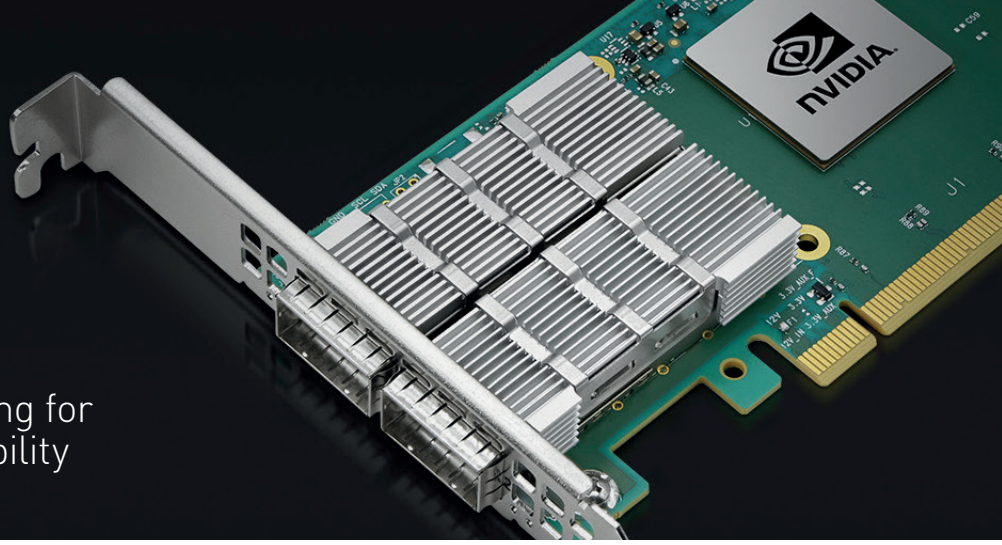
DATASHEET



NVIDIA CONNECTX-6

Featuring In-Network Computing for Enhanced Efficiency and Scalability

Complex workloads demand ultra-fast processing of high-resolution simulations, extreme-size datasets, and highly-parallelized algorithms. As these computing requirements continue to grow, NVIDIA Quantum InfiniBand—the world's only fully offloadable, In-Network Computing acceleration technology—provides the dramatic leap in performance needed to achieve unmatched results in high performance computing (HPC), AI, and hyperscale cloud infrastructures—with less cost and complexity.

NVIDIA® ConnectX®-6 InfiniBand smart adapter cards are a key element in the NVIDIA Quantum InfiniBand platform. ConnectX-6 provides up to two ports of 200Gb/s InfiniBand and Ethernet[1] connectivity with extremely low latency, high message rate, smart offloads, and NVIDIA In-Network Computing acceleration that improve performance and scalability.

High Performance Computing Environments

With its NVIDIA In-Network Computing and In-Network Memory capabilities, ConnectX-6 offloads computation even further to the network, saving CPU cycles and increasing network efficiency. ConnectX-6 utilizes remote direct memory access (RDMA) technology as defined in the InfiniBand Trade Association (IBTA) specification, delivering low latency, and high performance. ConnectX-6 enhances RDMA network capabilities even further by delivering end-to-end packet-level flow control.

Machine Learning and Big Data Environments

Data analytics has become an essential function within many enterprise data centers, clouds, and hyperscale platforms. Machine learning (ML) relies on especially high throughput and low latency to train deep neural networks and improve recognition and classification accuracy. With its 200Gb/s throughput, ConnectX-6 is an excellent solution to provide ML applications with the levels of performance and scalability that they require.

Security Including Block-Level Encryption

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. 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 eliminates 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 offer Federal Information Processing Standards (FIPS) compliance.

