



In-Network Computing and Next Generation HDR 200G InfiniBand

With the exponential growth of data that needs to be analyzed and the data resulting from ever-more complex workflows, the need for faster data movement has never been more challenging and critical to the worlds of High Performance Computing (HPC) and machine learning. Mellanox Technologies, the leading global supplier of end-to-end InfiniBand and Ethernet interconnect solutions and services for servers, storage, and hyper-converged infrastructure, has once again moved the bar forward with the introduction of an end-to-end HDR 200G InfiniBand product portfolio.

WHEN FAST ISN'T FAST ENOUGH

In an age of digital transformation and big data analytics, the need for HPC and deep learning platforms to move and analyze data both in real-time and at faster speeds is ever increasing. Machine learning enables enterprises to leverage the vast amounts of data being generated today to make faster and more accurate decisions. Whether for brain mapping or for homeland security, the most demanding supercomputers and data center applications need to produce astounding achievements, often in real time!

Until relatively recently, state-of-the-art applications analyzing automotive construction or weather simulations enjoyed the data performance and throughput speeds offered by 100G interconnect. Today's HPC, machine learning, storage and hyperscale now require both faster interconnect solutions and more intelligent networks to analyze data and run complex simulations with greater speed and efficiency.

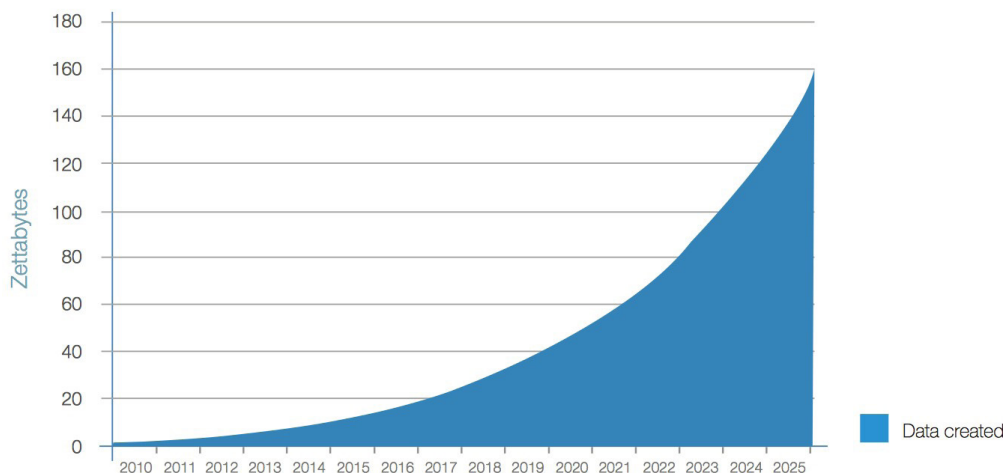


Figure 1: Exponential Data Growth - Everywhere

From the CPU to the Data

The need to support growing data speeds, throughput and simulation complexity accompanies a widespread recognition that the CPU has reached the limits of its scalability. Many have joined the ranks of those believing that it is no longer feasible to move data all of the way to the compute elements; rather computational operations should be performed on the data wherever the data is. Thus the start of a “Data-Centric” trend toward offloading network functions from the CPU to the network. By lightening the load on the server’s processors, the CPUs can devote all their cycles to the application. This approach increases system efficiency by allowing users to run algorithms on the data in-transit rather than waiting for the data to reach the CPU.

Next-Generation Machine Learning Interconnect Solutions

Mellanox’s HDR 200G InfiniBand solutions represents the industry’s most advanced interconnect solution for HPC and deep learning performance and scalability. Mellanox is the first company to enable 200G data speeds, doubling the previous data rate and expanding In-Network Computing capabilities to accommodate the larger message sizes typically found in deep learning application workloads. Utilizing Mellanox technology, the world’s most data-intensive applications and popular frameworks are leveraging Mellanox to accelerate the performance of their applications and frameworks: Yahoo has demonstrated 18X speedup for image recognition; Tencent has been able to achieve world record performance for data sorting; and NVIDIA has incorporated Mellanox solutions inside their deep learning DGX-1 appliance in order to provide 400Gb/s data throughput, and to build one of the most power-efficient machine learning supercomputers.

In-Network Computing and Security Offloads

Machine learning applications are based on training deep neural networks, which require complex computations and fast and efficient data delivery. Besides doubling the speed and providing the higher radix switch, Mellanox’s new HDR 200G switch and adapter hardware supports in-networking computing (application offload capability) and in-network memory. Mellanox’s HDR InfiniBand solution offers offloads beyond that of RDMA and GPUDirect® to computation for higher level communication framework collectives. This dramatically improves neural network training performance and overall machine learning applications, while saving on CPU cycles and increasing the efficiency of the network.

