FIBRE CHANNEL

SOLUTIONS GUIDE

2016

FCIA

FIBRE CHANNEL INDUSTRY ASSOCIATION

fibrechannel.org

Fibre Channel & FCoE

Powering the next generation private, public, and hybrid cloud storage networks

ABOUT THE FCIA

The Fibre Channel Industry Association (FCIA) is a non-profit international organization whose sole purpose is to be the independent technology and marketing voice of the Fibre Channel industry.

We are committed to helping member organizations promote and position Fibre Channel, and to providing a focal point for Fibre Channel information, standards advocacy, and education.

Today, Fibre Channel technology continues to be the data center standard for storage area networks and enterprise storage, with more than 80 percent market share.

CONTACT THE FCIA

For more information:
www.fibrechannel.org
info@fibrechannel.org

Table of Content

Fibre Channel's Growth Trend
Mark Jones, FCIA President, Director of Tech Marketing, Broadcom Limited

It's amazing to reflect on where Fibre Channel is today since the completion of the standard in 1994. From that time until now we have witnessed some amazing changes in the computing industry, no longer is there a single solution that fits all but continuous innovation has produced technologies and products that more accurately service the needs of a broad array of compute scenarios.

One trend that everyone can agree on is that the exploding growth of data is a constant that will challenge computing well into the future. Since 2001 more than 107.1 Million Fibre Channel ports[1] have shipped to storage customers and it is estimated that 46 million ports are in current operation in datacenters today[2]. Fibre Channel continues to be the most widely used transport for enterprise storage and is estimated to store nearly 25 Exabytes of data in 2016[3]. In fact, it is predicted that Fibre Channel will continue to lead enterprise storage capacity by 112% over all other block storage protocols at the end of 2019[3].

FIBRE CHANNEL: TIMELINE

Future: 64/256FC

Cloud Ready

Optimized for Flash

2015: Gen6 32/128FC

2012: 16Gb FC

Virtualization NPIV

2009: FCOE

Converged Networks

2008: 8Gb FC

2005: 4Gb FC

Arbitrated Loop

2001: 2Gb FC

1997: 1Gb FC SAN products emerge

Fabric Services

1988: Work begins on protocol

1. Dell'Oro June 2016
2. Dell'Oro June 2016 – assuming a typical five year hardware refresh cycle
3. IDC Worldwide 2015 External Storage Forecast

FIBRE CHANNEL INDUSTRY ASSOCIATION

The endurance of Fibre Channel as the go-to transport to carry data in the datacenter isn't by accident or the lack of other technology challengers; it was explicitly designed to meet the requirements of future storage networking needs by experienced industry professionals from the top companies in computing. In fact, Fibre Channel's continued popularity are directly attributable to a series of distinctive design attributes:

1. **Designed for Storage** – Unlike networks that are adapted for storage traffic, Fibre Channel was designed for storage since its inception. Fibre Channel was designed to ensure that all storage data arrives to its network destination without loss at extreme throughput and low latency, while operating at high utilization rates without data collisions or the need for retransmission.

2. **Designed for Compatibility** – Datacenter configurations often evolve organically, and other networking environments reflect a more ad hoc philosophy over time. This can lead to compatibility and interoperability issues. However, stringent industry standards require all conforming Fibre Channel products to be backwards compatible to at least two generations and ensure multi-vendor compatibility.

3. **Designed for Performance** – An industry roadmap has guided Fibre Channel to doubling speeds every three-to-four years, matching typical datacenter equipment refresh cycles. This means that as datacenters grow, the highest performance storage networking technology is always available for deployment. Current Gen6 products are available at 32GFC single lane and 128GFC multi-lane configurations for providing over 25GB/s of bi-directional data throughput.

4. **Designed for the Future** – As a transport protocol, Fibre Channel predominantly transports SCSI commands, but it was designed to support other upper level protocols as well. FICON is an example of a common alternate transport protocol often used. Future transports such as NVMe over Fabrics are in development for Fibre Channel and will allow flash storage devices to run native NVMe over a shared Fibre Channel network.

Overall, 2016 is a year for milestones for FC in the storage industry. Production shipments of 16GFC products have been ramping up and for the first time surpassed shipments of 8GFC. Gen6 Fibre Channel HBA and switches are just entering the marketplace from a variety of vendors.

We expect the draw to higher interface speeds will increase with the adoption of all flash storage products and the continued performance evolution of flash technologies and transports such as NVMe. Data center customers will continue to rely on the proven reliability record and performance of Fibre Channel as they plan deployments of these technologies. As sales of Gen6 Fibre Channel products accelerate and the FC-NVMe standard completes this year, the development of the next Fibre Channel speed (64/256GFC) is already underway. The future of the Fibre Channel industry is as bright as ever.

Fibre Channel and NVMe

The NVMe (Non-Volatile Memory Express) specification is being adopted as the interface of choice for high performance flash or SSD drive manufacturers. Used as a PCI Express storage device interface where SSDs connected directly to a server's PCIe complex, bypassing an operating system's SCSI stack altogether, avoiding significant I/O latency and performance penalties.

FC-NVMe leverages the new NVMe over Fabrics specification that was completed in 2016, enabling the networking of NVMe devices while preserving a significant amount of the performance benefit of local NVMe. FC-NVMe will allow native, end-to-end NVMe communication over a Fibre Channel network while concurrently passing legacy SCSI and FICON traffic. This means that NVMe devices can be deployed into existing Fibre Channel SANs, leveraging the industries philosophy of backwards compatibility and investment protection, robust interoperability and proven FC reliability.

This Solutions Guide provides some insight as to how this purpose-built storage networking technology is preparing customers for continued growth in storage environments, performance, and capacity. Inside, you'll find more information on:

- Roadmap trends
- How Fibre Channel and NVMe over Fabrics work together
- How Flash storage environments have unique challenges that are tailor-fit for Fibre Channel solutions
- How the industry is working to promote compatibility and interoperability, providing customers with ongoing reassurance of the reliability and security that they've come to expect

Without question, the needs and requirements of scalable data center storage will continue to rely on the rock-solid reliability and availability of Fibre Channel. Those who have come to count on the dependability of Fibre Channel – as well as newcomers seeking robust, sound technology for storage – can remain confident in its growth and endurance for Data Center storage needs.

The Fibre Channel Roadmap
A technology's heart and soul is its roadmap
Author: Scott Kipp

Just like the term suggests, a roadmap shows the story of a technology. It is a guide to where it has been, where it is, where it is going and when it is going to get there. The more accurate the roadmap, the more valuable it is to its three primary audiences: the IT user base that deploys the technology, the development, manufacturing and distribution base that supplies the technology, and the industry standards bodies that develop standards for the technology.

When a roadmap is consistent, it provides a reliable guide for suppliers, manufacturers and distributors of products to plan their product development and release cycles based upon the features and timing of the technology migration. Some technology developments outlined in reliable roadmaps are required building blocks for product development. For example, lasers in optical modules need to be developed before the transceiver modules can be developed, eventually used in switches or host bus adapters. With a solid roadmap and standards, multiple companies can develop products in parallel that will eventually interoperate when they reach the market.

