FIBRE CHANNEL INDUSTRY ASSOCIATION

SOLUTIONS GUIDE

2020

FIBRE CHANNEL

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.

CONTACT THE FCIA

For more information:
www.fibrechannel.org • office@fibrechannel.org

TABLE OF CONTENTS

Foreword 3

FCIA's President Introduction 4

The Benefits of NVMe™ & NVMe-oF™ 6

The 2020 Fibre Channel Roadmap 7

Fibre Channel Self-Healing Networks 9

The Need for Fibre Channel Standards 13

Fibre Channel and the Autonomous SAN 16

FOREWORD

By Rupin Mohan, FCIA Marketing Chairman; Director R&D and CTO (SAN), Hewlett-Packard Enterprise

Data center networking continues to be a critical piece of the puzzle in designing technology solutions to solve complex business problems for customers. Customer data and secure access to this data with high availability, reliability and performance sounds like a simple set of requirements but continues to be challenging to deliver in the real world.

In 2020, Fibre Channel continues to remain the premier storage fabric connectivity protocol in today's data centers delivering these requirements. Millions of ports of Fibre Channel are currently deployed by thousands of customers in their data centers around the world. Customers today have a wide variety of storage workloads driven by a similar wide variety of applications. The one common denominator that continues to remain proven and trustworthy to deliver these workloads is the Fibre Channel protocol.

FC-NVMe-2, the second version of the NVMe over Fibre Channel standard, is published and we are seeing products being announced and released in the market throughout the protocol stack. The announcement from VMware that vSphere 7 is supporting FC-NVMe with NVMe over Fabrics is proof that Fibre Channel will continue to remain dominant for SAN fabrics with NVMe-oF. Customers who have made investments in Fibre Channel will leverage the same hardware to run FC-NVMe. That just makes common sense. We have an article on this specific topic in this year's Solutions Guide. Please refer to "The Benefits of NVMe™ and NVMe-oF™".

Fibre Channel was built from the ground up for storage with a full featured highly available distributed name server built into the fabric. This gives FC-NVMe a unique edge over other network protocol standards in terms of rock-solid reliability, unmatched performance and massive scalability. FC-NVMe takes full advantage of this build-in name server, giving it the gold standard for reliability and "five nines of availability," meaning 99.999% of network uptime.

The Fibre Channel roadmap continues to be robust with the development of Gen 8 and other Fibre Channel technologies. Also, in this year's Solutions Guide, there are articles on autonomous SAN and the power of automation and orchestration in Fibre Channel with fabric notifications. Customer problems like fabric congestion are best solved at the protocol layer. These solutions at the protocol layer provide the most optimized and clean solutions without adding additional layers of software in the stack.

We cap off this year's guide with an article on the importance of technology standards in our industry. I hope you enjoy reading this year's Fibre Channel Solution Guide.

FCIA'S PRESIDENT INTRODUCTION

Mark Jones, President, FCIA and Director of Technical Marketing, Broadcom Inc.

Society is living through a difficult time. The crisis of the COVID-19 pandemic has left no one untouched, and the virus continues to claim human lives. We at the FCIA, wish all of our members, colleagues, and their families, good health and hope they stay safe.

In terms of the world's economies, the pandemic is putting industries and businesses under extraordinary stress rarely seen in peacetime. Now more than ever, we are relying on our datacenter storage networks to operate reliably at high levels of utilization as companies adapt to a remote workforce in response to personnel distancing requirements.

Rest assured, nothing has slowed our progress to continually improve Fibre Channel. While we have temporarily suspended meeting face-to-face, we are proud to report we have continued on-schedule with remote INCITS T11 standards meetings, FCIA member meetings and regular FCIA educational webcasts.

There is a resurgence of interest in Fibre Channel that we first experienced a few years ago, accelerated by companies repatriating their cloud storage back on premise and migrating to hybrid cloud solutions using Fibre Channel. Principal market analyst Casey Quillin of Quillin Research, estimates by the end of 2020, over 142M ports of Fibre Chanel will have shipped, and 37M Fibre Channel ports are in use today.

This data points to both the long and loyal customer base for Fibre Channel, as well as the significant increase in recent sales of Fibre Channel equipped solutions. IT organizations are discovering that Fibre Channel storage solutions are still the centerpiece in a balanced datacenter storage design, because they provide the security, reliability and performance for their organization's most premium data. All indications are this trend will continue due to both the massive growth of structured data and the increased focus on security and data governance.

The FCIA maintains a speed roadmap for Fibre Channel, and has forecasted and directed the development for increasing storage network speeds as required by the evolving needs of the datacenter. Since the first 32GFC Gen6 devices went into production in 2016, the amount of world wide data has increased by 4X, according to a report published by Statista. Organizations are going through a Business Intelligence (BI) revolution in which many more employees can access BI tools on their desktops to mine and analyze data from their data warehouses, once an application only reserved for a few data scientists.

