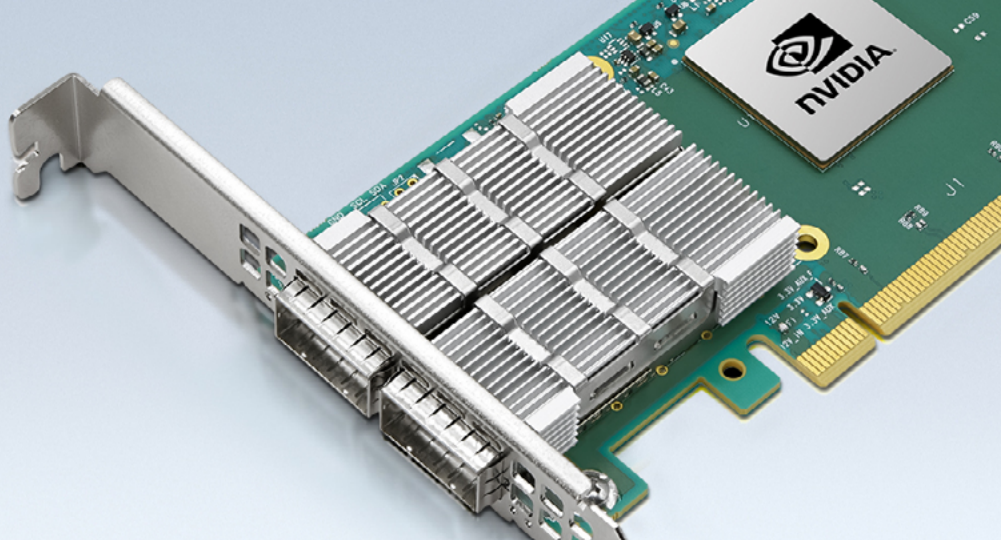# NVIDIA CONNECTX-6
2x200GbE Ethernet SmartNIC
for demanding applications

NVIDIA® ConnectX®-6 Ethernet smart network interface card (SmartNIC) provides up to two ports of 200GbE connectivity, extremely low latency, and a high message rate. These cards offer a high-performance and highly flexible solution for demanding applications in cloud, web 2.0, big data, storage, and machine learning markets.

In addition to all the existing innovative features of earlier generations, the ConnectX-6 offers enhancements to improve performance and scalability as well as PCIe Gen4 support. Moreover, ConnectX-6 Ethernet cards can connect up to 32 lanes of PCIe to achieve 200Gb/s of bandwidth, even on PCIe Gen3 systems.

Cloud and Web 2.0 Environments

Telco, cloud, and web 2.0 customers developing their platforms in software-defined network (SDN) environments are leveraging the virtual switching capabilities of the operating systems on their servers to enable maximum flexibility in their network management and routing protocols.

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 NVIDIA ASAP[2] - Accelerated Switch and Packet Processing™ technology to offload the virtual switch (vSwitch) and virtual router (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 and vRouter offload functions supported by ConnectX-6 include encapsulation and decapsulation of overlay network headers, as well as stateless offloads of inner packets, packet header rewrite (enabling network address translation functionality), hairpin, and more.

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

Key Features

- > Up to 200GbE connectivity per port
- > Maximum bandwidth of 200GbE
- > Up to 215 million messages/sec
- > Sub-800 nanosecond latency
- > Block-level XTS-AES mode hardware encryption
- > Optional FIPS-compliant adapter card
- > Supports 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
- > Open Data Center Committee (ODCC) compatible

Key Features

- > Intelligent, high-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
- > 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 by offering very fast access to storage media. The evolving NVMe over Fabrics (NVMe-oF) protocol leverages remote direct-memory access (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.

Improved 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. 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. With its 200GbE throughput, ConnectX-6 can provide machine learning applications with the levels of performance and scalability that they require. ConnectX-6 utilizes RDMA technology to deliver low latency and high performance. It also enhances RDMA network capabilities even further by delivering end-to-end, packet-level, flow control.

Host Management

NVIDIA 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.

Solutions

> Cloud-native, web 2.0, hyperscale
> Enterprise data centers
> Cybersecurity
> Big data analytics
> Scale-out compute and storage infrastructure
> Telco and NFV
> Cloud storage
> Machine learning and AI
> Media and entertainment
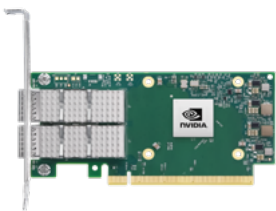
Standard PCIe Stand-up SmartNIC[†]
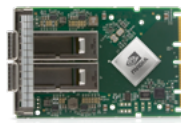
OCP3.0 SFF[†]

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

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
> NVIDIA 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 network address translation (NAT) router

Hardware-Based I/O Virtualization— NVIDIA 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 virtual machines

Storage Offloads

> Block-level encryption: XTS-AES 256/512-bit key
> NVMe-oF 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, NCSI over RBT in OCP cards
> 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

Remote Boot

> Remote boot over Ethernet
> Remote boot over iSCSI
> Unified extensible firmware interface (UEFI)
> Pre-execution environment (PXE)

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
> 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

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

Ordering Information

For NVIDIA ordering information, please contact your NVIDIA sales representative or visit the online ConnectX-6 user manuals:

PCIe HHHL form factor and **OCP 3.0 form factor**

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

[Learn more]

Find out more about ConnectX-6 at **www.NVIDIA.com/en-us/networking/ethernet/connectx-6**