Hewlett Packard Enterprise

Technical white paper

Enabling high-performance Ethernet storage connectivity

With HPE's StoreFabric M-series Storage-Optimized Ethernet Switches

Contents

Abstract

The need for high-performance storage networking is being fueled by rapid deployments of multi-core servers densely packed with virtual machines and a move to all-flash storage arrays capable of blazing speeds. Communication between virtualized servers and ultra-low latency flash-based storage is a challenge for legacy IP and Fibre Channel switching architectures. As a result, there is an increasing demand for an efficient, reliable, storage-aware, high-performance storage network connecting servers and storage. Recent innovations in StoreFabric M-series Ethernet Switches are addressing these demands and are keep pace with new storage trends such as the growing adoption of Cloud and Object storage, something Fibre Channel does not support.

The HPE StoreFabric M-series Ethernet Switches introduces the first true Ethernet Storage Fabric (ESF), with technology and features to support block, file, and object storage, thereby allowing customers to build a converged network capable of simultaneously handling compute and storage traffic. This Ethernet Storage Fabric supports converged and hyperconverged architectures with block, object, and file storage, and supports storage connectivity from iSCSI, to traditional arrays, to the newest NVMe over Fabrics arrays. Optimized for storage networking including dynamically shared buffers designed to handle bursty storage traffic, low latency and predictable performance maximize data delivery and scale-out network architecture—all crucial attributes for today's business-critical storage environments. Additionally, ESF provides support for storage offloads, such as RDMA, to free resources and increase performance. Not only are the M-series Ethernet switches specifically optimized for storage, but they also provide better value than traditional Storage Networking Switches. Simply put, faster storage requires faster networks and the M-series delivers just that.

Changes in the storage market

To say the storage market has been undergoing rapid change is an understatement. The traditional storage market of 10 years ago, the following could generally be assumed to be true:

1. All storage was based on spinning disk[1]
2. High-performance storage was block-accessed, connected via Fibre Channel
3. All enterprise storage was purchased as dedicated, integrated appliances from storage vendors
4. Cloud storage was not significant
5. Fibre Channel was faster than Ethernet

Faster storage demands more bandwidth

Today's market is much different. Increasing use of flash has dramatically raised the storage performance benchmark, where a single NVMe SSD can saturate two or more 10 Gb/s links and all-flash arrays routinely require multiple high-bandwidth interfaces to achieve full performance. The latency of storage devices and arrays has fallen by a factor of 5X to 50X, in some cases as low as 100 us (100 microseconds) instead of several milliseconds, and continues to drop.

Driven by multimedia, virtualization, technical computing, Artificial Intelligence (AI) and the Internet of Things, the bandwidth available on servers and required by workloads has rapidly increased as well. For example, editing and production of 4K video require as much as 20 Gb/s to handle one uncompressed stream.

[1] Flash as cache was introduced around 2008, for example the Fusion-io ioDrive was first shown in September 2007, and the NetApp "PAM" flash cache module was launched in June 2008. The first Pure Storage all-flash array was released in August 2011.

Change in storage types

In addition, most storage capacity growth is in file, object, and hyperconverged infrastructure (HCI) instead of only in block, with analysts reporting that up to 80% of storage capacity is in the "secondary storage" category[2]—this can include online backups, snapshots, replicas, test/development copies, and archive data. The use of analytics is driving the growth of distributed, file and object storage to support cloud, Big Data, and machine learning. Many customers are now deploying software-defined storage, which can be used as primary or secondary storage but is typically scale-out and is more likely to be used for file/object/HCI storage than for block storage. Scale-out and hyperconverged storage typically generate a lot of "east-west" traffic between server nodes and usually require Ethernet networks and a redesign of the network architecture.