Storage arrays have increasingly moved to high performance all-flash designs that have drastically increased the performance of even modest configurations, and server CPU core and memory capacities have increased drastically. We are once again seeing all the elements for the next storage IO transition come together to justify increasing storage network speeds to the next level. In 2018, FCIA announced the INCITS T11 standards committee completed the Fibre Channel Physical Interface (FC-PI-7) standard specifications for 64GFC, effectively doubling the data bandwidth over 32GFC. Since a number of FCIA member companies have released Gen 7 products capable of 64GFC for use when optics become available, FCIA predicts 2021 will be the year of 64GFC Gen 7!

FCIA is really proud how the industry has coalesced around NVMe over Fibre Channel solutions (FC-NVMe). It has been almost five years since the T11 Technical Committee began work on the standard, and the industry is now at a point now where FC-NVMe has near complete solution coverage across the ecosystem. All the component and fabric suppliers and most of the major array vendors are now offering production NVMe/FC options. This year, VMware ESXi 7.0 launched with support for NVMe/FC in-box, improving both IO performance and efficiency.

Work has been completed on the next iteration of the FC-NVMe standard to take advantage of Fibre Channel's inherent sequence level error recovery capabilities. FC-NVMe-2 detects and corrects bit level errors before the protocol detects them, thus preventing the need for the protocol association to reset. INCITS has now published the FC-NVMe-2 standard developed by the T11 Technical Committee, and the enhancements will work their way into production via product software updates.

Another recent important area of work in Fibre Channel standards is fabric notifications. Multiple standards define the purpose and use of Fabric Performance Impact Notifications or FPIN. From within, the fabric conditions effecting congestion or link integrity can be detected and notifications (FPIN) can be sent in-band to end points to be acted upon. Devices registered to receive the FPIN can choose to log the condition, send it to upper level software, or chose to correct the condition within the device.

One example is FPIN-CN. This is a congestion notification for when the fabric detects congestion from a port. The fabric then sends the FPIN-CN to the port where the device receives and corrects the congestion by altering its traffic flow. Because both fabric and end devices are communicating and responding together, we can see the makings of an autonomous SAN.

Another example is FPIN-LI, or link integrity notification. Imagine a case where a link is impaired, but not dead enough for the multi-path I/O (MPIO) software to detect the link is down and failover. Based upon fabric management policy, the fabric could detect the condition and send FPIN-LI in-band to either the target or server, and up their software stacks so the devices could intelligently redirect the I/O path to storage. It's not hard to imagine that how this mechanism could be leveraged for increased SAN self-healing capabilities.

While improved storage management applications and SAN automation will be great aids to help organizations manage their storage resources, there is no replacement for human knowledge. The trend in IT education is moving towards less specialists and more generalists. According to a recent ESG storage trend survey, 62% of new IT positions are being filled by generalists. As more seasoned storage administrators and architects retire, the need for Fibre Channel SAN education increases.

Fibre Channel SAN education is a key initiative within the FCIA and the most popular resource that the FCIA produces for the public. The FCIA BrightTalk webcast channel adds new content bimonthly, presented by technology experts crafting the industry standards for storage networking. The Fibrechannel.org website is also a key resource where you can find articles, blogs and the latest news on Fibre Channel technologies and products. This year the FCIA also launched its YouTube channel highlighting dozens of video presentations organized into playlists according to skill level. There's something for everyone, from Fibre Channel basic to expert courses.

Looking forward to 2021 and beyond, developments continue to position Fibre Channel as the network protocol of choice for datacenter storage. The FCIA roadmap points to next network speed transition, and work is already underway for the follow-on to Gen7 64GFC. Work study groups have also been formed to investigate Fibre Channel's role in a variety of new emerging data technologies.

Please stay up-to-date on Fibre Channel's progress by visiting www.fibrechannel.org, subscribing to FCIA's quarterly newsletter, or following FCIA on LinkedIn, Facebook, or Twitter at @FCIAnews.

The Benefits of NVMe™ & NVMe-oF™

By Casey Quillin, Founder and Principal Analyst, Quillin Research

Debates over the value of Non-Volatile Memory Express (NVMe™) seem to have faded a fair bit. The features and benefits are coming into better focus and enterprise's interest in the potential of coercing every bit of performance (and value) out of their storage infrastructure is undeniable. However, whether we need another article on the subject is probably debatable. So, I am not going to go over all the technology on the topic, but instead provide a brief outline of the benefits of NVMe and NVMe over Fabrics.