FCIA's Roadmap Committee is the committee that produces the FCIA Speedmap in concert with the ANSI INCITS T11.2 Task Group, the standards body that defines Fibre Channel speeds. Since the FCIA collocates with T11 meetings, and its roadmap committee includes many of the key T11.2 standards engineers as well as key Fibre Channel supplier technical engineering and marketing experts, the resulting roadmap is the refined product of an intense iterative process that pinpoints highly attractive market propositions balanced with sound engineering feasibility.

The end results are an official FCIA Speedmap and Marketing Requirement Documents (MRDs)that become T11.2's map of speeds and timelines. The MRDs define sets of features and benefits that are not only feasibly doable within the Speedmap timelines, but also results in actual products delivered in the prescribed timeframe that realize massive market success.

The Fibre Channel Industry Association's roadmap has helped the industry see the future of Fibre Channel for over 15 years. Fibre Channel has always had a clear road ahead where the link speeds double every 3-4 years when the speeds can be cost-effectively doubled. Figure 1 shows the history of Fibre Channel speeds and the future speeds through 2020.

Figure 1: Fibre Channel Speeds
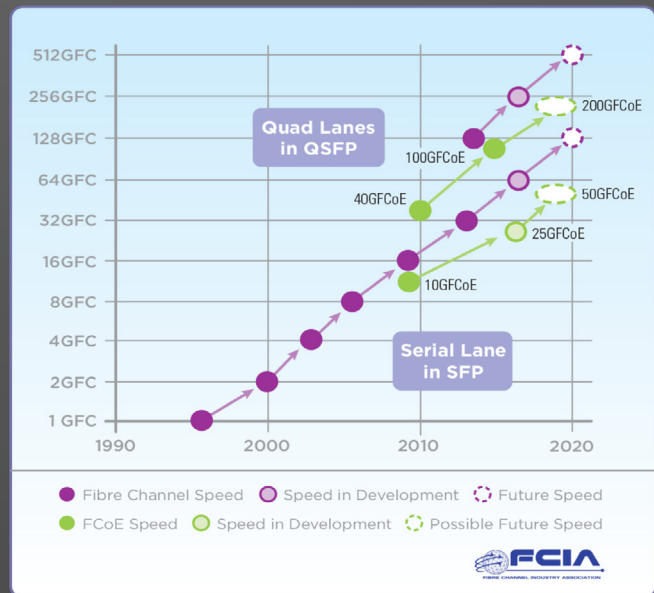
The Fibre Channel industry has embraced servers connecting to Fibre Channel storage via Ethernet with Fibre Channel over Ethernet (FCoE). FCoE uses the Ethernet physical layer and runs Fibre Channel frames and protocol over that physical layer of Ethernet.

Figure 1 shows how Fibre Channel has traditionally used serial speeds that use the venerable Small Form factor Pluggable (SFP) module. The sixth generation of Fibre Channel, known as Gen6 Fibre Channel, uses the SFP28 (an SFP that runs at 28Gb/s) for 32GFC as well as the Quad Small Form factor Pluggable (QSFP28) module for 128GFC. The T11 INCITS technical committee is defining new Fibre Channel speeds that will continue this tradition with 64GFC in an SFP and 256GFC in a QSFP.

The Fibre Channel roadmap doesn't stop there. In Figure 2, the roadmap extends to Terabit Fibre Channel – that's almost 1,000 Gigabits of data per second. Following the 1X/4X paradigm, Fibre Channel and Ethernet plan to double individual lane speeds repeatedly over the next decade. With Fibre Channel's focus on storage in the data center, Fibre Channel will continue to standardize speeds before Ethernet. While Fibre Channel will double speeds from 28Gb/s to 56Gb/s in 2017, Ethernet plans to double 25Gb/s to 50Gb/s between 2018 and 2020.

The trend will continue with Fibre Channel lanes doubling to 112Gb/s and then 224Gb/s. When 4 lanes of these speeds are aggregated, the combined speeds will deliver almost a terabit/second of data for what will be known as Terabit Fibre Channel (1TFC). While Fibre Channel standards are completed in advance of products being released by at least a year, some Ethernet products are released before the Ethernet standard is ratified. This odd comparison of standards and products means that Ethernet products of similar speeds are released at about the same time as similar Fibre Channel products. For example, 25GbE/100GbE products running at 25.78125 Gb/s and 32GFC/128GFC products running at 28.1Gb/s are both expected to be widely available in 2016 for the first time. High speed Ethernet and Fibre Channel products are basically running on similar physical layers.

The Fibre Channel Roadmap has been printed as a physical, folding roadmap and an electronic version can be downloaded at: www.fibrechannel.org/roadmap. The front side of the roadmap shows how Fibre Channel and FCoE is expected to sell over 12.5 million ports in 2015 for the first time according to Dell'Oro – an analyst firm. The backside of the map shows how Fibre Channel is used in data centers around to world to store and replicate data. Fibre Channel continues to grow and provide the most cost effective and reliable links for SANs.

Figure 2: Future Speeds for Fibre Channel and Ethernet
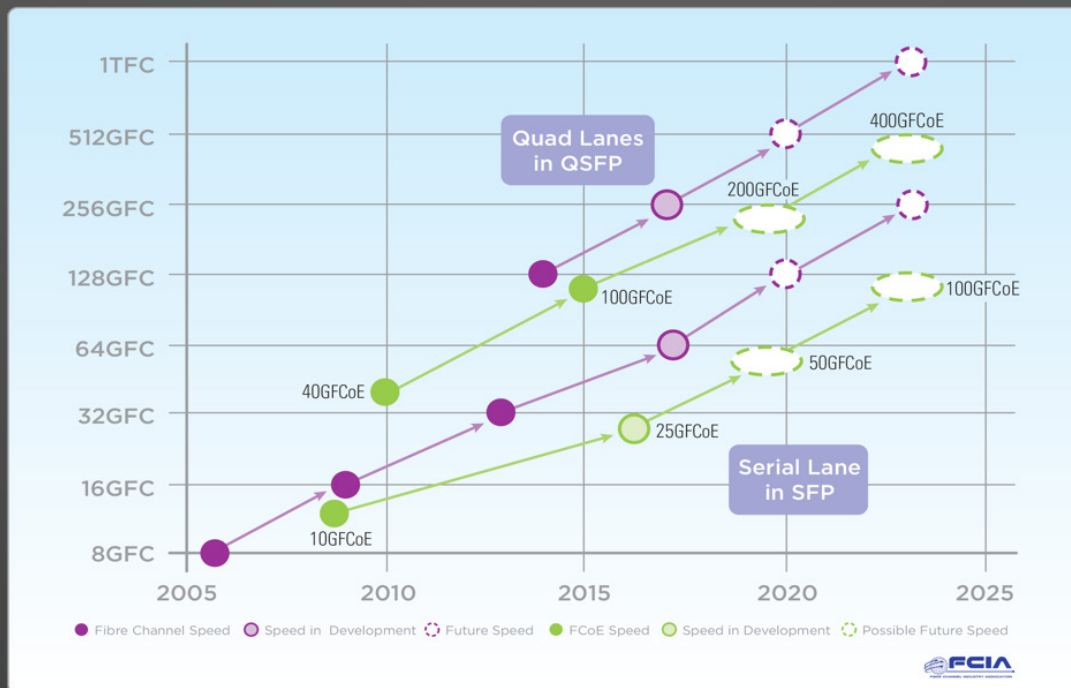
Shared Storage with NVMe
Rupin Mohan, Director R&D, Hewlett Packard Enterprise

Introduction

Storage has served as the core example of disruptive technologies in a paper by Prof. Clayton Christenson of Harvard Business School. Prof. Christenson practically invented the term "Disruptive Technologies" with his HBS paper "Disruptive Technologies, Catching the Wave" in 1995 using the disk storage drive industry as the prime example. The key metrics for the storage industry for the past two decades were 1) size of disk and 2) dollars per megabyte (MB). When the size of the disk fell from 14 inches to 3.5 inches, the leading and established disk manufacturers always were disrupted by a smaller and more nimble startup company.

