16	ConnectX-5 Ex, Mellanox Multi-Host, Thumbscrew (pull tab) Bracket	MCX565M-CDAB
2x **100**GbE	QSFP28	100,50,40,25,10,1	PCIe 3.0 x16	Internal Lock Bracket	MCX566A-CCAI
			PCIe 4.0 x16	ConnectX-5 Ex, Thumbscrew (pull tab) Bracket	MCX566A-CDAB
			PCIe 4.0 x16	ConnectX-5 Ex, Internal Lock Bracket	MCX566A-CDAI

Note: Mellanox OCP 3.0 cards that support Mellanox Multi-Host also support Mellanox Socket Direct.

General Note: ConnectX-5 Ex is an enhanced performance version that supports PCIe Gen 4.0 and higher throughput and lower latency.

350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com

† For illustration only. Actual products may vary.

53166PB
Rev 2.8

ConnectX®-4 Lx EN Card

Up to 50Gb/s Ethernet Adapter Cards

ConnectX·4 Lx

1/10/25/40/50 Gigabit Ethernet adapter cards supporting RDMA, Overlay Network Encapsulation/Decapsulation and more

ConnectX-4 Lx EN adapter cards with 1/10/25/40/50 Gb/s Ethernet connectivity addresses virtualized infrastructure challenges for today's demanding markets and applications. Providing true hardware-based I/O isolation with unmatched scalability and efficiency, ConnectX-4 Lx EN provides a cost-effective and flexible Ethernet adapter solution for Web 2.0, cloud, data analytics, database, and storage platforms.

With the exponential increase in usage of data and the creation of new applications, the demand for the highest throughput, lowest latency, virtualization and sophisticated data acceleration engines continues to rise. ConnectX-4 Lx EN adapter cards enable data centers to leverage leading interconnect adapters for increasing their operational efficiency, improving servers' utilization, maximizing applications productivity, while reducing total cost of ownership (TCO).

ConnectX-4 Lx EN adapter cards provide a combination of 1, 10, 25, 40, and 50 GbE bandwidth, sub-microsecond latency and a 75 million packets per second message rate. They include native hardware support for RDMA over Converged Ethernet (RoCE), Ethernet stateless offload engines, Overlay Networks, GPUDirect® technology and Mellanox Multi-Host® technology.

Mellanox Multi-Host® Technology

Innovative Mellanox Multi-Host technology enables data centers to design and build scale-out heterogeneous compute and storage racks, with direct connectivity between compute elements and the network. Significantly improving cost savings, flexibility, and total cost of ownership, Mellanox Multi-Host technology provides better power and performance, while achieving maximum data processing and data transfer at minimum capital and operational expenses.

Mellanox Multi-Host works by allowing multiple hosts to connect into a single interconnect adapter, by separating the adapter PCIe interface into several independent PCIe interfaces. Each interface connects to a separate host CPU—with no performance degradation. Reducing data center CAPEX and OPEX, Mellanox Multi-Host slashes switch port management and power usage by reducing the number of cables, NICs and switch ports required by four independent servers, from four to one of each. Additional features & benefits of Mellanox Multi-Host technology:

- Enables IT managers to remotely control the configuration and power state of each host individually; guaranteeing host security and isolation, the management of one host does not affect host traffic performance nor the management of other hosts.

- Lowering total cost of ownership (TCO), Mellanox Multi-Host uses a single BMC, with independent NC-SI/MCTP management channels for each of the managed hosts.

- Mellanox Multi-Host also supports a heterogeneous data center architecture; the various hosts connected to the single adapter can be x86, Power, GPU, Arm or FPGA, thereby removing any limitations in passing data or communicating between compute elements.