The influence of cloud

Cloud continues to grow rapidly both for public and private (on-premises) cloud and has a tremendous effect on the storage market. Large cloud service providers are buying and offering an increasing percentage of storage and generally use Ethernet as a single, converged network technology to connect all compute and storage. For all practical purposes, there is no Fibre Channel (FC) in the cloud,[3] this is driving the adoption and use of Ethernet-connected storage both from vendors and in large enterprises who copy cloud architectures in order to achieve cloud-like cost savings and flexibility. Cloud storage growth also promotes the use of file and object storage which is deployed on Ethernet-based networks, often using a scale-out and software-defined storage paradigm.

Rise of the Ethernet Storage Fabric

Changes in Ethernet technology

Ethernet technology has changed rapidly over the last few years, supporting new speeds and new features. The new speeds of 25, 40, 50, and 100GbE are being adopted very rapidly by servers, flash storage arrays, and in the cloud. Sales of high-speed Ethernet ports are projected to grow robustly over the next several years, and all major server vendors are already shipping 25, 40, and 100GbE adapters for their servers. Ethernet offers three times more bandwidth than Fibre Channel (100 Gb/s vs. 32 Gb/s) at lower latency and cost.

Storage switches must support lossless technology to ensure data delivery, Quality of Service (QoS) and advanced congestion control mechanisms, along with sophisticated telemetry, monitoring, and management tools. The high volume of Ethernet port shipments and a large number of competing vendors are driving a rapid advancement in Ethernet technology as well as falling prices, making Ethernet networks viable and cost-effective option for interconnecting storage.

Flexibility to support any storage deployment

Changes mentioned above in the storage market are driving a shift towards Ethernet to provide Storage connectivity. Ethernet offers more performance, lower pricing, and more flexibility than FC, which can be used only for block storage and its price typically precludes its use for secondary storage such as backup, data mining replicas, test/dev copies, or archives. Very large customers don't want to pay for new expanded or duplicate storage networks built on expensive hardware, as they would still need to build an Ethernet network for compute, management, and HCI traffic. Similarly, small customers don't want to pay high FC prices for new storage networks and frequently don't have on-site expertise.

Most new storage deployments on cloud and/or HCI are using Ethernet connectivity and the trends indicate an increasing percentage of networked storage will use Ethernet going forward. So while existing FC SANs and InfiniBand fabrics will generally remain in place for legacy applications running block storage, customers both large and small are looking to build their new storage networks on more flexible and cost-effective Ethernet Storage Fabrics.

[2] Reference Wikibon and whatever study the Scality folks quoted.
[3] Some public clouds are built on InfiniBand and a few small clouds use Fibre Channel, but none of the hyperscale providers or large Web 2.0 vendors use Fibre Channel for hosted data.