Fibre Channel (FC) continues to be a trusted, secure, and reliable network infrastructure, especially for mission-critical storage applications. The use of FC over the last 20 years has proven not only its value, but also its ability to evolve alongside the rest of the datacenter. That is not to say that some of these changes have not dampened FC market revenues, because some have – like Cloud Storage and increased network attached storage (NAS) due to the exponential growth of unstructured data. However, some emerging trends are positives for FC, such as all-flash arrays and NVMe.

All-flash arrays are an important driver for the use of NVMe, as increased speed and performance on these arrays has created bottlenecks at other points in the network. The NAND-based flash in arrays today are non-volatile memory (NVM), or persistent in the retention of stored data even in a powered-off state, and the foundation for NVMe. Since NVMe is designed and optimized for NVM drives, it can take advantage of the increased read and write operations. NVMe also supports a greater number of queues, offers higher IOPs with fewer CPU instructions, reduces latency, and significantly increases overall performance. Also, since NVMe communicates over PCI Express (PCIe) transfer rate caps are considerably higher than SATA interfaces.

So, where does FC fit in this picture? Enter NVMe over Fabrics (NVMe-oF™). The NVMe-oF standard allows for the use of NVMe over a variety of network transports, extending optimized performance to external networked storage. While the NVMe-oF standard is transport agnostic, there are multiple reasons that FC will be the initial solution of choice. The network needs to be NVMe ready (which FC is), it needs to be capable of high-speed and low latency to support all-flash arrays, and it must be stable and reliable enough to ensure that the network transport doesn't, by its use, create new bottlenecks. Those requirements sound a lot like FC. Maybe most importantly, NVMe over FC would allow users, especially large enterprises, to leverage their existing storage infrastructure while they design, deploy, and optimize their storage applications with NVMe. Not having to tear down the existing network is always music to a CFO's ears. The fact that FC SANs are air-gapped is also a benefit in this scenario, since a dedicated storage network absent the pesky employees playing with their iPhones and Facebook, means fewer potential problems to troubleshoot.

Will NVMe over FC always be the network transport of choice? It is hard to imagine that Ethernet will ever go quietly into the night. So, I never count it out. The way to think about that from a market share perspective is to expect the growth to move towards the solutions with the best economics. Ethernet speeds and feeds (and their associated average selling prices) are likely to outpace FC at some point. Network and datacenter evolution will undoubtably continue and which transport that change favors is difficult to predict but it favoring Ethernet is likely. In the meantime, NVMe and NVMe-oF will benefit the FC market and make a case for additional bandwidth and new refreshes.

The 2020 Fibre Channel Roadmap

By Craig Carlson, FCIA Board Member; Senior Technologist, Marvell Semiconductor Inc and Rupin Mohan, FCIA Board Member, Director R&D, CTO, HPE Storage Networking

When companies invest in a technology, they want to know that they will get a return on their investment for years to come. Fibre Channel has had a very accurate roadmap for over a decade, showing the past, present and future of the Fibre Channel physical layer. Fibre Channel has been progressing since 1996 by doubling the data rate every few years and the roadmap shows the progression will continue far into the future.

The ANSI INCITS T11.2 Task Group (T11.2), the standards body that defines Fibre Channel speeds, finished 64GFC in 2018. 64GFC runs 9% faster than 50GbE and has been defined for a Bit Error Ratio (BER) of 1E-15 that is 1,000 times better than Ethernet that has a 1W-12 BER. 64GFC products are expected to ship in 2020/2021. T11.2 is also working on 128GFC that runs 5% faster than 100GbE. 128GFC is expected to be completed in 2021 with products shipping in 2022. The Fibre Channel physical layer will continue to leverage the developments in the Ethernet physical layer. In 2020, FCIA has updated the Fibre Channel roadmap as part of our regular yearly cadence.

An accurate roadmap provides a reliable guide for suppliers, manufacturers and distributors of products to plan their product development and release cycles. The features and timing of the technology migration reflected in the roadmap are based on open standards that are technically stable and complete. 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 transceiver modules used in a switch or host bus adapter. 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 produces the FCIA Speedmap in concert with T11.2. The resulting roadmap is the refined product of an intense iterative process that pinpoints highly attractive market propositions balanced with sound engineering feasibility. It becomes the official FCIA Speedmap and MRDs (Marketing Requirement Documents) for T11.2's map of speeds and timelines. The MRDs define sets of features and benefits that are not only feasible within the Speedmap timelines, but also result in actual products delivered in the prescribed timeframe that realize massive market success.

FCIA'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 future speeds after 2020.

