Datasheet

ConnectX-8 InfiniBand SuperNIC

Highest-performance 800G networking designed for massive-scale AI.

The NVIDIA® ConnectX®-8 InfiniBand SuperNIC™ provides 800 gigabits per second (Gb/s) of data throughput with support for NVIDIA In-Network Computing to deliver the performance and robust feature set needed to power trillion-parameter-scale AI factories.

The ConnectX-8 SuperNIC accelerates AI workloads by providing high-speed, low-latency networking, enhancing system performance for diverse AI environments. The ConnectX-8 SuperNIC boasts adaptive routing for highest effective bandwidth, telemetry-based congestion control, and intelligent storage offloads. These features are essential for transferring large datasets or facilitating real-time AI processing in multi-tenant environments.

The ConnectX-8 SuperNIC provides the performance and feature set needed for high-performance computing environments. This includes hardware-based Message Passing Interface (MPI) tag matching and advanced MPI offload engines.

The ConnectX-8 SuperNIC, featuring octal small form-factor pluggable (OSFP) 224 and QSFP112 connectors, is compatible with PCI Express (PCIe) Gen6 Card Electromechanical (CEM) and includes both air-cooled and liquid-cooled options. The SuperNIC can also support an NVIDIA Socket Direct™ 16-lane auxiliary card to achieve 800Gb/s bandwidth utilizing 32 lanes of PCIe Gen5.

Product Specifications	
Maximum total bandwidth	800Gb/s
InfiniBand speeds	XDR 800Gb/s, NDR 400Gb/s, HDR 200Gb/s
Number of network ports	One or two
Host interface	PCIe Gen6, up to 48 lanes
Form factor	PCIe HHHL
Interface technologies	PAM4 (50G, 100G, 200G)

NVIDIA Quantum-X800 InfiniBand Platform

The ConnectX-8 SuperNIC is an integral part of the NVIDIA Quantum-X800 InfiniBand platform, achieving an end-to-end throughput of 800Gb/s for new AI infrastructures and supporting NVIDIA Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)™ to boost In-Network Computing. The SuperNIC connectivity options combined with the NVIDIA LinkX® interconnect portfolio of transceivers and cables provide the maximum flexibility to build a preferred network topology.

Features*

InfiniBand Interface

- Single port or dual ports with up to 800Gb/s of total bandwidth
- InfiniBand Trade Association Spec 1.7 compliant
- Supports XDR 800Gb/s and lower link speeds
- Remote direct-memory access (RDMA), send/receive semantics

Enhanced InfiniBand Networking

- Hardware-based reliable transport
- Extended Reliable Connected (XRC)
- Dynamically Connected Transport (DCT)
- NVIDIA GPUDirect® RDMA
- GPUDirect Storage
- Adaptive routing support
- Enhanced atomic operations
- Advanced memory mapping
- On-demand paging (ODP)
- Enhanced congestion control
- Burst buffer offload
- Single-root input/output (IO) virtualization (SR-IOV)
- Optimized for HPC software libraries, including NVIDIA HPC SDK, NVIDIA Unified Communication X (UCX®), NVIDIA Unified Collective Communication (UCC), NVIDIA Collective Communications Library (NCCL), OpenMPI, MVAPICH, MPICH, OpenSHMEM, partitioned global address space (PGAS)
- Collective operations offloads
- Rendezvous protocol offload
- In-network on-board memory
- 16 million IO channels
- 256 to 4,000 byte maximum transmission unit (MTU), 2GB messages

Management and Control

- Network controller sideboard interface (NC-SI), Management Component Transport Protocol (MCTP) over System Management Bus (SMBus), and MCTP over PCIe
- Platform-Level Data Model (PLDM) for Monitor and Control DSP0248
- PLDM for Firmware Update DSP0267
- PLDM for Redfish Device Enablement DSP0218
- PLDM for Field-Replaceable Unit (FRU) DSP0257
- Security Protocols and Data Models (SPDM) DSP0274
- Serial Peripheral Interface (SPI) to flash
- Joint Test Action Group (JTAG) Institute of Electrical and Electronics Engineers (IEEE) 1149.1 and IEEE 1149.6

Remote Boot

- Remote boot over InfiniBand
- Remote boot over Internet Small Computer Systems Interface (iSCSI)
- Unified Extensible Firmware Interface (UEFI)
- Preboot Execution Environment (PXE)

Cybersecurity

- Platform security:
  - Secure boot with hardware root of trust
  - Secure firmware update
  - Flash encryption
  - Device attestation

Compatibility

PCIe Interface

- PCIe Gen6 and backward compatible (Gen5, 4, 3), 48 lanes
- NVIDIA Multi-Host™ supports connection of up to four hosts
- PCIe switch Downstream Port Containment (DPC)
- Support for Message Signaled Interrupts (MSI)/MSI-X mechanisms

Operating Systems/Distributions

- In-box drivers for major operating systems:
  - Linux: RHEL, Ubuntu
  - Windows
- Virtualization and containers
  - VMware ESXi (SR-IOV)
  - Kubernetes

Ready to Get Started?

To learn more, contact an NVIDIA sales representative:
nvidia.com/en-us/contact/sales

*This section describes hardware features and capabilities.
Please refer to the **driver** and firmware release notes for feature availability.
Images are for illustration only; actual products may vary.

NVIDIA