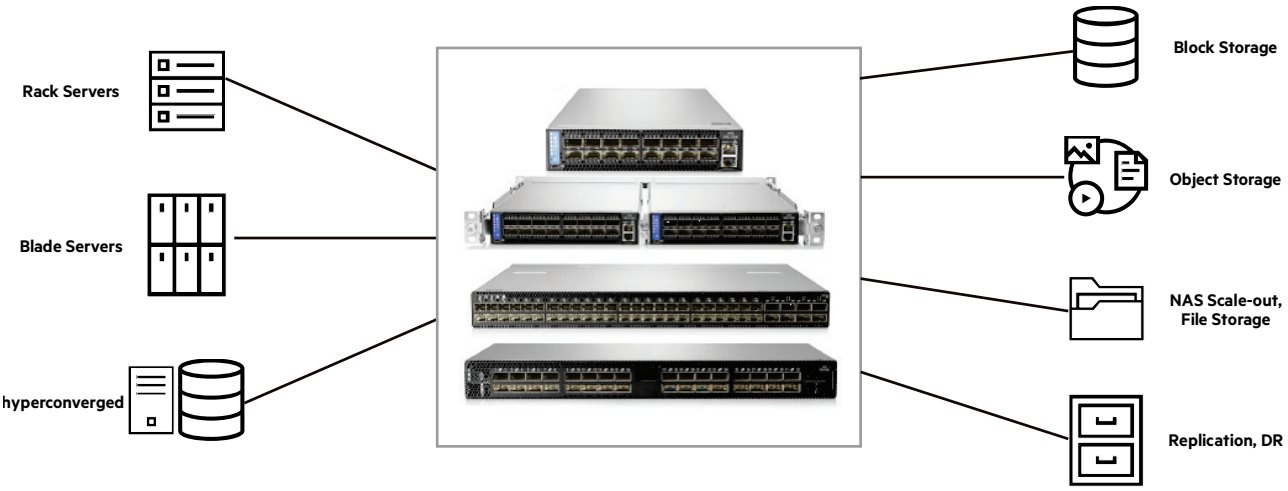

Figure 1. Ethernet supports any kind of servers, storage, and data type, as well as different tiers of storage.

Ethernet Storage Fabric

The ideal switch for storage would incorporate support for all storage types; file, block and object, and include support for HCI and cloud infrastructure and software-defined storage. Similarly, it should maintain high-bandwidth at low latency, ensure delivery of data, and support new media types such as NVMe. Ethernet is ideal to implement a true unified, converged fabric that can simultaneously support compute, communications, and storage traffic efficiently and cost-effectively.

The rise of HCI, such as HPE SimpliVity, is a great example of the efficiency, ease of deployment, and scalability that is capable when using a converged fabric. Combining your entire infrastructure into one simple, yet flexible fabric allows for reduced cost and complexity of your IT environment while delivering the bandwidth and reliability hyperconvergence requires. The true power of hyperconvergence comes from full consolidation of software and hardware, which is made possible by deploying on a unified, purpose-built converged network infrastructure capable of supporting the demanding requirements of today's high-performance storage. This is what is referred to as an Ethernet Storage Fabric. New switches dedicated to storage and HCI clusters should provide the following characteristics to be considered an Ethernet Storage Fabric:

1. **High-performance:** Including high bandwidth, non-blocking, low latency support for compute, communications and storage traffic
2. **Storage Fabric Intelligence:** Able to understand, manage, and monitor storage traffic, as well as integrate with storage and network management tools
3. **Efficiency:** This includes port density, power usage efficiency, flexibility of use, and price

In addition, Ethernet storage switches should be easy to purchase, scale, upgrade, and support, as well as backed by a trustworthy vendor with global reach and strong service offerings.

Figure 2. Ethernet Storage Fabric characteristics.

HPE M-series Ethernet Switch: Ideal for Ethernet Storage Fabrics

Storage-optimized with high performance

The new HPE StoreFabric M-series is the first storage-optimized Ethernet switch capable of providing an Ethernet Storage Fabric (ESF). It offers the highest performance measured in both bandwidth and latency, supporting from 8 to 64 ports per switch with speeds from 1GbE to 100GbE per port. All the M-series models are non-blocking and allow enough uplink ports to build a fully non-blocking fabric. The latency is not only the lowest (300 nanoseconds port-to-port) of any generally available Ethernet switch, but the silicon and software are designed to keep latency consistently low across any mix of port speeds, port combinations, and packet sizes.

M-series uses an intelligent buffer design to ensure buffer space is allocated to the ports that need it most. This ensures data I/O is treated fairly across all switch port combinations and packet sizes. Other switch designs typically segregate their buffer space into port groups, which makes them up to 4-times more likely to overflow and lose packets during a traffic microburst. This buffer segregation can also lead to unfair performance where different ports exhibit wildly different performance under load despite being rated for the same speed.

Storage Intelligence

The M-series family have features specific to optimizing current and future storage networking. These include support for Data Center Bridging (DCB), including DCBX, Enhanced Transmission Specification (ETS), and Priority Flow Control (PFC). iSCSI traffic can be specifically classified and prioritized using the iSCSI well known destination port 3260 based match criteria combined with DCB advertisements of the iSCSI application TLV. And the switches are designed to integrate with storage and network management tools as well as run a container on the switch to provide storage specific services.