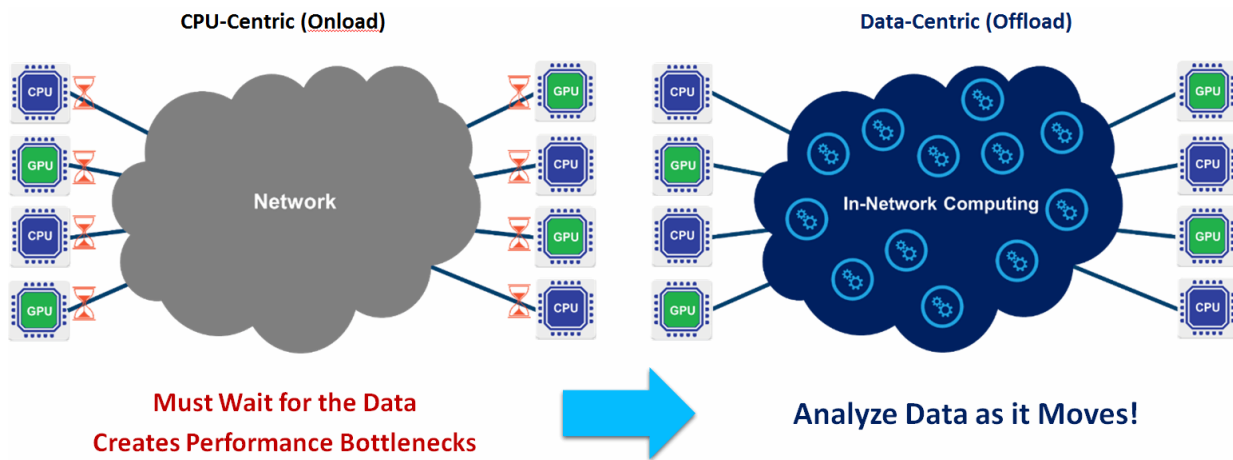


Figure 2. HDR100 Requires 1.6X Fewer Switches for 400 Nodes

OPTIMIZED SWITCHING – FROM 100G TO 200G

To show the improved performance in bandwidth, from 100G to 200G, we can compare Mellanox’s edge and chassis InfiniBand switch offerings, clearly indicating that 200G delivers more than twice the performance in the same QSFP package and 1RU edge switch:

EDR 100G InfiniBand Edge Switch			HDR 200G InfiniBand Edge Switch	
Height	Model	Switching Capacity	Model	Switching Capacity
1RU	SB7800	36 ports of 100G at 7.2Tb/s I/O	QM8700	<ul style="list-style-type: none"> • 80 Ports of HDR100 (100G) or • 40 Ports of 200G at 16.0TB/s I/O

EDR 100G InfiniBand Chassis Switch			HDR 200G InfiniBand Chassis Switch	
Height	Model	Switching Capacity	Model	Switching Capacity
28U	CS7500	<ul style="list-style-type: none"> • Up to 648 EDR (100G) ports • Up to 130TB/s switching capacity 	CS8500	<ul style="list-style-type: none"> • Up to 800 HDR (200G) ports • Up to 1600 HDR100 (100G) ports • Up to 320TB/s switching capacity

HIGH PERFORMANCE ADAPTERS

To complete its end-to-end HDR solution, Mellanox ConnectX®-6 delivers HDR 200G throughput with 200 million messages per second at under 600 nanoseconds of latency for both InfiniBand and Ethernet. Backward compatible, ConnectX-6 also supports HDR100, EDR, FDR, QDR, DDR and SDR InfiniBand as well as 200, 100, 50, 40, 25, and 10G Ethernet speeds.

ConnectX-6 also offers improvements in Mellanox’s Multi-Host® technology, allowing for up to eight hosts to be connected to a single adapter by segmenting the PCIe interface into multiple and independent interfaces. This leads to a variety of new rack design alternatives, lowering the total cost of ownership in the data center by reducing CAPEX (cables, NICs, and switch port expenses), and OPEX (cutting down on switch port management and overall power usage). ConnectX-6 200G InfiniBand and Ethernet (VPI) Network Adapter Storage customers will benefit from ConnectX-6’s embedded 16-lane PCIe switch, which allows them to create standalone appliances in which the adapter is directly connected to the SSDs. By leveraging ConnectX-6 PCIe Gen3/Gen4 capability, customers can build large, efficient, high-speed storage appliances with NVMe devices.

Rounding out Mellanox’s HDR 200G InfiniBand portfolio is its line of LinkX™ cables. Mellanox offers direct-attach 200G copper cables and 2 x 100G splitter breakout cables to enable HDR100 links, as well as 200G active optical cables that reach up to 100 meters. All LinkX cables in the 200G line come in standard QSFP packages.

To see how Mellanox fares against the competition, we can compare the benefits of Mellanox’s HDR100 versus Omni-Path:

ROI / TCO Advantages of Mellanox HDR100 (100G) Switch versus Omni-Path (100G) Switch		
	Mellanox HDR100 100G Switch)	Omni-Path Switch
Port radix	80p	48p
Chassis	1600p	768p (1152p in the future)

The higher port radix enables Mellanox to build a similar cluster size as Intel, using almost half the amount of switches and cables. Lowering TCO and increasing ROI, the Mellanox HDR solution reduces power consumption, thus lowering OPEX as well as reducing rack space and thus, CAPEX. Moreover, the higher density solution delivers performance improvements with a 50% reduction in latency, as well as a simpler interconnect solution that requires fewer cable connections and has HDR fat uplinks which are more robust for congestion and microbursts.