Figure 1: Fibre Channel Roadmap

Figure 1 also shows how Fibre Channel initially used only serial speeds for the earlier generations. These serial speeds have used 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. T11.2 finished the seventh generation of Fibre Channel speeds in 2019 that will continue this tradition with 64GFC in an SFP and 256GFC in a QSFP. The project for the eighth generation of Fibre Channel that supports 128GFC in the SFP is underway in 2020 and is keeping pace with 100 Gigabit Ethernet (100GbE) in an SFP module.

[1] *At the time of this article, the exact 128GFC design parameters have not been finalized and may change.*

The Fibre Channel Roadmap doesn't stop there. In Figure 1, the roadmap extends to Terabit Fibre Channel (1TFC) – that's almost 1,000 Gigabits of data per second. Following the 1X/4X lane 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 at approximately the same time as Ethernet, but Fibre Channel speeds will be 5% faster. While Fibre Channel doubled speeds from 28Gb/s to 56Gb/s, Ethernet plans to double 25Gb/s to 50Gb/s. 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 means that Ethernet products of similar speeds are released at about the same time as similar speed Fibre Channel products. For example, 50GbE products running at 53.125Gb/s and 64GFC products running at 57.8Gb/s are both expected to be widely available in 2020 for the first time. High-speed Ethernet and Fibre Channel products are basically running on similar physical layers.

The physical layers of Fibre Channel and Ethernet are marching at a similar pace now. While Fibre Channel has continuously doubled speeds from generation to generation, Ethernet used to grow by a factor of 10 until 40GbE came along. 40GbE, which is based on 4 lanes of 10G technology, broke the 10X paradigm and opened the door to more moderate steps in speed. Similar to technology progressions like Moore's Law and storage capacity, doubling of lane rates is the new norm. Individual lanes can then be grouped together to form new speeds. While Ethernet continues to use up to 16 lanes for 400GbE router applications, only 1X and 4X lanes are shown in the Fibre Channel Roadmap because these are the only relevant speeds for storage area networks (SANs).

The Fibre Channel Roadmap has been printed as a physical, folding roadmap and an electronic version can be downloaded at: fibrechannel.org/roadmap/. 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.

Besides the roadmap, the FCIA Roadmap subcommittee develops the MRD for new speeds like 128GFC. Important elements defined in the MRD include backward compatibility with previous speeds. For instance, just like 1/2/8/16GFC, and 32GFC edge connectivity, 64GFC and 128GFC are required to be backward compatible at least two generations. These speeds are auto-negotiated with no user intervention required, - i.e., 32GFC ports will automatically run at 8GFC and 16GFC, while 64GFC will automatically run at 32GFC and 16GFC. 128GFC continues Fibre Channel's long history of ensuring total backward compatibility by also operating at 32GFC or 64GFC. This important level of backward compatibility has been and will continue to be a major benefit in Fibre Channel's continued success.

Fibre Channel Self-Healing Networks

By Brandon Hoff, FCIA member, Principal Architect, Broadcom

The good old days: when we talked about congestion but didn't really see it in the wild. When we did, it was a big deal. We would sit down with the customer to work out a solution. By we, I mean the storage, server, HBA and switch vendors to name a few. This was a long and costly process. Teams of engineers going though endless data to pinpoint the root cause so it could be addressed. Then, addressing the issue took another team a lot of time and effort to fix. It is important to note that congestion isn't a Fibre Channel problem, it is a networking problem.

Congested Storage Networks

Today it is worse. All Flash Arrays and NVMe/FC storage arrays can stream 128GFC and soon 256GFC of data from a single HBA. Congestion happens when the server(s) cannot download the data fast enough from the storage network. In a lossy network, the congestion ends up in packet drops that either incur a protocol-level retry or an application-level retry, which add significant latency to the flow and cause application performance problems.

For lossless networks, the data fills up the network and consumes hardware resources across the fabric. The network of the past could easily identify when device behavior had an adverse impact on the performance of the specific device performance and other performance, but it was only able to do some of the mitigation by itself. Now with the ability to communicate behavior with adverse impact to the devices and the ability to react accordingly, the end-to-end system becomes self-healing.

The problem occurs when a workload's "eyes are bigger than its stomach" and the workload asks for too much data and becomes the Bully workload. This workload consumes hardware resources across the network. For lossless networks, buffer credits in the fabric switches are consumed by the Bully workload. Victim workloads, other workloads connected to the SAN, are starved and their performance is significantly impacted. It is becoming common for network administrators to plan for Bully Workloads to pop up in their storage network.

Causes of Congestion

When a new installation is architected, it is right-sized. That means it has enough compute, memory, storage, and bandwidth to handle the workloads deployed on the hardware platform.