Storage optimizations in the HPE M-series include the following:

- Ability to expand and attach beyond the primary array
- Designed to support Ethernet storage connectivity to HPE 3PAR, HPE Nimble storage, HPE SimpliVity, and MSA arrays, as well as HPE storage partners Qumulo, Scality, and Cohesity
- DCB, including DCBX, ETS, and PFC
- iSCSI-TLV
- Support for all Ethernet speeds including 1, 10, 25, 40, 50, and 100GbE
- Lowest latency of any mainstream Ethernet switch
- Differentiated storage ports—ports can be locked down for storage traffic
- The highest levels of (predictable) performance
- Zero avoidable packet loss regardless of packet size
- Performance fairness across any combination of ports
- Shared smart buffer
- Smart cut-through, allowing rapid packet forwarding even with mixed speeds

Efficiency

HPE's Ethernet storage solutions using M-series are specifically optimized for storage efficiency to offer unmatched performance and value. On a price-performance basis, the M-series delivers 3 times the performance at 1/3 the price of traditional storage networking solutions, they have the lowest power consumption of any commercial Ethernet switch and high port density.

Special half-width form factor of the M-series SN2100M allows for two switches to be deployed side-by-side for high availability and up to 128 ports in one rack unit (RU) of space, making them optimal for storage and HCI deployments.

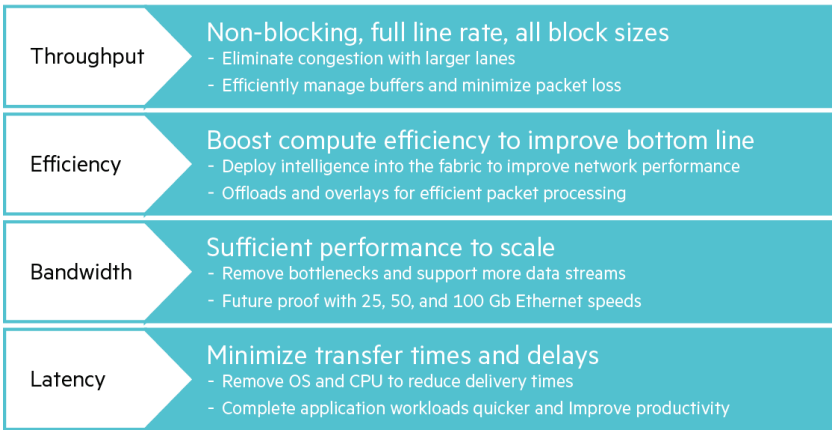

Figure 3. HPE StoreFabric M-series Ethernet Switches provide a solid foundation for an Ethernet Storage Fabric for hyperconverged and storage environments.

Future proofing

The M-series switches offer future proofing for customers who will ultimately require higher densities or upgrades to 40 and 100 Gb/s topologies. Density can be increased in all M-series switches with high-speed ports through unique breakout cables, which can be used to split out individual switch ports to multiple device ports, increasing port density and allowing for more connections. Additionally, future proofing can be achieved through the following:

- Physical connectivity: The same box supports 1GE—100GE
- 10GE connectivity today and upgrade to next generation of Host/Target connectivity with 25GE, 40GE, 50GE, 100GE
- Software/OS: The same box supports iSCSI-based connectivity today and NVMe over Fabrics for future target arrays
- Supports containers to run custom applications for storage management or services

