



# High-Performance Server SAN for Web-Scale Data Centers

## BACKGROUND

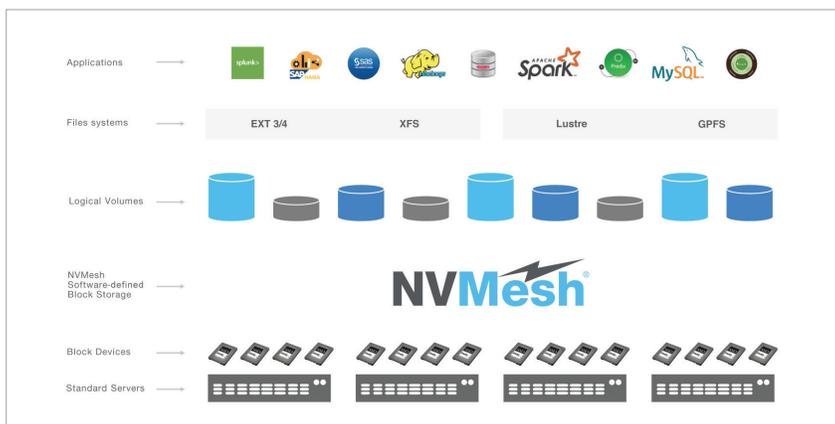
New application workloads are driving the need for new storage architectures. Cloud and mobile applications, Industrial IoT and machine learning have a massive impact on the volumes of data that need to be transferred, stored and processed. Real-time analytics demand the lowest levels of latency. To meet these new requirements, enterprises and service providers are seeking to optimize their infrastructures in the same way as the hyperscale data centers have done. For storage, this means they want to deploy scale-out storage infrastructures using standard servers.

Mellanox and Excelero have joined forces to deliver a high-performance, low-latency Server SAN solution leveraging industry standard servers with NVMe flash, Mellanox end-to-end RMDA-accelerated networking and Excelero NVMesh® Virtual SAN software – a Software-Defined Block Storage solution that allows unmodified applications to utilize pooled NVMe storage devices across a network at local speeds and latencies.

## EXCELERO NVMesh

Excelero's Software-Defined Storage platform "NVMesh" enables customers to design Server SAN infrastructures for the most demanding enterprise and cloud-scale applications. NVMesh is a Software-Defined Block Storage solution that forms a Non-Volatile Mesh, a distributed block layer that allows unmodified applications to utilize pooled NVMe storage devices across a network at local speeds and latencies. Distributed NVMe storage resources are pooled with the ability to create arbitrary, dynamic block volumes that can be utilized by any host running the NVMesh block client. These virtual volumes can be striped, mirrored or both while enjoying centralized management, monitoring and administration. In short, applications can enjoy the latency, throughput and IOPs of a local NVMe device while at the same time getting the benefits of centralized, redundant storage.

A key component of Excelero's NVMesh is the patented Remote Direct Drive Access (RDDA)



## KEY BENEFITS

- Use standard servers with state-of-the-art flash components and high-performance network adapters
- Pool NVMe over the network at local speeds to maximize utilization and avoid data localization
- Scale performance & capacity linearly, across data centers without limits
- No noisy neighbors through full logical disaggregation of storage and compute

*"Excelero and Mellanox have joined forces to enable customers to leverage the full potential of NVMe flash at scale. With Mellanox ConnectX and NVMesh, customers can achieve full NVMe capacity utilization without giving up local performance characteristics effectively, which lowers the storage cost dramatically."*

*- Lior Gal  
CEO, Excelero*

functionality, which leverages RDMA to bypass the target CPU and, thus, avoids the noisy neighbor effect for the application. The shift of data services from centralized CPU to client side distribution enables unlimited linear scalability, provides deterministic performance for applications and enables customers to maximize the utilization of their flash drives. Elastic Virtual SAN is deployed as a virtual, distributed non-volatile array and supports both converged and disaggregated architectures, giving customers full freedom in their architectural design.