KEY FEATURES

- Up to 200Gb/s connectivity per port
- Max bandwidth of 200Gb/s
- Up to 215 million messages/sec
- Extremely low latency
- Block-level XTS-AES mode hardware encryption
- Federal Information Processing Standards (FIPS) compliant
- Supports both 50G SerDes (PAM4)- and 25G SerDes (NRZ)-based ports
- Best-in-class packet pacing with sub-nanosecond accuracy
- PCIe Gen 3.0 and Gen 4.0 support
- In-Network Compute acceleration engines
- RoHS compliant
- Open Data Center Committee (ODCC) compatible

KEY APPLICATIONS

- Industry-leading throughput, low CPU utilization, and high message rate
- High performance and intelligent fabric for compute and storage infrastructures
- Cutting-edge performance in virtualized networks, including network function virtualization (NFV)
- Host chaining technology for economical rack design
- Smart interconnect for x86, Power, Arm, GPU, and FPGA-based compute and storage platforms
- Flexible programmable pipeline for new network flows
- Efficient service chaining enablement
- Increased I/O consolidation, reducing data center costs and complexity

Bring NVMe-oF to Storage Environments

NVMe storage devices are gaining momentum, offering very fast access to storage media. The evolving NVMe over Fabrics (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.

Portfolio of Smart Adapters

ConnectX-6 is available in two form factors: low-profile stand-up PCIe and Open Compute Project (OCP) Spec 3.0 cards with QSFP connectors. Single-port, HDR, stand-up PCIe adapters are available based on either ConnectX-6 or ConnectX-6 DE (ConnectX-6 Dx enhanced for HPC applications).

In addition, specific PCIe stand-up cards are available with a cold plate for insertion into liquid-cooled Intel Server System D50TNP platforms.

NVIDIA Socket Direct

ConnectX-6 also provides options for NVIDIA Socket Direct™ configurations, which improves the performance of multi-socket servers by enabling each of their CPUs to access 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.

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

Socket Direct technology is enabled by a main card housing the ConnectX-6 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 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 200Gb/s to servers with PCIe Gen3-only support.

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.

Broad Software Support

All ConnectX adapters are supported by a full suite of drivers for major Linux distributions, as well as Microsoft® Windows® Server and VMware vSphere®.

HPC software libraries supported include HPC-X, OpenMPI, MVAPICH, MPICH, OpenSHMEM, PGAS, and varied commercial packages.

[1] Please refer to the **ConnectX-6 Ethernet SmartNIC** for Ethernet features and functionality.

Features[2]

InfiniBand

- 200Gb/s and lower rates
- IBTA Specification 1.3 compliant
- RDMA, send/receive semantics
- Hardware-based congestion control
- Atomic operations
- 16 million I/O channels
- 256 to 4Kbyte MTU, 2Gbyte messages
- 8 virtual lanes + VL15

Ethernet

- 200 / 100 / 50 / 40 / 25 / 10 / 1 GbE

Enhanced Features

- Hardware-based reliable transport
- Collective operations offloads
- Vector collective operations offloads
- NVIDIA PeerDirect® RDMA (a.k.a. NVIDIA 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 applications
- 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—NVIDIA ASAP[2]

- Single Root IOV (SR-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 Fabrics 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, and 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

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

Compatibility[2]

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
- Transaction Layer Packet (TLP) Processing Hints (TPH)
- PCIe switch Downstream Port Containment (DPC)
- 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

- Interoperability with InfiniBand switches (up to 200Gb/s, at 4 lanes of 50Gb/s data rate)
- Interoperability with Ethernet switches (up to 200Gb/s, at 4 lanes of 50Gb/s data rate)
- Passive copper cable with ESD protection
- Powered connectors for optical and active cable support

Ethernet Standards

- 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)

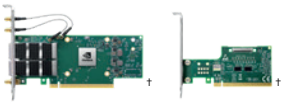
[2] This section describes hardware features and capabilities. Please refer to the driver and firmware release notes for feature availability.