Fast forward to 2016, and the key metrics, in focus, for storage performance have transitioned to:

1. I/O's per second (IOPs)
2. Bandwidth in MBs/sec
3. I/O operation Latency

IOPS – I/O's per second
– A measure of the total I/O operations (reads and writes) issued by the application servers.

Bandwidth
– A measure of the data transfer rate, or I/O throughput, measured in bytes per second or Megabytes per second (MBPS).

Latency
– A measure of the time taken to complete an I/O request, also known as response time. This is frequently measured in milliseconds (one thousandth of a second) or, for extremely high-performance technology, in microseconds (one millionth of a second). Latency is introduced into the SAN at many points, including the server operating system and HBA, SAN switching, and at the storage target(s) and media and is, in general, undesirable in any system.

While the physical size of the disks have stabilized, the dollars-per-MB has transitioned to dollars-per-gigabyte (GB) now as storage continues to substantially drop in price. That metric is, of course, still very important, but there is clear expectation prices will drop as volumes for new drives increase.

Similar to the electric car re-inventing the automobile industry, there are transformations in the storage industry that profoundly affect advances in latency, IOPs, and bandwidth.

Take, for example, the growing popularity of flash storage and solid state drives (SSDs) in the enterprise storage market. Flash has clearly disrupted the spinning media/ classic disk drive industry. This time, the disk drive industry was disrupted not by size of the disk, but by a complete shift in storage technology. Instead of storing data using ferrite heads on oxide disks, these new flash drives store data on non-volatile memory semiconductor chips. Additionally, there is one big difference with these new SSD disks: disk access is extremely fast, as the media does not have to spin at thousands of RPMs and the disk heads don't have to move around.

Historically, the spinning disk drive was the lowest common denominator in storage performance. Storage experts were quick in calculating the max performance of a virtualized storage array based on the number of spindles (a.k.a disk drives) behind the storage controllers. However, this tended to mask areas of improvement that were necessary inside the storage stack, areas exposed by the invention and use of SSDs. No longer is the spinning disk drive the lowest common denominator in storage performance. Now, we find that several other technologies and protocols need to be improved and enhanced.

One of these areas involves the actual method by which the storage media is accessed. Using a protocol designed for spinning disks to access flash leaves us with a key question: can we improve the way that host applications and CPUs can access non-volatile (NV) storage media?

Advent of NVMe

The SCSI protocol has been the bed rock foundation of all storage for nearly three decades and it has served customers admirably. SCSI protocol stacks are ubiquitous across all host operating systems, storage arrays, devices, test tools etc.

It's not hard to understand why: SCSI is a high performance protocol. Flash and SSDs have challenged the performance limits of SCSI as these disks do not have to rotate media and move disk heads. Hence what you find is that traditional max I/O queue depth of 32 or 64 outstanding SCSI READ or WRITE commands are now proving to be insufficient, as SSDs are capable of servicing a much higher number of READ or WRITE commands in parallel.

To address this, a consortium of industry vendors began work on the development of the Non Volatile Memory Express (NVM Express, or NVMe) protocol. The key benefits of this new protocol is that a storage subsystem or storage stack will be able to issue and service thousands of disk READ or WRITE commands in parallel, with greater scalability than traditional SCSI implementations. The effects are greatly reduced latency as well as dramatically increased IOPs and MB/sec metrics.

Shared Storage with NVMe

The next hurdle facing the storage industry is how to deliver this level of storage performance, given the new metrics, over a storage area network (SAN). While there are a number of pundits forecasting the demise of the SAN, sharing storage over a network has a number of benefits that many enterprise storage customers enjoy and are reluctant to give up. Without having to go into details, the benefits of shared storage over a storage area network are:

1. More efficient use of storage, which can help avoid "storage islands"
2. Offering a full featured, mature storage services like snapshots, backup, replication, thin provisioning, de-duplication, encryption, compression, etc.
3. Enabling advanced cluster applications
4. Offering multiple levels of disk virtualization and RAID levels
5. Offering no single point of failure
6. Ease of management with storage consolidation

The challenge facing the storage industry is to develop a really low latency storage area network (SAN) that can potentially deliver these improved I/O metrics.

The latency step function challenge is described in the following chart:

Latency Step Function Chart
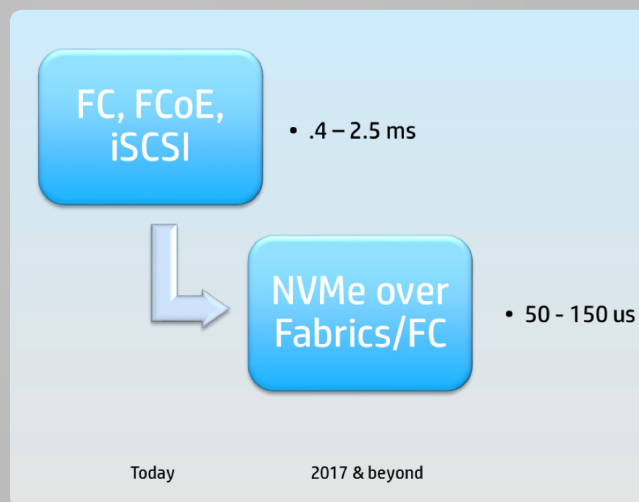
Latency Step Function

There are two new projects in development in the industry that can solve this challenge. These two projects are:

1. **NVMe over Fabrics** – Defined by the NVM Express group
2. **NVMe over Fibre Channel (FC-NVMe)** – New T11 project to define an NVMe over Fibre Channel Protocol mapping

NVMe over Fibre Channel is a new T11 project that has engineers from leading storage companies actively working on a standard. Fibre Channel is a transport that has traditionally solved the problem of SCSI over longer distances to enable shared storage. Fibre Channel, in a simple way, transports SCSI READ or WRITE commands, and corresponding data statuses over a network. The T11 group is actively working on enabling the protocol to compatibly transport NVMe READ or WRITE commands over the same FC transport.

Conclusion

NVMe will increase storage performance by orders of magnitude as the protocol and products come to life and mature with product lifecycles. Transporting NVMe over a network to enable shared storage is an additional area of focus in the industry. Using an 80/20% example, Fibre channel protocols solve 80% of this FC-NVMe over fabrics problem with existing protocol constructs, and the T11 group has drafted a protocol mapping standard and is actively working on solving the remaining 20% of this problem.