HIGHLIGHTS

FEATURES
- 1/10/25/40/50 Gb/s speeds
- Single and dual-port options
- Virtualization
- Low latency RDMA over Converged Ethernet (RoCE)
- Mellanox Multi-Host technology connects up to 4 independent hosts
- CPU offloading of transport operations
- Application offloading
- Mellanox PeerDirect® communication acceleration
- Hardware offloads for NVGRE, VXLAN and GENEVE encapsulated traffic
- End-to-end QoS and congestion control
- Hardware-based I/O virtualization
- RoHS compliant
- ODCC compatible
- Various form factors available

BENEFITS
- High performance boards for applications requiring high bandwidth, low latency and high message rate
- Industry leading throughput and latency for Web 2.0, Cloud and Big Data applications
- Smart interconnect for x86, Power, ARM, and GPU-based compute and storage platforms
- Cutting-edge performance in virtualized overlay networks
- Efficient I/O consolidation, lowering data center costs and complexity
- Virtualization acceleration
- Power efficiency

†For illustration only. Actual products may vary.

Wide Selection of Ethernet Adapter Cards

ConnectX-4 Lx EN adapter cards offer a cost-effective Ethernet adapter solution for 1, 10, 25, 40 and 50 Gb/s Ethernet speeds, enabling seamless networking, clustering, or storage. The adapter reduces application runtime, and offers the flexibility and scalability to make infrastructure run as efficiently and productively as possible.

ConnectX-4 Lx Ethernet adapter cards are available in several form factors including: low-profile stand-up PCIe, OCP 2.0 Type 1, OCP 2.0 Type 2, and OCP 3.0 small form factor. (See the portfolio on the last page.)

I/O Virtualization

ConnectX-4 Lx EN SR-IOV technology provides dedicated adapter resources and guaranteed isolation and protection for virtual machines (VMs) within the server. I/O virtualization with ConnectX-4 Lx EN gives data center administrators better server utilization while reducing cost, power, and cable complexity, allowing more virtual machines and more tenants on the same hardware.

Overlay Networks

In order to better scale their networks, data center operators often create overlay networks that carry traffic from individual virtual machines over logical tunnels in encapsulated formats such as NVGRE and VXLAN. While this solves network scalability issues, it hides the TCP packet from the hardware offloading engines, placing higher loads on the host CPU. ConnectX-4 Lx EN effectively addresses this by providing advanced NVGRE, VXLAN and GENEVE hardware offloading engines that encapsulate and de-capsulate the overlay protocol headers, enabling the traditional offloads to be performed on the encapsulated traffic for these and other tunneling protocols (GENEVE, MPLS, QinQ, and so on). With ConnectX-4 Lx EN, data center operators can achieve native performance in the new network architecture.

RDMA over Converged Ethernet (RoCE)

ConnectX-4 Lx EN supports RoCE specifications delivering low-latency and high- performance over Ethernet networks. Leveraging data center bridging (DCB) capabilities as well as ConnectX-4 Lx EN advanced congestion control hardware mechanisms, RoCE provides efficient low-latency RDMA services over Layer 2 and Layer 3 networks.

Mellanox PeerDirect®

Mellanox PeerDirect communication provides high efficiency RDMA access by eliminating unnecessary internal data copies between components on the PCIe bus (for example, from GPU to CPU), and therefore significantly reduces application run time. ConnectX-4 Lx EN advanced acceleration technology enables higher cluster efficiency and scalability to tens of thousands of nodes.

Storage Acceleration

Storage applications will see improved performance with the higher bandwidth ConnectX-4 Lx EN delivers. Moreover, standard block and file access protocols can leverage RoCE for high-performance storage access. A consolidated compute and storage network achieves significant cost-performance advantages over multi-fabric networks.

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 Monitoring and Control DSP0248 and PLDM for Firmware Update DSP0267.

Software Support

All Mellanox adapter cards are supported by Windows, Linux distributions, VMware, FreeBSD, and Citrix XENServer. ConnectX-4 Lx EN supports various management interfaces and has a rich set of tools for configuration and management across operating systems.

Additionally, ConnectX-4 Lx EN provides the option for a secure firmware update check using digital signatures to prevent remote attackers from uploading malicious firmware images; this ensures that only officially authentic images produced by Mellanox can be installed, regardless whether the source of the installation is the host, the network, or a BMC.