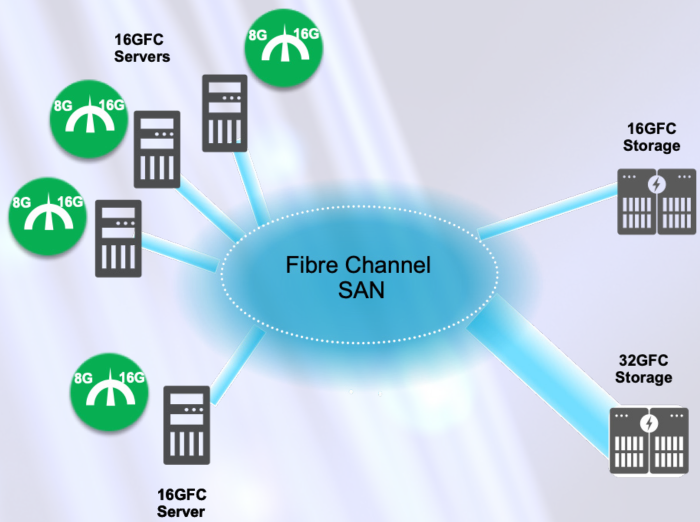


Figure 1: Well-architected data center

Over time, things change and there are two key causes of congestion:

1. A workload gets too big for its hardware footprint and results in an over-utilized server. It consumes either 100% of the CPU, memory, PCIe bandwidth, or Fibre Channel bandwidth. Examples of this are an administrator moving too many VMs onto a server or from the normal growth of data used in applications.

2. A 32GFC All Flash Array is installed on the network and now can feed the demands of over-utilized servers.

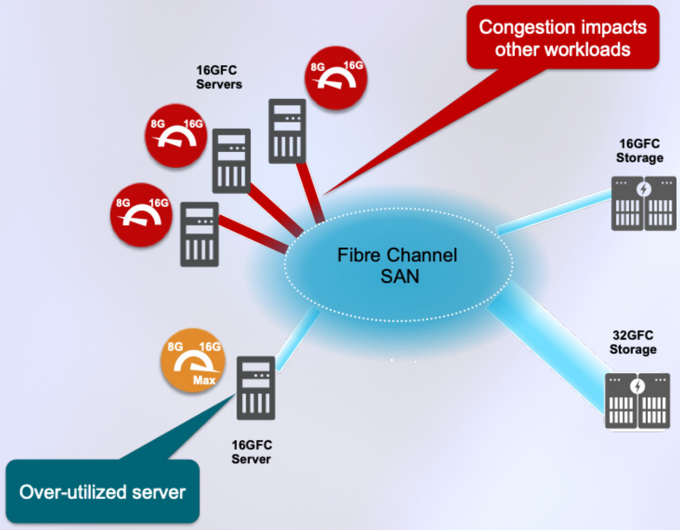


Figure 1: Well-architected data center

The over-utilized workload, or the Bully workload, looks like it is running well, and the lossless network is doing its job, so it isn't obvious that it is causing problems.

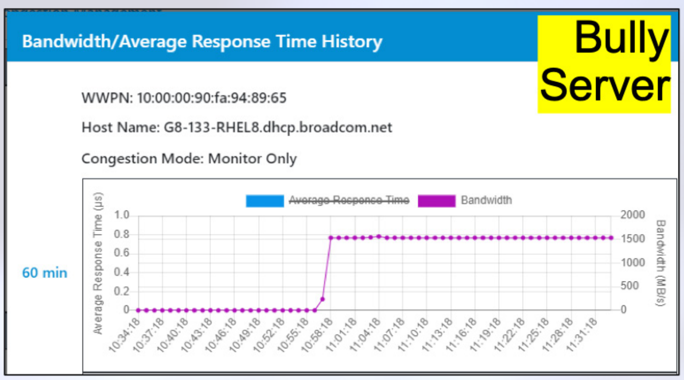


Figure 3: The Bully Workload comes online

Here is the challenge: the hardware resources of the SAN are holding data and waiting for the Bully workload to download them. If you connect more workloads to the network, Victim workloads are created. Victim workloads take a significant performance hit. In Figure 4, the performance degradation of the Victim workload is cut in half.

At this point, many other workloads on the SAN are being negatively impacted. Customer wait times go up as application administrators file trouble tickets. Everybody points to the SAN as being the bad guy.

This scenario applies to any lossless network: IP, InfiniBand, and Fibre Channel. Identifying where the problem is and how to solve it is very difficult, until now.