Improving High Throughput Applications Performance with Gen6 Fibre Channel

Mark Jones, FCIA President, Director of Tech Marketing, Broadcom Limited

In early 2015, the Fibre Channel Industry Association (FCIA) announced the completion of a collection of new INCITS T11 standards referred to as Gen6 Fibre Channel. As the name suggests, Gen6 points to the sixth generation on the Fibre Channel roadmap fundamentally defining the single lane speed of 32GFC. As has been tradition with Fibre Channel every new speed step doubles the link throughput over the previous generation with the current 32GFC capable of 3200MB/s of single lane half-duplex data throughput.
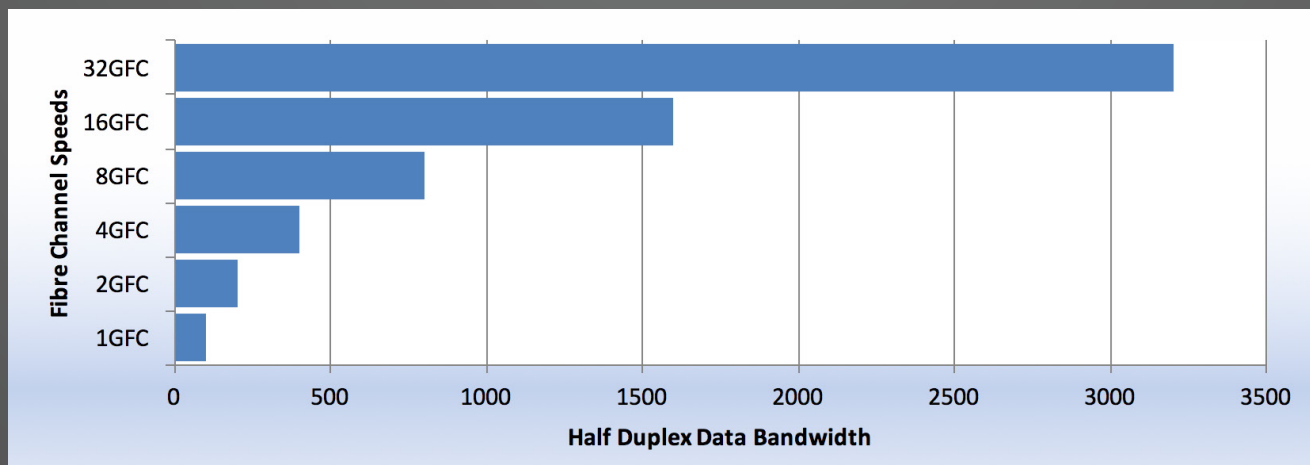
Gen6 Fibre Channel products are now being shipped today by most of the major industry suppliers and we will see many major server and storage vendors shipping solutions with these components in the near future.

Because Gen6 Fibre Channel is backwards compatible at least two generations of speeds, customers have the opportunity to use these products at the full 32GFC speed or connect Gen6 products into previous generation products. Auto-speed negotiation means that the connection will operate at the highest common speed between the two connections.
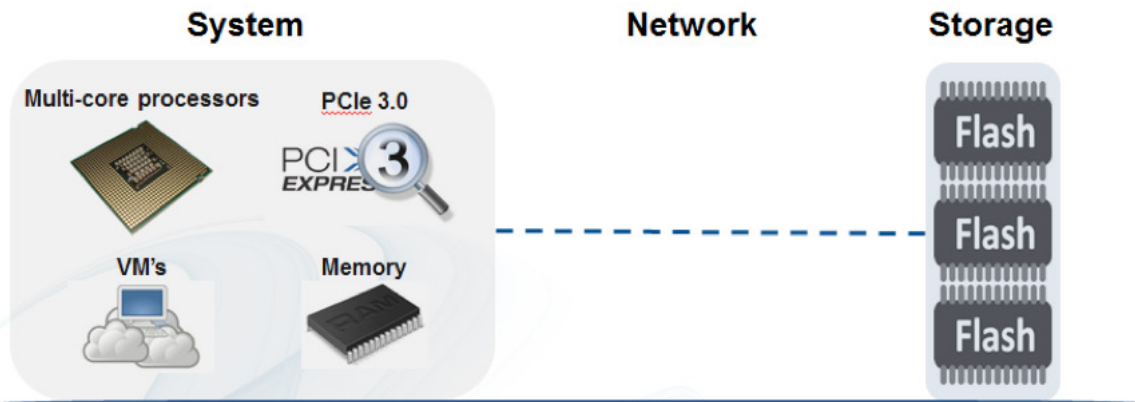
The connection combination for maximum application performance is going to occur when Gen6 is deployed at 32GFC from server to switch and switch to storage, this will allow for end to end throughput of up to 3200MB/s on a single lane. This same application performance level is possible to be achieved in mixed speed configurations, most storage arrays ship with multiple target ports for redundancy and performance, using multiple lower speed (16GFC or 8GFC) target ports in a load balanced multipath configuration will provide the same throughput level to the application. As a result, there is no reason to delay adopting Gen6 products now as they become available for both maximum future-proofing as well as achieving maximum application performance.

Achieving optimum application performance is a balancing act between server and storage performance. In the last 20 years there has been continuous improvement within the server: CPU speed/cores, memory and IO performance have all improved at staggering rates. Storage array performance has improved at a slower pace until recently; the physical performance limitations of spinning disk drives has meant that delivering performance levels that match the server could only come from expensive solutions with large spindle counts and cache sizes. This imbalance between server and storage has kept the pressure off of the SAN link speed which often has had more than adequate performance to deliver the data at a pace that traditional storage arrays were capable of.

Solid State Drives (SSD), on the other hand, and all flash technologies in storage arrays has brought tremendous performance capabilities to low- and high-cost storage segments alike, has quickly becoming mainstream, and will likely surpass spinning disk arrays during the lifespan of Gen6. Now with mega-performance servers and storage on both ends of the SAN the Fibre Channel link speed is now the critical element in the configuration.

FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

Elements of Application Performance
Flash Storage is the Super Catalyst for Gen6

System   **Network**   **Storage**

Multi-core processors   PCIe 3.0

VM's   Memory

Flash
Flash
Flash

- Performance is balancing system, network, and storage capabilities
- System and network capabilities are increasing by a *linear* rate
- Storage (Flash specifically) capabilities are increasing by an *exponential* rate
- As Flash Array growth increases, it has a multiplier effect on Gen6 adoption

With all elements in place, performance-matched servers, storage and Gen6 Fibre Channel the only question remaining is can common datacenter applications take advantage of the performance that Gen6 Fibre Channel is capable of delivering? The FCIA member companies using Gen6 32GFC HBAs and switches connected to all flash storage arrays conducted a series of application tests to prove this point.

