



HPE and Mellanox: Advanced Technology Solutions for HPC



HPE server and storage systems address a host of next generation HPC challenges, offering an optimal approach to meet the technology needs of organizations that operate massive HPC installations with tens of thousands of servers. The HPE product portfolio incorporates a broad range of solutions that address unique challenges in every area of discipline, including energy research, financial services, government and education, aerospace and manufacturing, and of course life sciences. HPE products and solutions, at their core, feature Mellanox ConnectX®-4 100Gb/s InfiniBand and 100Gb Ethernet connectivity, which enables extreme computing, real-time response, I/O consolidation, and power savings for a broad range of HPC application environments.

For nearly a decade, HPE and Mellanox have been improving high-performance computing, enabling game-changing capabilities in the world's most critical technical arenas. As the performance leader in the interconnect industry, Mellanox paves the way for faster, more efficient, and more intelligent data transfers. Not only does Mellanox provide industry standard-based native RDMA technology that ensures the lowest latencies in the market, but it also further improves performance with advanced hardware-based acceleration engines.

Perhaps even more significantly, Mellanox designs to a CPU-offload architecture, saving valuable CPU cycles for compute tasks, while ensuring that networking tasks are handled by the network, instead of leveraging the CPU to perform network operations. Mellanox solutions are highlighted in nearly every HPE Benchmarking and Customer Solution Center around the world, just as prominently as they are found in the elite of the Top500 supercomputers list.

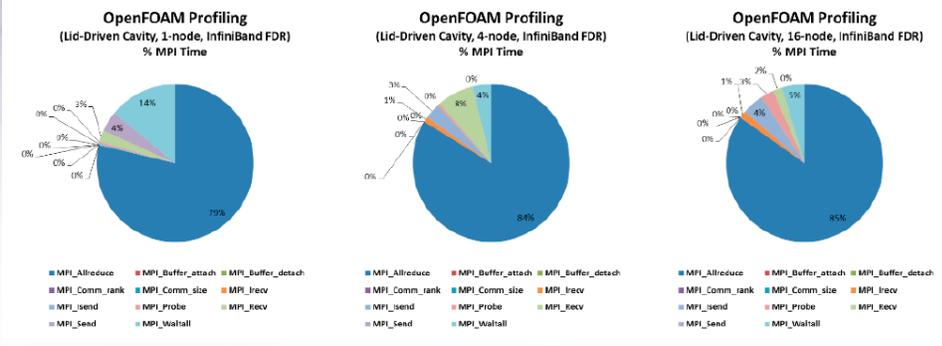
Just as HPE leads the industry with superior class system solutions, Mellanox has always been the obvious choice for state-of-the-art networking capabilities. What might not be so clear-cut is which interconnect product is right for a given

cluster's unique tasks. More specifically, why should a data center make the move to EDR 100Gb/s versus FDR 56Gb/s interconnect? On the surface, it seems like a rhetorical question, considering that typically higher bandwidth equates with moving more data and is ultimately a faster interconnect. Because EDR improves upon the overall bandwidth and latency, additional capabilities and improved acceleration engines can also enable better scaling; it should be a foregone conclusion that a data center would benefit from EDR over FDR.

However, not every organization can justify the commitment to such an upgrade without being able to prove significant benefits above and beyond the "speeds & feeds." EDR, in a respect, is still like the shiny new sports car that can go from 0-60 in mere seconds, but that speed alone is not necessarily enough to make the car a top-performer.

Just as HPE leads the industry with superior class system solutions, Mellanox has always been the obvious choice for state-of-the-art network capabilities. What might not be so clear-cut is which interconnect product is right for a given cluster's unique tasks...

Only HPE solutions that include Mellanox end-to-end 100Gb/s EDR deliver efficiency, scalability, and overall system performance that results in maximum performance per TCO dollar.



Upon deeper analysis, however, EDR 100Gb/s technology has a lot more going for it than just the higher bandwidth, with advantages that make it a clear choice over FDR 56Gb/s performance. The benefits of EDR 100Gb/s are notably brought forth to real application performance through additional improvements to the acceleration engines, and in the software that exposes the underlying hardware architecture. Because nearly all HPC applications are written around MPIs, which heavily leverage collective operations such as AllReduce and All-to-All, Mellanox acceleration capabilities for these operations drastically improve performance and the ability to continue scaling as you add more compute resources.

Mellanox offers an enhanced software package called HPC-X™ Scalable Toolkit, to couple the advanced capabilities of software along with the underlying hardware. HPC-X is an MPI that is heavily optimized for collective communication operations, and when used with the included MXM and MCA accelerator libraries, it further reduces runtime and increases scalability.

In recent benchmark tests run on HP ProLiant XL170r and XL230a Gen9 32-node (1024-core) clusters, significantly improved application performance and scalability was achieved with Mellanox EDR than would have been possible on Mellanox FDR. In the below graphs, the XL170r cluster, represented by the dark blue bars, used Mellanox ConnectX-4 100Gb/s EDR InfiniBand adapters and a Mellanox Switch-IB™ SB7700 36-port 100Gb/s EDR InfiniBand switch, while the XL230a cluster used Mellanox ConnectX-3 56Gb/s FDR InfiniBand adapters and a Mellanox SwitchX®-2 SX6036 36-port 56Gb/s FDR InfiniBand switch.

More than just the enhancements to the bandwidth and acceleration engines of the host channel adapters, the Mellanox EDR 100Gb/s switches also have much improved port-to-port latency of less than 90ns, guaranteeing

very predictable top-notch performance for any HPC cluster. To further the performance per TCO dollar, Mellanox and HPE are already on the path to the next generation of network capabilities that is, yet again, an industry game-changer.

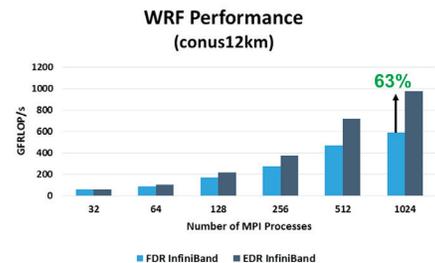
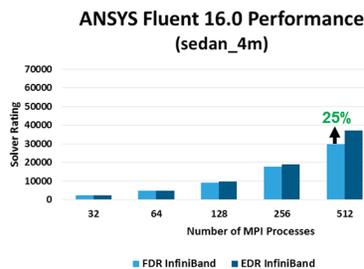
To return to the sports car analogy, Mellanox EDR is not only capable of 0-60 in mere seconds, handles beautifully, and gets great gas

mileage... but can you imagine if the car also has onboard intelligence that could optimize traffic patterns to ensure clear roads and green lights at all times?

Mellanox's Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)™ already adds the benefit of such intelligence already, improving upon the performance of collective operations by processing the data as it traverses the network and eliminating the need to send data multiple times between endpoints. This innovative approach will decrease the amount of data traversing the network as aggregation nodes are reached. Implementing collective communication algorithms in the network also has additional benefits, such as freeing up valuable CPU resources for computation rather than using them to process communication.

Switch-IB™ 2 is the first InfiniBand switch to employ Mellanox SHARP technology, which enables the switch to operate as a co-processor and introduces more intelligence and offloading capabilities to the network fabric.

So while every company must weigh the cost and commitment of upgrading its data center or HPC cluster to EDR, the benefits of such an upgrade go well beyond the increase in bandwidth. Only HPE solutions that include Mellanox end-to-end 100Gb/s EDR deliver efficiency, scalability, and overall system performance that results in maximum performance per TCO dollar.



350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
 Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com