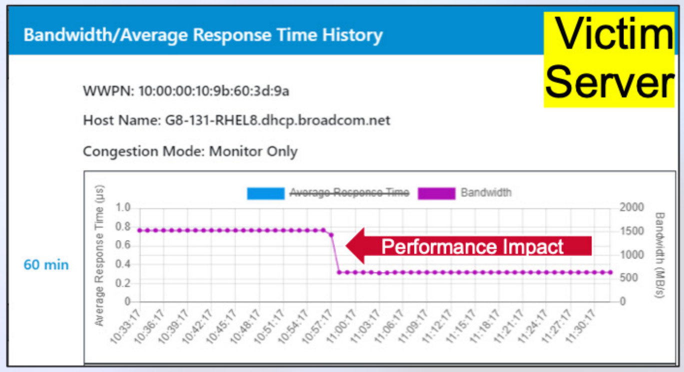
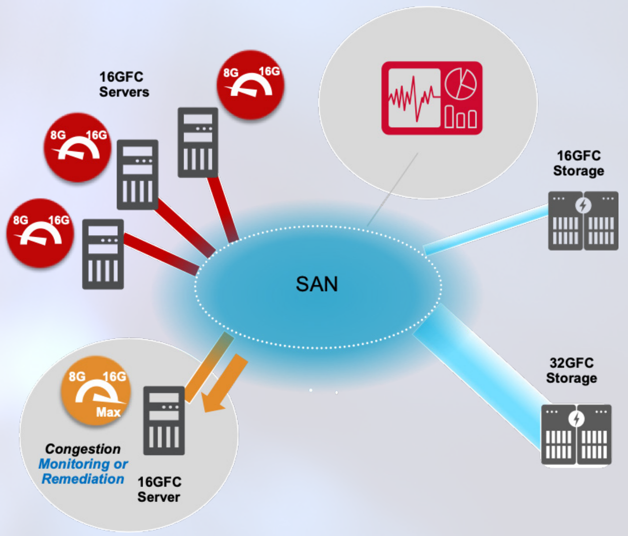


Figure 4: Performance impact on Victim Workloads

Based on the requests from enterprise customers, the Fibre Channel industry developed Fabric Notifications. Fabric Notifications is a new set of standards from INCITS/T11 that enables Fibre Channel Fabric to collaborate with Fibre Channel end-points (HBAs) and Fibre Channel end-points to collaborate with the Fibre Channel Fabric. This article discusses how Fabric Notifications can be used to detect congestion and remediate it. The rule is that the Fabric is the best place to detect congestion, and the end-point is the best place to fix congestion.

Detecting Congestion in the Fabric



Part of the Fabrics Notifications are two new signals. One is called FPIN-CN that stands for Fabric Performance Impact Notification, specifically, Congestion Notification. FPIN-CN is sent to the HBA as an ELS Frame and requires a buffer credit. The second is congestion signaling that is part of the 64GFC Gen 7 Fibre Channel standard which provides the same information, but at lower level in the networking stack and doesn't require a buffer credit. Both signals tell the HBA that its server is causing congestion.

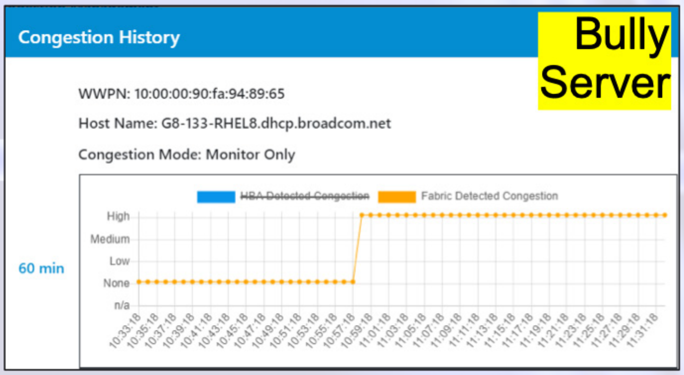


Figure 5: FPIN-CNs received by the HBA

Once the switch detects congestion it sends FPIN-CNs to the HBA. The Fabric now collaborates with the endpoint, a new networking technique first deployed in Fibre Channel networks.

Fabric Notifications is a new technology, and as with any technology, the administrator needs to review data from the Fabric and the HBA. They can set the policy to monitor only and review performance over time. They can turn on congestion management for low priority workloads, like noisy neighbors, and monitor high priority workloads.