As an aid to product readiness the FCIA hosted it's 36th Fibre Channel Plugfest at the University of New Hampshire Interoperability lab, eleven FCIA member companies participated in the event to test Gen6 Fibre Channel products. The Plugfest also was the venue for another milestone for the Fibre Channel industry, it was the first industry wide proof-of-concept demonstration of the new NVMe over Fabrics specification as transported over Fibre Channel, part of the FC-NVMe specification being developed within INCITS T11.
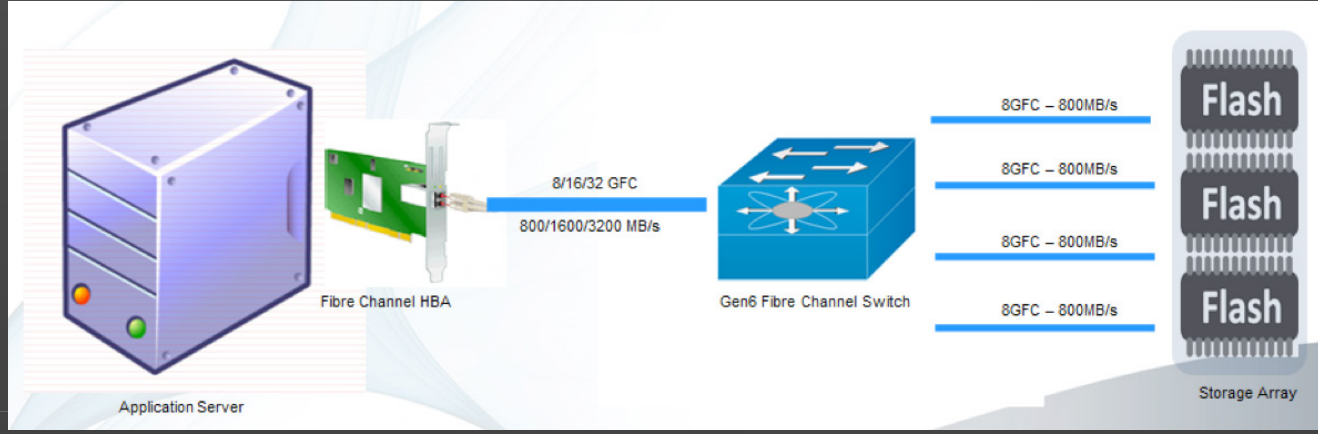
One key result from the testing was that it isn't necessary to have all components in the SAN to be of the same link speed in order to take advantage of the most current Fibre Channel link speed. The configurations of the tests used all-flash storage arrays with Fibre Channel interface speeds of just 8GFC, using four ports of 8GFC the array is capable of equaling the throughput of 32GFC when host multipathing is enabled. Key to the configuration is the Gen6 switch that is capable of 32GFC and is backward compatible of speeds down to 8GFC. In general, three speeds of Fibre Channel are currently dominant in the marketplace, so our application benchmarks compared the performance 8/16/32GFC speeds by using HBA models from each generation.

Database – Data Warehousing

During testing we used a benchmark that closely follows the industry standard - decision support benchmark TPC-H on both Microsoft SQL server 2016 and Oracle 12c database servers. The benchmark uses a series of 22 business oriented queries to a data warehouse instance stored on the SAN attached flash array. The operation must read in the data from the SAN into the database server's memory in order to perform the queries, with the dataset being many times larger than the server's memory. Since little opportunity exists to maintain the data in the server's cache, SAN performance is key to how fast the queries can be completed. Completing the queries in a shorter amount of time means that enterprises can run decision support queries on larger datasets, run more complex queries and deliver greater intelligence to the business.
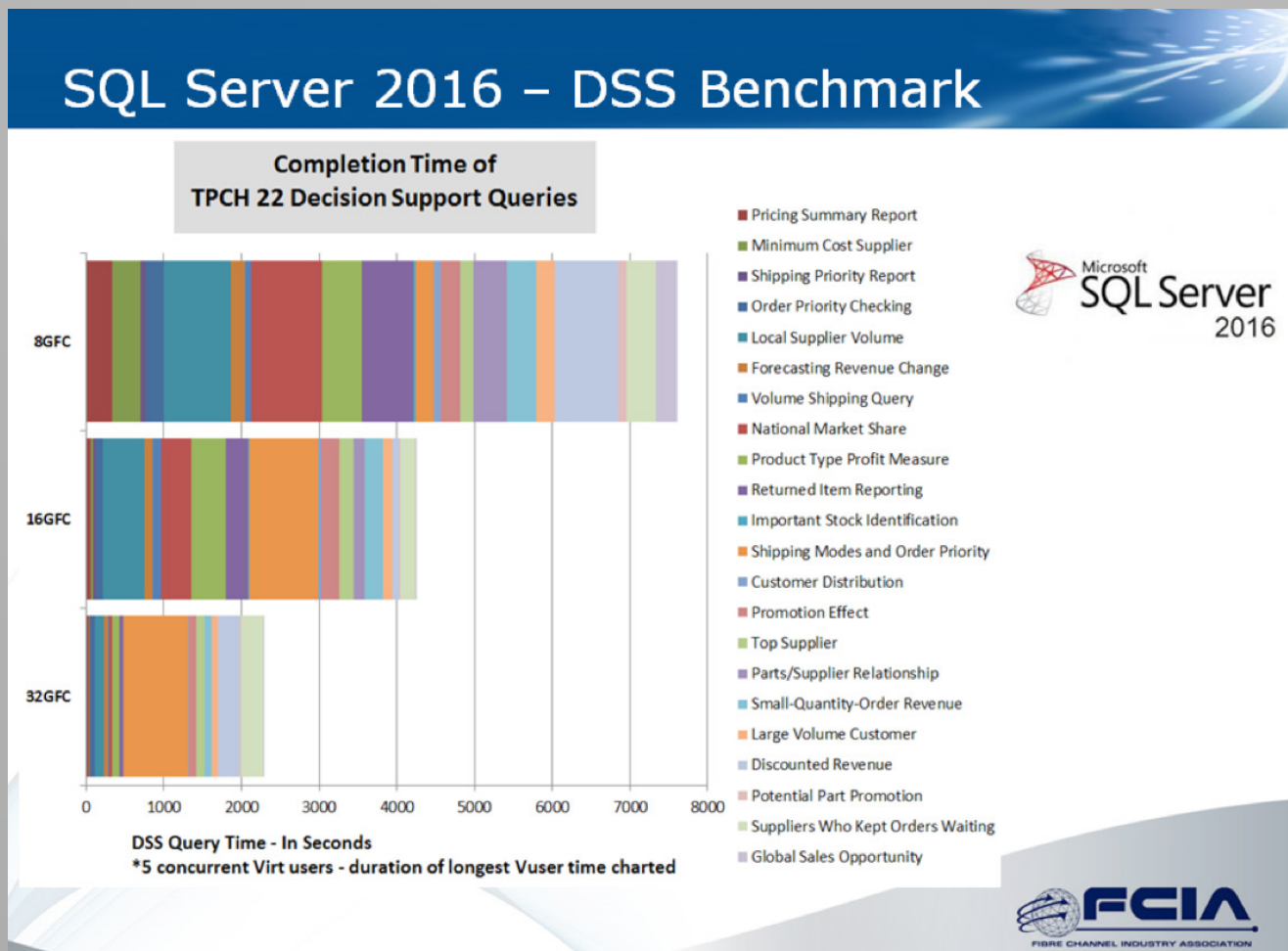
Oracle 12c DSS Benchmark
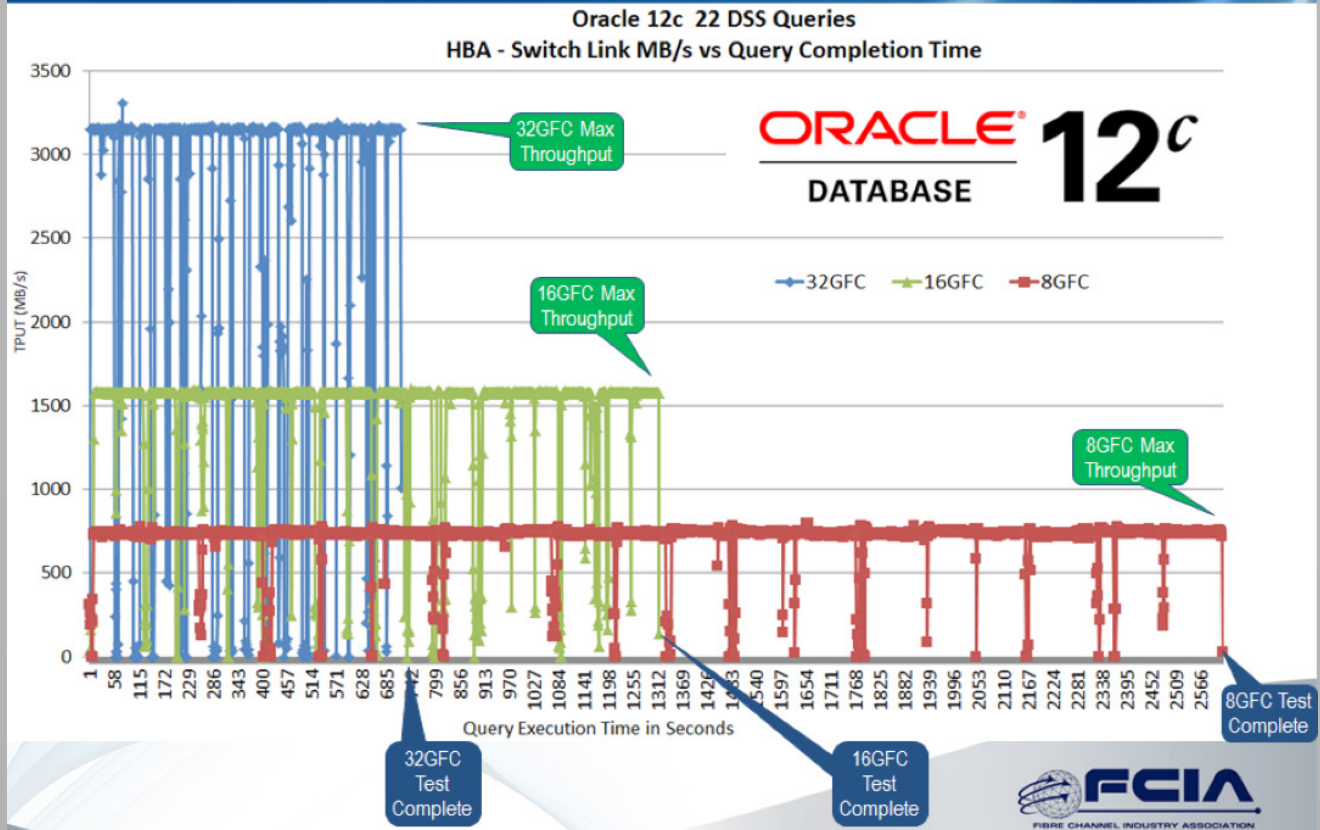
Virtualization VDI "Bootstorm"

