

Applications Acceleration with Mellanox RDMA enabled Networking Solutions

Realize Unprecedented Application Performance by Offloading Compute-Intense Data Communication

Business Benefits

Increase Server Efficiencies



Increase CPU efficiency by 33%



Achieve over 1 Million IOPS

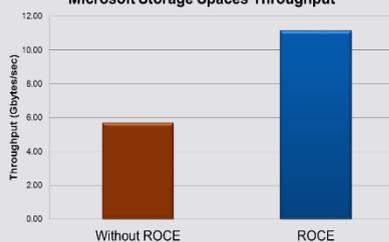


Reduce latencies by 50%



Sustain near line rate bandwidth

Microsoft Storage Spaces Throughput



- 2X better performance with RoCE
 - 2X higher bandwidth & 2X better CPU efficiency
- RoCE* achieves full Flash storage bandwidth
 - Remote storage without compromises

Executive Summary

The expansion of business-critical information, unstructured data, and rich content within enterprises continues to change storage dynamics. This trend is driving the need for higher connectivity speeds to keep up with applications requirement for access to data. In order to address these market conditions, Microsoft and Mellanox have joined forces to solve the storage access problem by removing CPU overhead and providing faster and more efficient access to data for enterprise applications.

Microsoft SMB

The solution comes upon enhances to Microsoft's Server Message Block (SMB) 3.0 protocol in Windows Server 2012 R2. The SMB protocol was extended to include two new features, SMB Direct, and SMB Multichannel. SMB Direct implemented the use of various high-speed RDMA methods to transfer large amounts of data with little CPU intervention. SMB Multichannel allows file servers to use multiple network connections simultaneously and provides fault tolerance through automatic discovery of network paths to dynamically add connections as required. The addition of these two features allows Mellanox RDMA enabled IO adapters to deliver line-rate performance and improve availability by optimizing data transfer between server and storage over standard Ethernet. By offering built-in support for Mellanox RDMA, Microsoft has enabled customers to deploy storage on low-cost file servers, while delivering higher performance. This enables storage

that rivals costly Fibre Channel SANs in efficiency, with lower latency and operating cost.

Increasing Server Efficiency

Mellanox ConnectX®-4 Lx EN Ethernet adapters support RDMA at 10/25/40/50 GbE speeds while the ConnectX®-4 network adapters offer 10/25/40/50 and 100Gb/s Ethernet connectivity and provide the highest performance and most flexible solution for high-performance RDMA implementations. RDMA provides low latency by removing processing of the protocol stack for data transmission from the operating system, allowing applications to directly read and write to remote virtual memory and to directly exchange messages. This drastically reduces the CPU's involvement in data I/O requests and reduces memory bandwidth bottlenecks. Mellanox RDMA technologies process all transport protocol in the adapter hardware, completely bypassing the host OS. This permits the adapters to transfer data between servers and storage with minimum involvement from the host CPU. As a result, SMB Direct is extremely fast with client-to-file server performance, almost equaling solution that use locally attached storage. The CPU reduction in this process leaves more cycles available for server applications. The result is larger numbers of hosted VMs per physical servers, more VDI instances running simultaneously and SQL environments achieve high-performance file sharing to complete queries quicker. Data Centers efficiencies and scalability increases in each of these instances.

Hyper-V Use Case

The purpose of the use case is to show the efficiency of the CPU when running workloads within Hyper-V. The demonstration utilized two servers, a client and file server, that were connected back-to-back with Mellanox ConnectX®-4 100Gb/s Ethernet adapters running SMB Direct over RoCE (RDMA over Converged Ethernet). Three Micron NVMe SSD 300GB 6G SAS adapters were used to form an all-SSD storage pool and IOMeter for the primary I/O generator and benchmarking tool. To simulate a real-world environment, different workload were used to provide data streams with both static and mixed I/O. Four cores were allocated to running the data communication task and the proof of concept run with and without RoCE enabled. The results demonstrated that workloads running on Hyper-V VMs over RoCE can achieve over 1 million IOPS of throughput and sustain a near line rate of 100Gbps of aggregate bandwidth without taxing the CPU. When compared against the result with RoCE disabled, latency decreased nearly 50% and CPU efficiency by more than 33%. The CPU was bound due to the requirement to process the data transport protocol; sequence checking, packet loss recovery, CRC

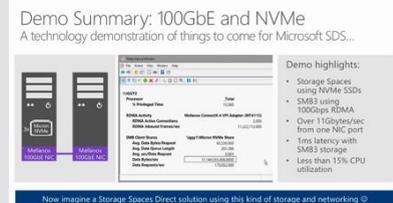
checking, and passing kernel to user space transmission. Running the data transport stack in the CPU steals cycles from applications and limits overall productivity. The Mellanox adapter frees the CPU from this process, delivering improved efficiencies.

RoCE Advantages

By using Microsoft’s SMB 3.0 over Mellanox RDMA technologies, organizations can deliver extremely high levels of storage performance and resiliency. RoCE requires very low CPU utilization and takes advantage of Priority Flow Control in Data Center Bridging Ethernet for lossless connectivity. Configuring Microsoft SMB with Mellanox RDMA enabled workloads running in Hyper-V to achieve over 1 million IOPS. This is not a novel idea, the gaming industry has been offloading graphic processing for a long time. And GPU-accelerated computing has been used to accelerate scientific, analytics, and engineering applications since the mid-2000s. Adding these powerful capabilities into your arsenal enables a comprehensive platform that is poised to tackle the storage demands of the modern data center.

Technology Demonstration

More information about the technology demonstration can be seen in the Microsoft 2015 Ignite conference presentation given by Claus Joergensen, Principal Program Manager, Microsoft



“Enabling Private Cloud Storage Using Servers with Local Disks”

For more information:

Download the [Microsoft White Paper](#)

Download a Mellanox Case Study:

http://www.mellanox.com/related-docs/case_studies/CS_MHA_Cloud.pdf

http://www.mellanox.com/related-docs/case_studies/CS_Financial_Institute.pdf

To learn more about Microsoft SMB visit: [https://technet.microsoft.com/enus/library/hh831795\(v=ws.11\).aspx](https://technet.microsoft.com/enus/library/hh831795(v=ws.11).aspx)

Network Offload Frees Up CPU for Application Processing

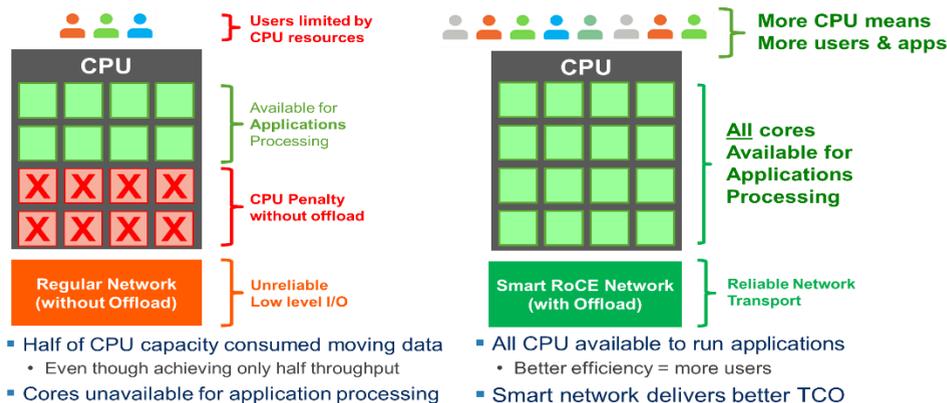


Figure 1. – Microsoft and Mellanox 100Gb/s cloud efficiency demonstration



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