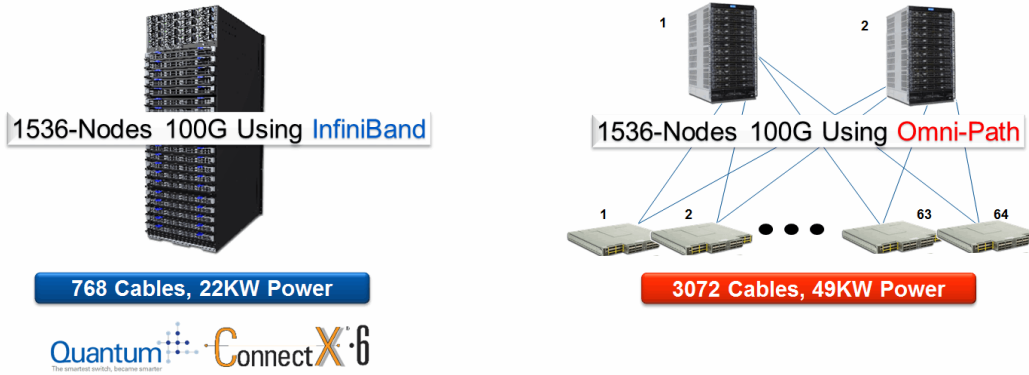


Figure 3: HDR100 Savings vs. Competition: 3X Real Estate Saving, 4X Cable Saving, 2X Power & Latency Saving

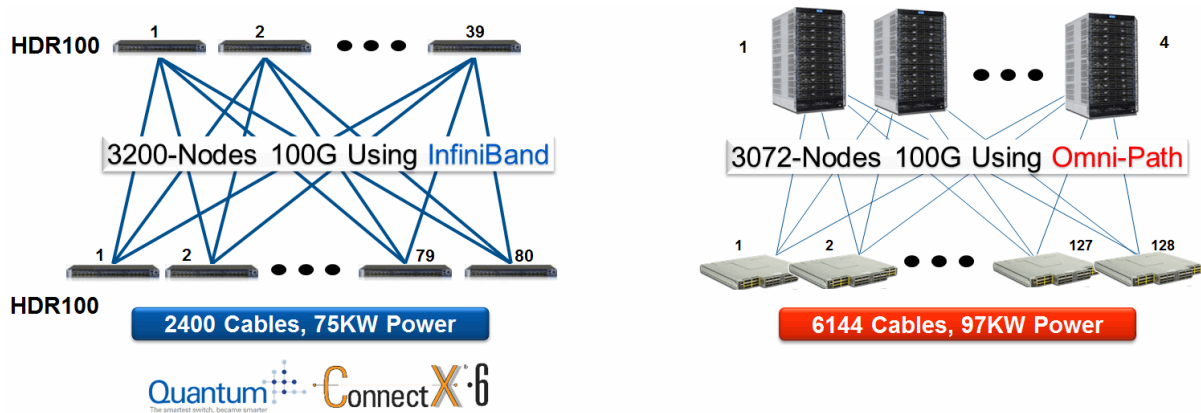


Figure 4. HDR100 Savings of Mellanox vs. the Competition: 1.6X Switch Saving, 2X Cable Saving, Wider Uplink Pipes

EMPOWERING NEXT-GENERATION DATA CENTERS

As the requirement for intensive data analytics increases, there is a corresponding demand for higher bandwidth. Even 100Gb/s is insufficient for some of today's most demanding data centers and clusters. Moreover, the traditional CPU-centric approach to networking has proven to be too inefficient for such complex applications. The Mellanox HDR 200G solution address these issues by providing the world's first 200Gb switches, adapters, and cables and software, and by enabling In-Network Computing to handle data throughout the network instead of exclusively in the CPU. With its 200G solution, Mellanox continues to push the industry toward Exascale computing and remains a generation ahead of the competition.

The advantages of leveraging an all-Mellanox ecosystem range from inter-product compatibility of switches, adapters and cables to simpler integration and implementation with the customer's pre-existing systems, and streamlined training sessions for IT staff.

About Mellanox

Mellanox is a leading supplier of end-to-end intelligent interconnect solutions and services for a wide range of markets including high performance computing, machine learning, enterprise data centers, Web 2.0, cloud, storage, network security, telecom and financial services. Mellanox adapters, switches, cables and software implement the world's fastest and most robust InfiniBand and Ethernet networking solutions for a complete, high-performance machine learning infrastructure. These capabilities ensure optimum application performance.

For more information, please visit: <http://www.mellanox.com/solutions/hpc/>