### NVMeSH Features

Data Management & Protection	
Multiple Transports	Patented RDDA NVMe-ready
Logical Volumes	Concatenated RAID 0 RAID 1 RAID 10
Multiple Drive Types	NVMe SATA SAS
Failure Domains	Host, rack & row aware
Management & Monitoring	
Interactive Interfaces	Web GUI CLI commands
Automated Provisioning	RESTful API Docker Persistent Volumes
Next-Gen Data Center	
Flexible Topologies	Physically converged Disaggregated Mixed
High Performance	Client, target and management scale independently
Scale-Out Architecture	Intelligent clients utilize multiple hosts, drives, network paths and racks
Advanced Networking	Ethernet (RoCE v2) InfiniBand

## MELLANOX END-TO-END ETHERNET SOLUTION

The Mellanox end-to-end Ethernet solution greatly simplifies the deployment and management of high-speed networking in the data center. Optimized for ultra-low latency lossless fabrics, the Mellanox Ethernet switches support the requirements of today's performance-demanding data centers and work as the ideal top-of-rack switches for various hyper-converged infrastructure deployments.

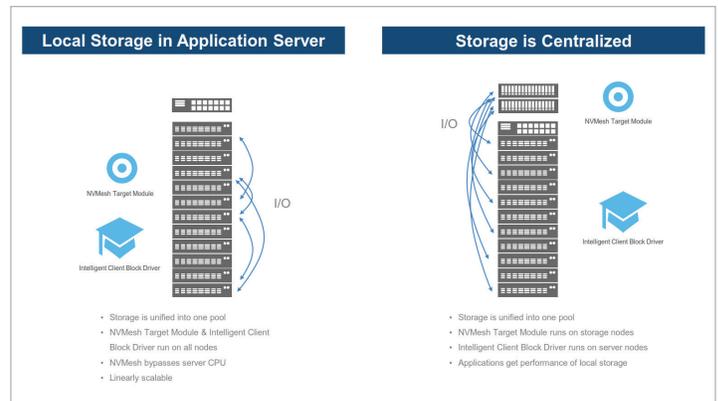
Equipped with intelligent ASICs, Mellanox Ethernet adapter cards provide offloading mechanisms such as Erasure Coding, T10/DIF, TCP and UDP offloads, and overlay offloads. By bypassing the CPU, RDMA over Converged Ethernet (RoCE) frees up the CPU's resources, especially in a CPU-bound environment, for the necessary storage and compute tasks, allowing for higher scalability and efficiency with the data center.

Leveraging its industry-leading networking technologies in high-performance

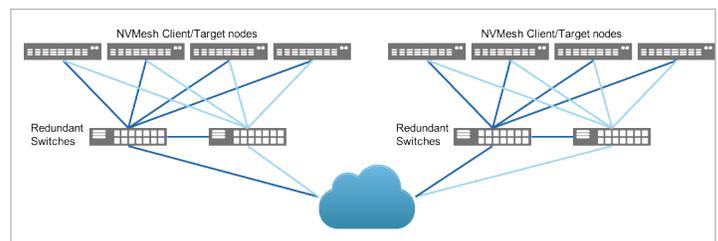
computing, Mellanox builds cutting-edge 25/50/100G Ethernet switches with line-rate performance. Mellanox Spectrum™ switches, powered by Mellanox's own ASICs, do non-blocking, cut-through switching at line speeds of 1/10/25/40/50/100Gb/s with no packet loss. The dynamically shared switch buffer provides the best microburst adoption thus enables the storage solution to deliver guaranteed throughput and latency. Combined with faster congestion notification, Mellanox switches form a network fabric that unleash the maximal power of NVMe storage pool across the network.