A Virtual Desktop infrastructure (VDI) is a datacenter compute environment that hosts desktop computers within a virtual machine running in a centralized virtual server. This infrastructure often stores the desktop data in a centralized Fibre Channel SAN connected to the VDI virtual server. There are significant organizational benefits to deploying VDI from costs saving and centralized IT efficiencies.

Storage IO performance between the VDI virtual server and storage infrastructure can be a critical factor for user satisfaction and efficiencies. A VDI "Bootstorm" is the event when many VDI users attempt to boot their desktops and begin work at the same time, such as when work begins in the morning, after lunch or recovering from an outage.

For our benchmark we focused on "Heavy" VDI users, which are users that will not only boot their environment but also open numerous large applications and load large amounts of data into their applications to begin work. This particular event will put a heavy load on the Fibre Channel link between the VDI server and storage array, so in our benchmark we compared the time for the boot operations to complete when 8/16/32GFC speeds are used.

As can be seen in the figure below, the results confirm that using Gen6 32GFC can help organizations VDI users become more efficient, potentially support more desktop users per server (VM Density) and ultimately lead to more datacenter cost savings.
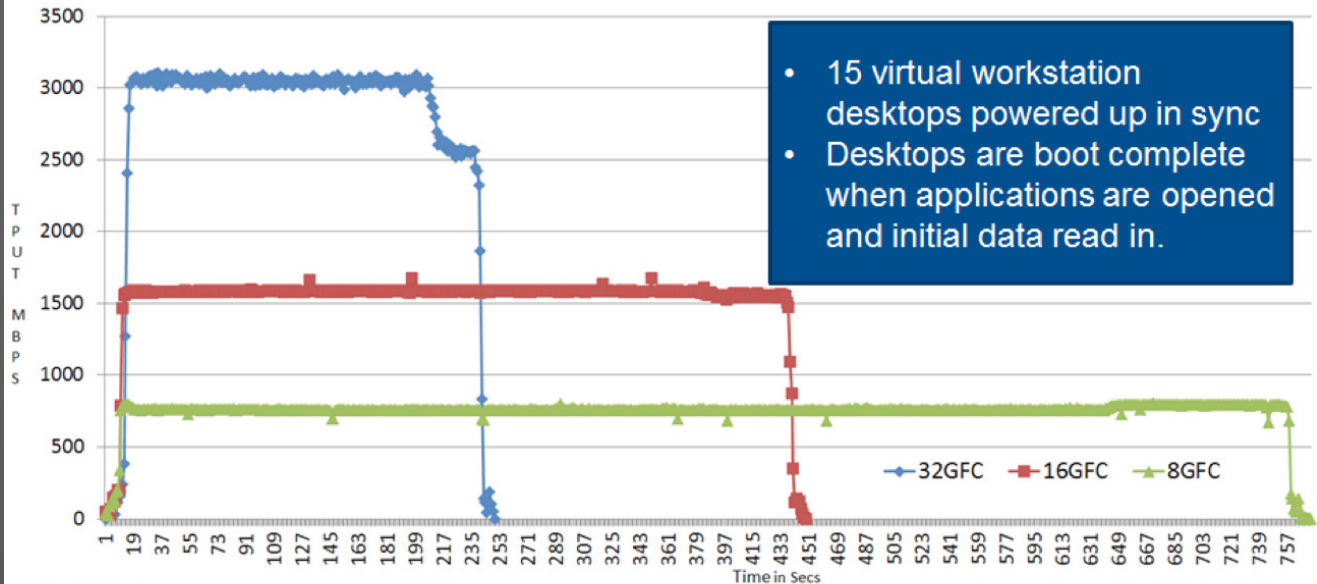
Virtualization VM Storage Migration

Live migration is one of the great benefits of server virtualization. It allows for actively running virtual machines to move to another nearby virtual server for resource-balancing and serviceability.