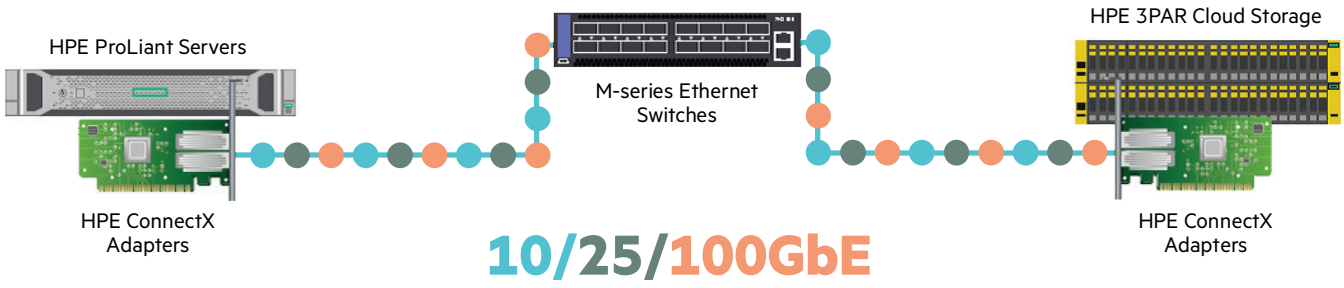

Figure 4. Investment protection with M-series Ethernet Switches.

Purchase flexibility

The M-series switches feature flexible port licensing, allowing customers to license half the ports on the switch and activate the remaining ports whenever desired with a simple software license. This lowers the entry cost for small storage and HCI clusters which typically need fewer than 8 ports per switch (fewer than 16 ports total when using two switches for HA), and allows easy port expansion when the cluster grows.

Flexible financing—buy as OPEX or CAPEX

HPE Financial Services offers consumption-based pricing via leases and "ports-on-demand" financing options. Leasing allows M-series switches to be consumed with a simple monthly payment instead of as an upfront capital expense, allowing most organizations to classify the fabric cost as a monthly operating expense (OPEX) instead of as a depreciating capital expense (CAPEX). This conserves precious cash for customers and simplifies purchasing and accounting for the Ethernet Storage Fabric. The "ports-on-demand" option allows customers to pre-provision Ethernet network capacity while only paying for the ports used—periodically HPE adjusts the monthly payment upward or downward as more or fewer ports are actually used. This empowers customers to only pay for the amount of networking they actually need at any given time.