Features*

Ethernet
– 50GbE / 40GbE / 25GbE / 10GbE / 1GbE
– IEEE 802.3bj, 802.3bm 100 Gigabit Ethernet
– 25G Ethernet Consortium 25, 50 Gigabit Ethernet
– 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
– Proprietary Ethernet protocols (20/40GBASE-R2, 50GBASE-R4)
– 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)

Mellanox Multi-Host
– Up to 4 separate PCIe interfaces to 4 independent hosts
– Two PCIe x4 to two hosts, or four PCIe x4 to four hosts, or four PCIe x2 to four hosts

– Independent NC-SI SMBus interfaces
– Independent stand-by and wake-on-LAN signals

Enhanced Features
– Hardware-based reliable transport
– Collective operations offloads
– Vector collective operations offloads
– PeerDirect RDMA (aka GPUDirect® communication acceleration)
– 64/66 encoding
– Extended Reliable Connected transport (XRC)
– Dynamically Connected transport (DCT)
– Enhanced Atomic operations
– Advanced memory mapping support, allowing user mode registration and remapping of memory (UMR)
– On demand paging (ODP) – registration free RDMA memory access

CPU Offloads
– RDMA over Converged Ethernet (RoCE)
– TCP/UDP/IP stateless offload
– LSO, LRO, checksum offload
– RSS (can be done on encapsulated packet), TSS, VLAN insertion/ stripping, Receive flow steering
– Intelligent interrupt coalescence

Overlay Networks
– Stateless offloads for overlay networks and tunneling protocols
– Hardware offload of encapsulation and decapsulation of NVGRE and VXLAN overlay networks

Hardware-Based I/O Virtualization
– Single Root IOV
– Multi-function per port
– Address translation and protection
– Multiple queues per virtual machine
– Enhanced QoS for vNICs
– VMware NetQueue support

Virtualization
– SR-IOV: Up to 256 Virtual Functions
– SR-IOV: Up to 8 Physical Functions per port
– Virtualization hierarchies (e.g. NPAR)
  • Virtualizing Physical Functions on a physical port
  • SR-IOV on every Physical Function
– Ingress and egress QoS levels
– Guaranteed QoS for VMs

Protocol Support
– OpenMPI, IBM PE, OSU MPI (MVAPICH/2), Intel MPI
– Platform MPI, UPC, Open SHMEM
– TCP/UDP, MPLS, VxLAN, NVGRE, GENEVE
– iSER, NFS RDMA, SMB Direct
– uDAPL

Management and Control Interfaces
– NC-SI over MCTP over SMBus and NC-SI over MCTP over PCIe - Baseboard Management Controller interface
– SDN management interface for managing the eSwitch
– I²C 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
– PXE and UEFI

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

Compatibility

PCI Express Interface
– PCIe Gen 3.0 compliant, 1.1 and 2.0 compatible
– 2.5, 5.0, or 8.0GT/s link rate x8
– Auto-negotiates to x8, x4, x2, or x1
– Support for MSI/MSI-X mechanisms

Operating Systems/Distributions*
– RHEL/CentOS
– Windows
– FreeBSD
– VMware
– OpenFabrics Enterprise Distribution (OFED)
– OpenFabrics Windows Distribution (WinOF-2)

Connectivity
– Interoperable with 1/10/25/40/50 Gb/s Ethernet switches
– Passive copper cable with ESD protection
– Powered connectors for optical and active cable support

Standards*

–IEEE 802.3cd, 50,100 and 200 Gb/s
–IEEE 802.3by, 25, 50 Gb/s supporting all FEC modes
–IEEE 802.3ba 40 Gigabit Ethernet
–IEEE 802.3ae 10 Gigabit Ethernet

–IEEE 802.3az Energy Efficient Ethernet (supports only "Fast-Wake" mode)
–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 802.1Qaz (ETS)
–IEEE 802.1Qbb (PFC)
–IEEE 802.1Qbg
–IEEE 1588v2