Monitoring and Remediating Congestion at the HBA

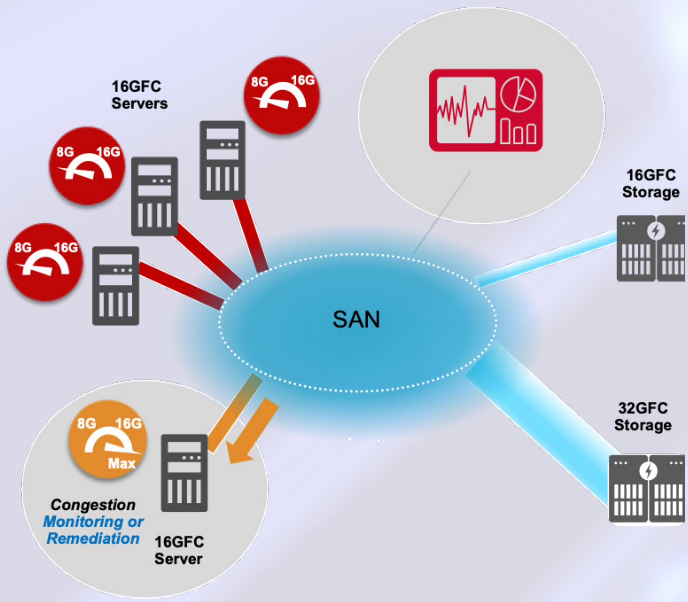


Figure 6: Fabric Notifications at work

Monitoring Congestion

When deploying a new technology, it is essential to be able to see the before and after of a solution like congestion management. So, the first step is to monitor congestion, identify impacted workloads, and decide how to proceed. For a tier 1 workload, you may decide not to touch it and just monitor it over time. Or you may decide to turn on congestion management and measure the effects of this new feature. For low priority workloads, you may turn on congestion management by default and monitor those workloads for performance issues. Fabric Notifications give administrators a lot of flexibility around detecting and remediating congestion.

Remediating Congestion

If you decide to turn on congestion management for the Bully workloads, congestion management can restore performance to the Victim workloads. This is done by slowing down the Bully workload at the Fibre Channel HBA.

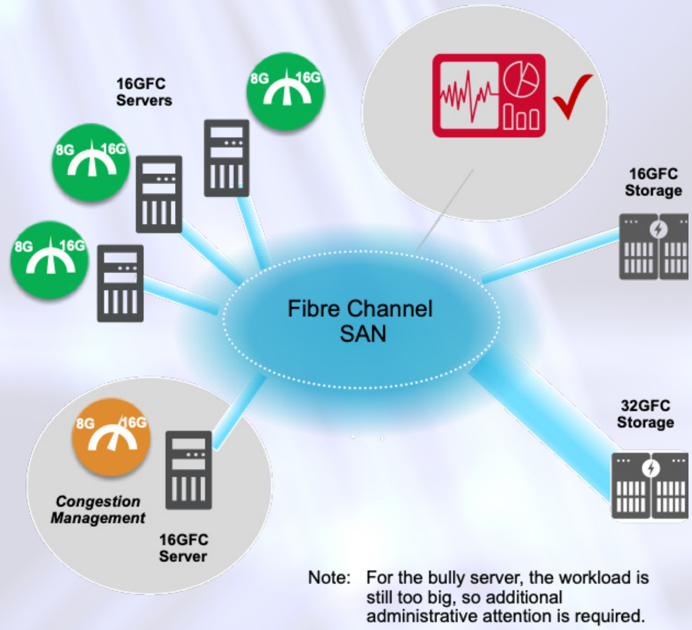


Figure 7: Performance is restored

With congestion management, the demands on the storage network are balanced to maximize resource utilization on the SAN. Victim workloads are no longer impacted by the Bully workload. The Bully workload has slowed down, which may seem like a bad thing, but, if you dig into the data, you'll see that the latency for the Bully workload has dropped significantly, thus, even the Bully server runs better.

The result is balanced resource utilization maximizing performance on the hardware footprint. Note that the oversized workload still needs to be addressed by either moving the workload to a faster machine or by moving VMs off the server.

The Proof is in the Data

Figure 8 and Figure 9 show the complete cycle from congestion, enabling congestion management, and remediating congestion.

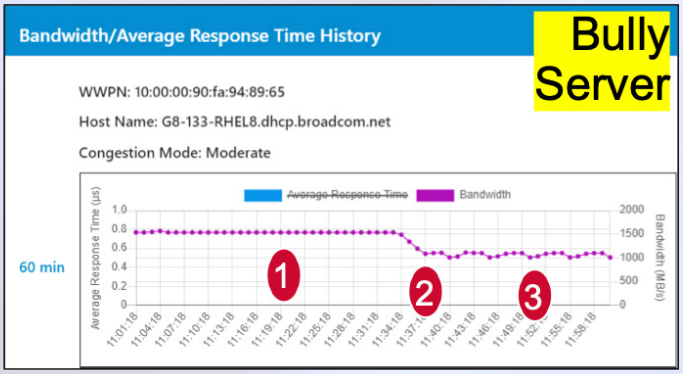


Figure 8: Bully workload before and after congestion management is enabled

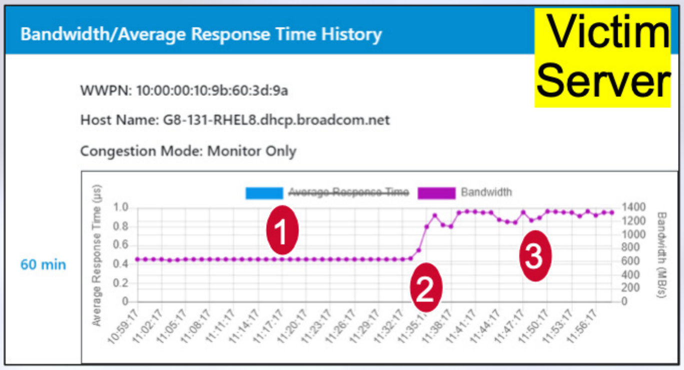


Figure 9: Victim workload before and after congestion management is enabled

The different points on the graphs are:

