



# NVIDIA Quantum-X800 InfiniBand Switches

Accelerate AI workloads with 800G InfiniBand.



NVIDIA Quantum-X800 InfiniBand switches deliver 800 gigabits per second (Gb/s) throughput, with ultra-low latency and advanced NVIDIA In-Network Computing, which is essential for handling trillion-parameter-scale generative AI.

These switches incorporate advanced features, including remote direct-memory access (RDMA), the fourth-generation NVIDIA® Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)™, adaptive routing, telemetry-based congestion control, and self-healing technologies. Such enhancements elevate overall application performance within high-performance computing (HPC) and AI data centers.

## Key Benefits

- > **Highest scale for AI:** NVIDIA Quantum-X800 switches enable 2X faster speeds and 5X higher scalability for AI compute fabrics. A two-tier Quantum-X800 fat-tree topology can support over 10,000 800Gb/s host connections.
- > **Next-generation In-Network Computing:** Quantum-X800 switches support SHARP for efficient offloading of compute operations to the network, boosting performance by up to 9X. The fourth-generation of SHARP adds support for FP8 precision and new collective operations such as ReduceScatter and ScatterGather.
- > **Higher effective bandwidth and performance isolation:** Quantum-X800 switches support enhanced adaptive routing and telemetry-based congestion control. This enables nearly perfect effective bandwidth, as well as performance isolation for multi-tenant and multi-job environments.
- > **Enhanced software operations:** Quantum-X800 switches include NVIDIA Networking OS Software (NVOS) for comprehensive chassis management and system configuration. NVOS supports various interfaces, including a command-line interface (CLI), REST APIs, Simple Network Management Protocol (SNMP), and gRPC Network Management Interface (gNMI) telemetry.

## Key Features

- > 800Gb/s speed
- > Ultra-low latency
- > 4th generation NVIDIA SHARP
- > Self-healing technology
- > Enhanced adaptive routing
- > Telemetry-based congestion control
- > Communications frameworks offloads

“NVIDIA Quantum-X800 InfiniBand switches are pivotal for achieving trillion-parameter scale generative AI.”

## NVIDIA Quantum-X800 Switches

The NVIDIA Quantum-X800 Q3400-RA/Q3401-RD 4U switches, the first to leverage 200Gb/s-per-lane serializer/deserializer (SerDes) technology, significantly enhances network performance and bandwidth. It features 144 ports at 800Gb/s distributed across 72 octal small form-factor pluggable (OSFP) cages. The switch's high radix supports a two-level fat-tree topology capable of connecting up to 10,368 network interface cards (NICs) with minimal latency and optimal job locality, as well as other topologies providing connectivity to tens of thousands of GPUs. The Q3400-RA is an air-cooled system designed for standard 19-inch rack cabinets, while the Q3401-RD offers the same air-cooled design with support for rack cabinets utilizing DC power.

For smaller-scale platforms or integration with existing infrastructures, the NVIDIA Quantum-X800 Q3200-RA 2U air-cooled configuration switch is ideal. This system houses two independent switches within a single enclosure, each providing 36 ports at 800Gb/s. The Q3200-RA fixed-configuration switches are well-suited for connecting new compute clusters to previous-generation Quantum and Quantum-2 InfiniBand storage infrastructure.

All NVIDIA Quantum-X800 switches include a dedicated InfiniBand in-band management port specifically for NVIDIA Unified Fabric Manager (UFM®) management, separated on the front panel from the other ports. This separation allows the full set of standard ports to be used for data network connectivity, simplifying port allocation and streamlining topology design.

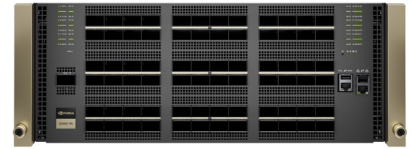
Additionally, NVIDIA Quantum-X800 switches feature optional router capabilities, facilitating the expansion of InfiniBand clusters to support a large scale of nodes located across multiple sites.