–25G/50G Ethernet Consortium "Low Latency FEC" for 50/100/200GE PAM4 links
–PCI Express Gen 3.0

Card Portfolio & Ordering Information

Table 1 - PCIe HHL Form Factor

Max Network Speed	Interface Type	Supported Ethernet Speeds (GbE)	Host Interface	Additional Features	OPN
1x **10**GbE	SFP28	10, 1	PCIe 3.0 x8		MCX4111A-XCAT
2x **10**GbE	SFP28	10, 1	PCIe 3.0 x8		MCX4121A-XCAT
			PCIe 3.0 x8	Host Management, UEFI Enabled	MCX4121A-XCHT
1x **25**GbE	SFP28	25, 10, 1	PCIe 3.0 x8		MCX4111A-ACAT
			PCIe 3.0 x8	UEFI Enabled	MCX4111A-ACUT
2x **25**GbE	SFP28	25, 10, 1	PCIe 3.0 x8		MCX4121A-ACAT
			PCIe 3.0 x8	UEFI Enabled	MCX4121A-ACUT
1x **40**GbE	QSFP28	40, 25, 10, 1	PCIe 3.0 x8		MCX4131A-BCAT
1x **50**GbE	QSFP28	50, 40, 25, 10, 1	PCIe 3.0 x8		MCX4131A-GCAT

Notes: All tall-bracket adapters are shipped with the tall bracket mounted and a short bracket as an accessory.
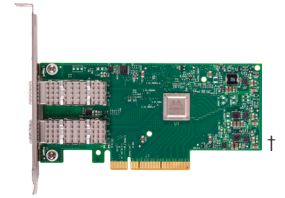Dimensions without bracket: 14.2cm x 6.9cm (low profile)

Table 2 - OCP 2.0 Type 1 Form Factor

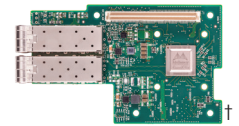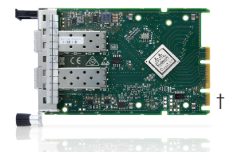| Max Network Speed	Interface Type	Supported Ethernet Speeds (GbE)	Host Interface	Additional Features	OPN
2x **10**GbE	SFP28	10, 1	PCIe 3.0 x8	Host Management	MCX4421A-XCQN
1x **25**GbE	SFP28	25, 10, 1	PCIe 3.0 x8		MCX4411A-ACAN
			PCIe 3.0 x8	Host Management	MCX4411A-ACQN
			PCIe 3.0 x8	UEFI Enabled	MCX4411A-ACUN
2x **25**GbE	SFP28	25, 10, 1	PCIe 3.0 x8		MCX4421A-ACAN
			PCIe 3.0 x8	UEFI Enabled	MCX4421A-ACUN
			PCIe 3.0 x8	Host Management	MCX4421A-ACQN
1x **50**GbE	QSFP28	50, 40, 25, 10, 1	PCIe 3.0 x8	Host Management	MCX4431A-GCAN
			PCIe 3.0 x8	Host Management, UEFI Enabled	MCX4431A-GCUN

Table 3 - OCP 2.0 Type 2 Form Factor

Max Network Speed	Interface Type	Supported Ethernet Speeds (GbE)	Host Interface	Additional Features	OPN
1x **40**GbE	QSFP28	40, 25, 10, 1	PCIe 3.0 x8	Mellanox Multi-Host, Host Management	MCX4431M-BCAN
1x **50**GbE	QSFP28	50, 40, 25, 10, 1	PCIe 3.0 x8	Mellanox Multi-Host, Host Management	MCX4431M-GCAN

Note: OCP2.0 cards are shipped without a bracket.

Table 4 - OCP 3.0 Small Form Factor

Max Network Speed	Interface Type	Supported Ethernet Speeds (GbE)	Host Interface	Additional Features	OPN
2x **25**GbE	SFP28	25, 10, 1	PCIe 3.0 x8	Host Management, Thumbscrew bracket	MCX4621A-ACAB

[†] For illustration only. Actual products may vary.

350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com