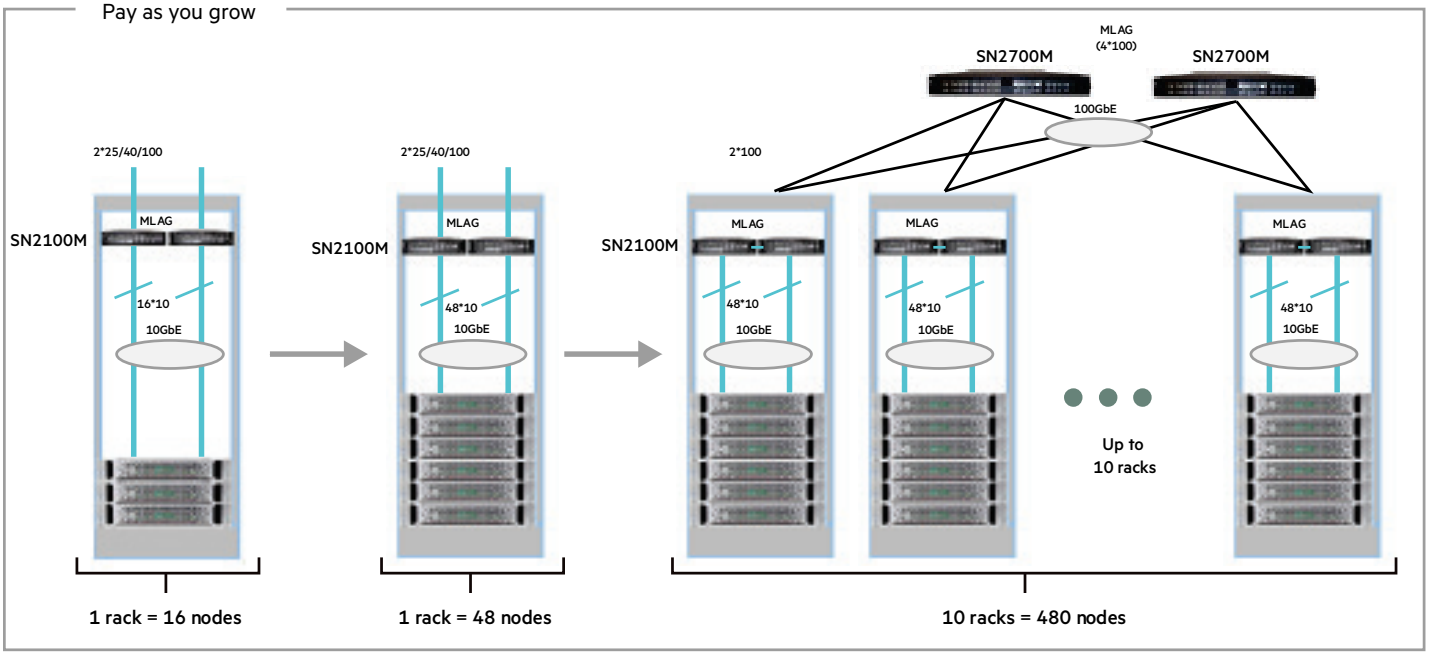

Figure 5. Consumption-based pricing to pay for the ports you need now and license more as needed, or chose attractive lease options.

The risks of using existing switch ports

As more storage networks move to Ethernet, more ports are needed, often at higher speeds including 25, 40, 50, and 100GbE. It doesn't make sense to use existing ports on general purpose switches, nor does it make sense to deploy more of these legacy infrastructures for storage.

Standard networking switches are typically not optimized for storage, often suffering from high latency, packet loss, and lack impartial fairness in data I/O transfers which can result in delays, requiring retransmission of data packets and disruption on a storage network. Existing legacy switches are often oversubscribed, either internally or lack enough bandwidth for uplinks to other switches. Using leftover ports on existing switches risks overwhelming the switch backplanes and uplinks. When a switch fills up, new storage nodes or storage clients must attach into different switches on different racks. This means storage traffic must compete with other traffic and risks becoming bottlenecked in the switches, or often needs to travel extra hops over core switches or routers, leading to more chances for congestion, dropped packets, and high latency, which storage administrators want to avoid at all costs.

Existing switches may not support higher speeds now being adopted by servers and storage, such as 25/40/50GbE. When newer nodes arrive, they would be forced to connect using more 10GbE ports instead of a few 25/40/50GbE ports, making network configuration and efficient bandwidth utilization more difficult achieve, or those switches will need to be replaced in a disruptive process. Many general-purpose Ethernet switches are not designed for storage and exhibit higher latency, more congestion, and a higher rate of dropped packets than would be ideal for a storage fabric.

Conclusion

Storage networks built on M-series provide an ideal Ethernet Storage Fabric (ESF) optimized to deliver the highest levels of performance, industry-best latency, zero packet loss, and unique storage aware features and form factors. By leveraging ESF it's possible to build a converged network capable of simultaneously handling compute and storage traffic, iSCSI storage networking as well as support for file, block, object and the latest NVMe over fabric storage. Optimized for storage networking including dynamically shared buffers, low latency and offering predictable performance makes the M-series ideal for storage environments. Future-proof your storage environment with support for faster speeds, new protocols like NVMe-oF, and the ability to add storage services by either running them on the switches or integrating with the Switch OS. Put your storage on the fast-track with HPE StoreFabric M-series Ethernet Switches and an Ethernet Storage Fabric.

Learn more at

HPE Storage Networking
HPE StoreFabric SN2100M
HPE StoreFabric SN2410M
HPE StoreFabric SN2700M

Sign up for updates

a00038928ENW, December 2017