Moving VMs beyond their local resource pools or clusters is the next level of benefit but it involves moving the storage out of the current storage resource and into another, which can affect the storage I/O connections and can cause congestion to other running virtual servers and VMs. Storage migration is a feature offered by all the major server virtualization suppliers and our benchmarking has again shown that Gen6 32G Fibre Channel can offer significant performance improvements.

As indicated in the graphic below, our benchmarking used the storage migration features of these virtualization environments to time how long it takes to migrate a number of virtual machines storage pools from one storage array to another. Again, we changes the HBA to Switch link rate by changing HBA models that support 8, 16 and 32GFC speeds within the virtual servers.

VIRTUALIZATION: VM STORAGE MIGRATION
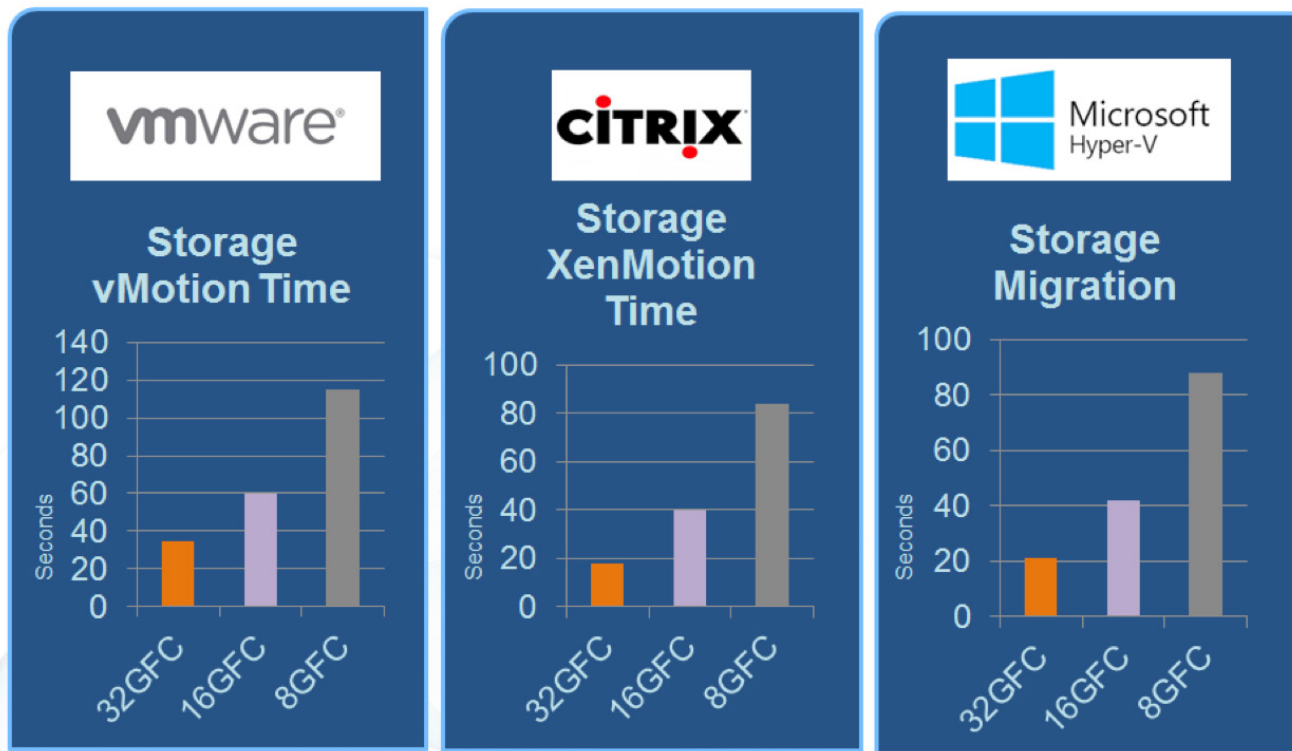
With these benchmark results we have made the case that Gen6 Fibre Channel should be considered for many common high storage throughput applications. Also we conclusively showed that end-to-end speed matching of your SAN is not necessary in order to take advantage of application benefits today with Gen6. Gen6 Fibre Channel products are available today and will be more widely available from major server and storage OEMs in the near future.

Fibre Channel – The Most Trusted Fabric Delivers NVMe

Author: Nishant Lodha, Product Marketing, QLogic Corp

The Evolution of Disk and Fabric

Ever since IBM shipped the world's first hard disk drive (HDD), the RAMAC 305 in 1956, persistent storage technology has steadily evolved. In the early 1990s, various manufacturers introduced storage devices known today as flash-based or dynamic random access memory (DRAM)-based solid state disk (SSD). The SSDs had no moving (mechanical) components, which allowed them to deliver lower latency and significantly faster access times.

While SSDs have made great strides in boosting performance, its interface—the 6Gb per second (Gb/s) SAS 3 bus—began to hinder further advances in performance. Storage devices then moved to the PCI Express (PCIe) bus, which is capable of up to 500 MB/s per lane for PCIe 2.0 and up to 1000 MB/s per lane for PCIe 3.0, with up to 16 lanes. In addition to improved bandwidth, latency is reduced by several microseconds due to a faster interface as well as the ability to directly attach to the chip set or CPU. Today, PCIe SSDs are widely available from an array of manufacturers.

Additionally, PCIe SSDs removed the hardware bottleneck of using the SATA interface, but these devices continued to use the Advanced Host Controller Interface (AHCI) protocol/command set, which dates back to 2004 and was designed with rotating hard drives (HDDs) in mind. AHCI addressed the need for multiple commands to read the data, but SSDs do not have this need. Because the first PCIe SSDs used the AHCI command set, they were burdened with the overhead that comes with AHCI. Obviously, in order to become more efficient, the industry had to develop an interface that eliminated the limits imposed by AHCI.

It wasn't only the overheads of AHCI that challenged PCIe SSD adoption, though, since each SSD vendor provided a unique driver for each operating system (OS), with a varying subset of features, creating complexity for customer looking for a homogeneous high speed flash solution for enterprise data centers.

To enable an optimized command set and usher in faster adoption and interoperability of PCIe SSDs, industry leaders have defined the Non Volatile Memory Express (NVMe) standard. NVMe defines an optimized, scalable command set that avoids burdening the device with legacy support requirements. It also enables standard drivers to be written for each OS and enables interoperability between implementations, reducing complexity and simplifying management.

SCSI and NVMe Differences

While the SCSI/AHCI interface comes with the benefit of wide software compatibility, it cannot deliver optimal performance when used with SSDs connected via the PCIe bus. As a logical interface, AHCI was developed when the purpose was to connect the CPU/memory subsystem with a much slower storage subsystem based on rotating magnetic media. As a result, AHCI introduces certain inefficiencies when used with SSD devices, whose characteristics are far closer to behaving like DRAM than like spinning media.

