



NVIDIA QUANTUM QM8700 SERIES

Accelerating Data Centers with 200G InfiniBand Smart Switches

Scaling Out High-Performance Computing (HPC), AI, Cloud Applications, and Storage Infrastructures with 200G InfiniBand Fixed-Configuration Switches

Faster servers combined with high-performance storage and applications that use increasingly complex computations are causing data bandwidth requirements to spiral upward. As servers are deployed with next-generation processors, HPC, AI, and hyperscale cloud environments will need every last bit of network bandwidth. NVIDIA Quantum QM8700 and QM8790 200 gigabits per second (Gb/s) InfiniBand fixed-configuration switches provide up to forty 200Gb/s ports with 16 terabits per second (Tb/s) of non-blocking bandwidth or eighty 100Gb/s ports with full bidirectional bandwidth per port with extremely low latency—an ideal choice for building flexible and scalable data centers at any size.

Collective Communication Acceleration

The QM8700 and QM8790 fixed-configuration switches are designed to enable NVIDIA In-Network Computing offload engines through NVIDIA Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)[™] technology. Using in-switch silicon-embedded hardware, the switch architecture enables the usage of all active data center devices to accelerate the communications frameworks, resulting in order-of-magnitude performance improvements.

NVIDIA Quantum switches improve the performance of select collective operations commonly used in HPC and deep learning communication frameworks, such as Message Passing Interface (MPI), Shared Memory (SHMEM), and NVIDIA Collective Communication Library (NCCL). By processing data as it traverses the network, NVIDIA Quantum switches eliminate the need to send data multiple times between server endpoints. They also support the aggregation of large data vectors at wire speed, which are crucial for machine learning applications.

Streamlining Network Design and Topologies

The QM8700, together with the NVIDIA[®] ConnectX[®]-6 adapter card, also supports 100Gb/s. With its support for up to 80 ports of 100Gb/s, this unique technology provides double-density radix for 100Gb/s data speeds, thus reducing the cost of network design and topologies. The QM8700 InfiniBand fixed-configuration switch enables four-lane 200Gb/s ports to be split into two distinct two-lane 100Gb/s ports. As the switch creates a highly dense top-of-rack (ToR) switch, it's the perfect solution for double-dense racks. In addition, NVIDIA Quantum-based fixed-configuration switches provide maximum flexibility, supporting a variety of network topologies, including Fat Tree, DragonFly+, multi-dimensional Torus, and more. NVIDIA Quantum fixed-configuration switches are backwards compatible to previous generations and include extensive software ecosystem support.

SYSTEM SPECIFICATIONS

Performance	200Gb/s per port
Switch radix	40 non-blocking ports with aggregate data throughput up to 16Tb/s
Connectors and cabling	Quad small form-factor pluggable (QSFP56) connectors; passive or active copper or active fiber cable; optical module
Power supply	1+1 redundant and hot-swappable power Input range: 100-127VAC, 200-240VAC 80 Gold+ and ENERGY STAR certified power supplies
Management ports	1x RJ45 1x console port: RS232 1x micro USB
CPU	Broadwell ComEx D-1508 2.2GHZ
System memory	Single 8GB
Software	MLNX-OS
System weight	1 PSU: 11.4kg 2 PSUs: 12.48kg
System dimensions	Height: 1.7 in (43.6 mm) Width: 17 in (433.2 mm) Depth: 23.2 in (590.6 mm)
Mount rack	1U rack mount
Operating Conditions	Temperature: <ul style="list-style-type: none"> > Operating: 0°C-40°C > Non-operating: -40°C-70°C Humidity: <ul style="list-style-type: none"> > Operating: 10%-85% non-condensing > Non-operating: 10%-90% non-condensing Altitude: up to 3,200m
EMC (Emissions)	CE, FCC, VCCI, ICES, and RCM
Safety	RoHS Compliant
Warranty	1 year

Enhanced Management

The internally managed QM8700 switch features an on-board subnet manager that enables simple, out-of-the-box bring-up for up to 2,000 nodes. NVIDIA MLNX-OS® delivers full chassis management through command-line interface (CLI), web-based user (WebUI), Simple Network Management Protocol (SNMP), or JavaScript Object Notation (JSON) interfaces.

The externally managed QM8790 switch can utilize the advanced NVIDIA Unified Fabric Manager (UFM®) platform to empower data center operators to efficiently provision, monitor, manage, preventatively troubleshoot, and maintain the modern data center network to realize higher utilization and reduce overall opex.

Building Efficient Clusters

NVIDIA Quantum-based QM8700 and QM8790 switches are used as cost-effective building blocks for deploying high-performance and computing-efficient data centers. With features such as static routing, adaptive routing, congestion control, and quality of service that enable modern topologies, the switches ensure maximum effective fabric bandwidth by eliminating congestion hot spots. Featuring best-in-class design to support low power consumption, power is further reduced upon partial port utilization. Whether looking at price to performance or energy to performance, the QM8700 series offers superior performance, low power, and scalability, reducing capital and operating expenses and providing the best return on investment.

Ordering Information

ORDERABLE PART NUMBER (OPN)	DESCRIPTION
MQM8700-HS2F	NVIDIA Quantum 200Gb/s InfiniBand switch, 40 QSFP56 ports, two power supplies (AC), x86 dual core, standard depth, P2C airflow, rail kit
MQM8700-HS2R	NVIDIA Quantum 200Gb/s InfiniBand switch, 40 QSFP56 ports, two power supplies (AC), x86 dual core, standard depth, C2P airflow, rail kit
MQM8790-HS2F	NVIDIA Quantum 200Gb/s InfiniBand switch, 40 QSFP56 ports, two power supplies (AC), unmanaged, x86 dual core, standard depth, P2C airflow, rail kit
MQM8790-HS2R	NVIDIA Quantum 200Gb/s InfiniBand switch, 40 QSFP56 ports, two power supplies (AC), unmanaged, x86 dual core, standard depth, C2P airflow, rail kit

[Learn more](#)

Learn more about **NVIDIA Quantum InfiniBand switches**.

Learn more about **NVIDIA UFM software**.