## JOINT SOLUTION

The joint Mellanox-Exceero solution provides remote, high-speed, low-latency shared storage with in-server flash performance. Customers can deploy the solution on existing application clusters and turn them into converged infrastructure with compute and a high-performance storage without affecting the application performance. With this "just add software" approach, customers remove storage bottlenecks without the need to purchase external storage arrays or servers. Alternatively, customers can choose for a disaggregated architecture and add NVMe storage nodes that can be accessed by application servers with the same performance as in-server NVMe.



The solution scales naturally, targeting larger data centers with upwards of 100,000 converged nodes. NVMeSH has a pure software-defined architecture, avoiding any centralized capability for the data path. This eliminates centralized metadata and lock management. Client, target and management functionalities are built as scale-out technologies. As it is critical to ensure that the networking pattern is suitable for scale, clients are completely independent. Rare knowledge sharing occurs indirectly and anonymously via targets. Client independence enables easy client scaling.



The Mellanox-Excelero solution was designed to let applications enjoy full performance, capacity and processing power of underlying servers and storage. Leveraging the NVMe RDMA technology, the solution adds a mere 5µs latency over local NVMe drive latency, allowing customers to leverage full performance of their hardware at data center scale. The solution enables customers to achieve higher performance levels than proprietary all-flash-arrays at a much lower \$/IOP, or \$/GB/s.

## SOLUTION BENEFITS

### Virtual SAN

NVMe is the only virtual SAN approach for shared NVMe: it is the only solution that is 100% software-only. Customers can choose any standard servers and state-of-the-art flash drives to meet all their storage requirements.

### Unified NVMe

NVMe allows unmodified applications to utilize pooled NVMe storage across a network at local speeds and latencies. A unified pool of NVMe enables customers to maximize NVMe utilization and avoids data locality issues for the application.

### Zero CPU Utilization

NVMe enables 100% converged infrastructure by full logical disaggregation of storage and compute. The patented Remote Direct Drive Access (RDMA) bypasses the CPU and avoids the noisy neighbors effect for the application. Applications can leverage the full capability of CPU.

### Flexible

NVMe is the industry's most flexible deployment model for Server SAN: it is deployed as a virtual, distributed non-volatile array and supports both converged and disaggregated architectures, and even mixed environments, giving customers full freedom in their architectural design.

### 100% Efficiency

NVMe is the only NVMe sharing technology that scales performance linearly at near 100% efficiency. The shift of data services from centralized CPU to complete client side distribution provides deterministic performance for applications and enables customers to maximize the utilization of their flash drives.

## ABOUT EXCELERO

Excelero enables enterprises and service providers to design scale-out storage infrastructures leveraging standard servers and high-performance flash storage.

Founded in 2014 by a team of storage veterans and inspired by the Tech Giants' shared-nothing architectures for web-scale applications, the company has designed a Software-Defined Block Storage solution that meets performance and scalability requirements of the largest web-scale and enterprise applications

With Excelero's NVMe, customers can build distributed, high-performance Server SAN for mixed application workloads. Customers benefit from the performance of local flash, with the convenience of centralized storage while avoiding proprietary hardware lock-in and reducing the overall storage TCO. The solution has been deployed for hyper-scale Industrial IoT services, machine learning applications and massive-scale simulation visualization.

## ABOUT MELLANOX

Mellanox Technologies (NASDAQ: [MLNX](#)) is a leading supplier of end-to-end Ethernet and InfiniBand intelligent interconnect solutions and services for servers, storage, and hyper-converged infrastructure. Mellanox intelligent interconnect solutions increase data center efficiency by providing the highest throughput and lowest latency, delivering data faster to applications and unlocking system performance. Mellanox offers a choice of high performance solutions: network and multi-core processors, network adapters, switches, cables, software and silicon, that accelerate application runtime and maximize business results for a wide range of markets including high performance computing, enterprise data centers, Web 2.0, cloud, storage, network security, telecom and financial services. More information is available at [www.mellanox.com](http://www.mellanox.com).



350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085  
Tel: 408-970-3400 • Fax: 408-970-3403  
[www.mellanox.com](http://www.mellanox.com)