The NVMe device interface has been designed from the ground up, capitalizing on the low latency and parallelism of PCIe SSDs, and complementing the parallelism of contemporary CPUs, platforms and applications. At a high level, the basic advantages of NVMe relate to its ability to exploit parallelism in host hardware and software, manifested by the differences in command queue depths, efficiency of interrupt processing, and the number of un-cacheable register accesses, etc., resulting in significant performance improvements across a variety of dimensions.

FCIA

FIBRE CHANNEL INDUSTRY ASSOCIATION

NVMe Deep Dive

The NVMe interface is defined in a scalable fashion such that it can support the needs of Enterprise and Client (e.g., consumer devices) in a flexible way. NVM Express has been developed by an industry consortium, the NVM Express Workgroup. Version 1.0 of the interface specification was released on March 1, 2011 and a version 1.2 released on Nov. 12, 2014. Expansion of the standard to include NVMe over Fabrics (including Fibre Channel) was completed in June, 2016.

Today more than 100 companies participate in the definition of the interface, which some very impressive characteristics:

- Architected from the ground up for this and next generation Non-Volatile Memory to address Enterprise and Client system needs

- Developed by an open industry consortium, directed by a 13 company Promoter Group

- Architected for on-motherboard PCIe connectivity

- Capitalize on multi-channel memory access with scalable port width and scalable link speed

As a result of the simplicity, parallelism and efficiently of NVMe, it delivers significant performance gains vs. SCSI. Some metrics include -

- For 100% random reads, NVMe has 3x better IOPS than 12Gbps SAS[1]

- For 70% random reads, NVMe has 2x better IOPS than 12Gbps SAS [1]

- For 100% random writes, NVMe has 1.5x better IOPs than 12Gbps SAS [1]

- For 100% sequential reads: NVMe has 2x higher throughput than 12Gbps SAS [1]

- For 100% sequential writes: NVMe has 2.5x higher throughput than 12Gbps SAS [1]

In addition to just IOPS and throughput, the efficiencies of command structure described above also reduce CPU cycles by half, as well as reduce latency by more than 200 microseconds than 12 Gbps SAS.

NVMe and Fibre Channel (FC-NVMe)

The Fibre Channel protocol (FCP) has been the dominant protocol used to connect servers with remote shared storage comprising of HDDs and SSDs. FCP transports SCSI commands encapsulated into the Fibre Channel frame and is one of most reliable and trusted networks in the data center for accessing SCSI-based storage. While FCP can be used to access remote shared NVMe-based storage, such a mechanism requires the interpretation and translation of the SCSI commands encapsulated and transported by FCP into NVMe commands that can be processed by the NVMe storage array. This translation and interpretation can impose performance penalties when accessing NVMe storage and in turn negates the benefits of efficiency and simplicity of NVMe.

FC-NVMe extends the simplicity, efficiency and end-to-end NVMe model where NVMe commands and structures are transferred end-to-end, requiring no translations. Fibre Channel's inherent multi-queue capability, parallelism, deep queues, and battle-hardened reliability make it an ideal transport for NVMe across the fabric. FC-NVMe implementations will be backward compatible with Fibre Channel Protocol (FCP) transporting SCSI so a single FCNVMe adapter will support both SCSI-based HDDs and SSDs, as well as NVMe-based PCIe SSDs.

Standardization

New T11 project to define an NVMe over Fibre Channel Protocol mapping has been initiated. In August 2014, the INCITS T11 started a FC-NVMe working group and it is expected that the specification will be complete by end of calendar year 2016. Fibre Channel networks are a natural fit for shared remote access to high speed NVMe due to their trusted, lossless and high performance characteristics.

Conclusion

As next-generation data-intensive workloads transition to low-latency NVMe flash-based storage to meet increasing user demand, the Fibre Channel industry is combining the lossless, highly deterministic nature of Fibre Channel with NVMe. FC-NVMe targets the performance, application response time, and scalability needed for next generation data centers, while leveraging existing Fibre Channel infrastructures. FCIA is pioneering this effort with industry leaders, which in time, will yield significant operational benefits to data center operators and IT managers.

Storage Forces
J Metz, R&D Engineer, Cisco

Storage is hard. Well, doing storage *right* is hard. In fact, it's more than just hard – it's *extremely* difficult.

Why? Because storage is the one part of the data center that is absolutely unforgiving. When you lose data, you don't just lose a bit or a byte, you lose something of value. You lose something with meaning. It's no surprise that people are far more paranoid when their data crashes than, say, when their laptop operating system crashes.

There are many different factors that affect long-term solutions to storage problems in data centers. Things change over time – the need for capacity grows, the need for scale, the need for performance, the need for better and easier management. All of this happens while simultaneously requiring the same level of comfort and consideration for being reliable.

This is why there's a temptation to suggest that "new" systems will automatically and necessarily replace long-standing technologies, like the Fibre Channel protocol. And yet, when it comes to being able to wholly substitute the scalability, reliability, and performance for dedicated storage systems, there always seems to be something missing. There always seems to be something that needs to be done that Fibre Channel solves (and has always solved) better and easier "out of the box."

Think about it this way: many storage solutions are non-deterministic. What that means is that a general-purpose network is created, and storage traffic is dropped on top of it, intermixing with other types of network traffic. Networks are not static, however, and there is congestion, "hot spots," and other kinds of challenges that need to be overcome during their lifespan. Without special consideration, these types of network designs affect the storage durability/performance tradeoff in negative ways.

Deterministic storage networks, on the other hand, are specifically architected to handle these kinds of fluctuations as part of their design principles. Generally speaking, when you have storage systems that grow larger, they become more difficult to manage and their performance suffers. It becomes very difficult to keep the high benchmark of quality and integrity of the system, *unless* that system is prepared for those changes from the beginning.

FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

Fibre Channel, being a deterministic storage network technology, has been tested repeatedly and come through with flying colors at both the smallest and very large scales (thousands upon thousands of nodes). Moreover, it's been able to do this *predictably*.

Without question this is why the Fibre Channel protocol has been - and continues to remain - the dominant storage transport technology in data centers. Things change over time. There is a need for more capacity, more performance – all the while having to do so with fewer resources, less money, and fewer people. When the forces that affect the development and growth of data centers start to apply pressure, Fibre Channel has been the one technology that consistently provides the reassurance of predictability and reliability, regardless of which direction those forces take you.

In this solutions guide, we have looked specifically at how Fibre Channel has done this, with the powerful benchmarks that emerged from the 2016 plugfest at the University of New Hampshire. We've seen how the ongoing roadmap has been a constant guide for developers and customers both in the past and in the future, with an incredible vision of reliable storage traffic running at over 1 Terabit. We've also seen how Fibre Channel is a major transport for NVMe, being worked on by both T11 and the NVM Express group.

If there's one theme that runs through this Solutions Guide, it's that reliability and dependability for storage networks is not new, nor is it fading away. Fibre Channel has been, and continues to be, the gold standard for storage networks – regardless of storage device type. The fact that it's applicable to multiple storage types, from SCSI to FICON to NVMe, underscores the well-designed nature of the protocol.

All in all, it's worth reiterating the tremendous value that Fibre Channel has brought to storage networks and the preservation of not only the bits and bytes, but the meaning and significance of data. More to the point, it's worth taking a closer look at the technology that has consistently proven its resilience and reliability, in the past, present and future.