Adapter Card Portfolio and Ordering Information

PCIe Stand-Up Form Factor

InfiniBand Supported Speed (Gb/s)	Ethernet Supported Speed (Gb/s)	Network Ports and Cages	Host Interface (PCIe)	Ordering Part Number (OPN)
100Gb/s and lower	100Gb/s and lower	1x QSFP56	PCIe Gen 3.0/4.0 x16	MCX653105A-ECAT
		1x QSFP56	PCIe Gen 4.0 x8	MCX651105A-EDAT
		2x QSFP56	PCIe Gen 3.0/4.0 x16	MCX653106A-ECAT
200Gb/s and lower	200Gb/s and lower	1x QSFP56	PCIe Gen 3.0/4.0 x16	MCX653105A-HDAT, MCX653105A-HDAL[1]
		2x QSFP56	PCIe Gen 3.0/4.0 x16	MCX653106A-HDAT, MCX653106A-HDAL[1]
200Gb/s and lower	N/A	1x QSFP56	PCIe Gen 3.0/4.0 x16	MCX683105AN-HDAT[2]
100Gb/s and lower	100Gb/s and lower	1x QSFP56	PCIe Gen 3.0/4.0 x16 Socket Direct 2x8 in a row	MCX653105A-EFAT
		2x QSFP56		MCX653106A-EFAT
200Gb/s and lower	200Gb/s and lower	1x QSFP56	Socket Direct 2x PCIe 3.0 x16	MCX654105A-HCAT
		2x QSFP56		MCX654106A-HCAT
100Gb/s and lower	100Gb/s and lower	2x QSFP56		-

Standard PCIe Stand-up SmartNIC



NVIDIA Socket Direct Architecture

[1] Include a cold plate for insertion into liquid-cooled Intel Server System D50TN.
[2] Based on ConnectX-6 DE

Dimensions without brackets are 167.65mm x 68.90mm. All tall-bracket adapters are shipped with the tall bracket mounted and a short bracket as an accessory.

OCP 3.0 Small Form Factor

InfiniBand Supported Speed (Gb/s)	Ethernet Supported Speed (Gb/s)	Network Ports and Cages	Host Interface (PCIe)	Ordering Part Number (OPN)
200Gb/s and lower	200Gb/s and lower	1x QSFP56	PCIe Gen 3.0/4.0 x16	MCX653435A-HDAI
200Gb/s and lower	200Gb/s and lower	2x QSFP56		MCX653436A-HDAI MCX653436A-HDAB
200Gb/s and lower	200Gb/s and lower	1x QSFP56		MCX653435A-HDAE
100Gb/s and lower	100Gb/s and lower	1x QSFP56		MCX653435A-EDAI

OCP 3.0 SFF Form Factor

The last digit of the OCP 3.0 OPN-suffix displays the OPN's default bracket option: I = Internal Lock; E = Ejector Latch.
For other bracket types, contact NVIDIA.

Standalone IC Ordering Information

Max. Speeds Supported	Network Ports	Host Interface (PCIe)	Feature Set	Ordering Part Number (OPN)
100Gb/s	2-port IC	Gen 3.0/4.0 x32	No Crypto	MT28908A0-NCCF-EV
	2-port IC		NVIDIA Multi-Host™, No Crypto	MT28908A0-NCCF-EVM
	2-port IC		NVIDIA Multi-Host, Crypto	MT28908A0-CCCF-EVM
200Gb/s	1-port IC	Gen 3.0/4.0 x32	NVIDIA Multi-Host, No Crypto	MT28904A0-NCCF-HVM
	2-port IC		Crypto	MT28908A0-CCCF-HV
	2-port IC		No Crypto	MT28908A0-NCCF-HV
	2-port IC		NVIDIA Multi-Host, Crypto	MT28908A0-CCCF-HVM
	2-port IC		NVIDIA Multi-Host, No Crypto	MT28908A0-NCCF-HVM
	2-port IC		ConnectX-6 Ex*, NVIDIA Multi-Host, No Crypto	MT28918A0-NCCF-HVM
	2-port IC		ConnectX-6 Ex*, NVIDIA Multi-Host, Crypto	MT28918A0-CCCF-HVM

ConnectX-6 Ex is an enhanced performance ConnectX-6 device that supports higher throughput and lower latency.

† For illustration only. Actual products may vary.

Learn more

Learn more at **www.pny.eu**

PNY Technologies Europe
9 rue Joseph Cugnot, 33708 Mérignac cedex | France
T +33 (0) 5 56 13 75 75 | pnypro@pny.eu
For more information visit: www.pny.eu