## NVIDIA Quantum-X800 InfiniBand Platform

The Quantum-X800 InfiniBand platform includes Q3400-RA, Q3401-RD, and Q3200-RA switches, the NVIDIA ConnectX®-8 SuperNIC and NVIDIA LinkX® interconnect portfolio of transceivers and cables. The platform achieves end-to-end throughput of 800Gb/s from switch to host. For fabric-scale platform management and monitoring, Quantum-X800 features UFM, which enables true software-defined networking with powerful visibility and insights into the performance and health of the network. This end-to-end network platform is purpose-built to deliver the highest performance for the scale-out compute fabrics, enabling massive-scale AI.



NVIDIA Quantum-X800 Q3200-RA  
InfiniBand switch



NVIDIA Quantum-X800 Q3400-RA  
InfiniBand switch

System Specifications			
	Q3200-RA	Q3400-RA	Q3401-RD
<b>Connectivity</b>	Pluggable	Pluggable	Pluggable
<b>Performance</b>	Two switches each of 28.8Tb/s throughput	115.2Tb/s throughput	115.2Tb/s throughput
<b>Switch radix</b>	Two switches of 36 800Gb/s non-blocking ports	144 800Gb/s non-blocking ports	144 800Gb/s non-blocking ports
<b>Connectors and cabling</b>	Two groups of 18 OSFP connectors	72 OSFP connectors	72 OSFP connectors
<b>Management ports</b>	Separate OSFP 400Gb/s InfiniBand in-band management port (UFM)	Separate OSFP 400Gb/s InfiniBand in-band management port (UFM)	Separate OSFP 400Gb/s InfiniBand in-band management port (UFM)
<b>CPU</b>	Intel CFL 4 Cores i3-8100H 3GHz	Intel CFL 4 Cores i3-8100H 3GHz	Intel CFL 4 Cores i3-8100H 3GHz
<b>Security</b>	CPU/CPLD/Switch IC based on IROt	CPU/CPLD/Switch IC based on IROt	CPLD/Switch IC based on IROt
<b>Software</b>	NVOS	NVOS	NVOS
<b>Rack mount</b>	2U	4U	4U
<b>Cooling Mechanism</b>	Air-cooled	Air-cooled	Air-cooled
<b>EMC (emissions)</b>	CE, FCC, VCCI, ICES, and RCM	CE, FCC, VCCI, ICES, and RCM	CE, FCC, VCCI, ICES, and RCM
<b>Product safety compliant/certified</b>	RoHS, CB, cTUVus, CE, and CU	RoHS, CB, cTUVus, CE, and CU	RoHS, CB, cTUVus, CE, and CU
<b>Power Feed</b>	200-240V AC	200-240V AC	48-54V DC
<b>Warranty</b>	One year	One year	One year

## IT Lifecycle Services for Networking

### Experts, insights and ease

Our highly trained experts, with innovative tools and proven processes, help you transform your IT investments into strategic advantages.



#### Plan & Design

Let us analyze your multivendor environment and deliver a comprehensive report and action plan to build upon the existing network and improve performance.



#### Deploy & Integrate

Get new wired or wireless network technology installed and configured with ProDeploy. Reduce costs, save time, and get up and running fast.



#### Educate

Ensure your staff builds the right skills for long-term success. Get certified on Dell Networking technology and learn how to increase performance and optimize infrastructure.



#### Manage & Support

Gain access to technical experts and quickly resolve multivendor networking challenges with ProSupport. Spend less time resolving network issues and more time innovating.



#### Optimize

Maximize performance for dynamic IT environments with Dell Optimize. Benefit from in-depth predictive analysis, remote monitoring and a dedicated systems analyst for your network.



#### Retire

We can help you resell or retire excess hardware while meeting local regulatory guidelines and acting in an environmentally responsible way.

Learn more at [DellTechnologies.com/Services](https://www.dell.com/services)

## Ready to Get Started?

To learn more about Dell Networking solutions, visit [www.dell.com/networking](https://www.dell.com/networking).

© 2025 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA, the NVIDIA logo, ConnectX, LinkX, Scalable Hierarchical Aggregation and Reduction Protocol (SHARP), and UFM are trademarks and/or registered trademarks of NVIDIA Corporation and its affiliates in the U.S. and other countries. Dell and other trademarks are trademarks of Dell Inc. or its subsidiaries. All other trademarks are property of their respective owners. 3555250. AUG25