1. The Bully workload is causing congestion and the Victim's performance is cut in half.
2. Fabric Notifications is turned on in the Fibre Channel Fabric Congestion management is enabled on the Fibre Channel HBA.
3. Congestion management reduces the throughput on the Bully workload by about 30%. Performance is restored for the Victim workloads on the SAN. The throughput on the Victim flow is doubled and near line-rate.

Note that the Bully workload, at point 3 on the graph, bumps up and down a little bit. This is because this solution is adaptive and can address transient and steady state congestion events.

Of course, this process doesn't need to be manual. It can be integrated with enterprise management tools and automated making it easy to deploy and manage.

Summary

Today, cogestion is becoming more common. Driven by All Flash Arrays, the velocity of data has increased, and congestion creates performance problems around the data center.

In a lossless network, that data fills up the network and consumes all of the network's resources, impacting other workloads and causing performance problems across the data center.

Fabric Notifications is a powerful set of new tools to solve data center networking issues for Fibre Channel storage networks. Congestion management directly address a key customer point of pain in their storage networks – the performance degradation due to congestion.

Availability:

- This solution works transparently with any Fibre Channel storage array.
- This solution is supported in VMware, RedHat, SUSE, and Windows today.
- This solution is available on servers from many server vendors including HPE, Lenovo, Dell, and others.

For more information on congestion:

gblogs.cisco.com/in/slow-drain-in-fibre-channel-better-solutions-in-sight

fibrechannel.org/wp-content/uploads/2018/02/FCIA_Fibre_Channel_Performance_Congestion_Slowdrain_and_Over_Utilization_Final.pdf

The Need for Fibre Channel Standards

By Barry Maskas, FCIA member and Senior Technical Staff, HPE

In the world of information technology, companies investing in Fibre Channel (FC) SANs must ensure that they use products and product components that work interchangeably with other products from other companies. Having multiple suppliers is often considered essential for business continuity purposes.

In other words, there is a business-critical demand for standards that document requirements, specifications, guidelines or characteristics that can be used consistently to ensure that materials, products, processes and services are fit for their purpose.

International Standards Organization (ISO) standards are developed according to strict rules to ensure that they are transparent and fair. It can take time to develop consensus among the materially interested parties and for the resulting agreements to go through the public review process in the ISO member countries. However, the resulting standards underpin the deployment of FC SAN technologies.

For some users of standards, particularly those working in fast-changing information technology sectors, it may be more important to agree on a technical specification and publish it quickly, before going through the various checks and balances needed to win the status of a full international standard. Therefore, to meet such needs, ISO has developed a new range of deliverables defined by four different categories of specifications.

These four categories allow for publication of a specification at intermediate stages (with associated adoption risks of development) before full consensus is reached:

1. Publicly Available Specification (PAS)
2. Technical Specification (TS)
3. Technical Report (TR)
4. International Workshop Agreement (IWA)

It must also be noted that not all technology standards we talk about are issued by ISO or one of its Technical Committees. There are a few different standards settings organizations (SSOs), which may have differing structures and rules with implied levels of adoption risk.

The Fibre Channel Industry Association (FCIA) supports, participates, and promotes FC T11 standard developments because full consensus standardization is an integral part of FC technologies. FC standards help companies realize many benefits because of the provided measurable conformance criteria. Standardization improves efficiency in design, development, and material acquisition. Products built using FC standards facilitate interchangeability and compatibility. This results in FC SAN technologies being the right choice for business critical demands and provides increased business continuity, reduced inventory, reduced number of supply sources, increased worker productivity and economies of scale operations.

Because of FC standards, FC SAN designers can be confident that they are designing with products that embody recognized and measurable practices. Standards allow technology to work seamlessly and establish trust so that technical risks are minimized by using components that are proven to work for the intended application. Full consensus-based standardized FC SAN technologies are mature, well known, and provide qualified solutions for high-speed storage communications. Additionally, consumers (and producers!) of FC standards have the added knowledge that a breach of these standards can mean answering to the representative standards bodies.

In summary, FC standards:

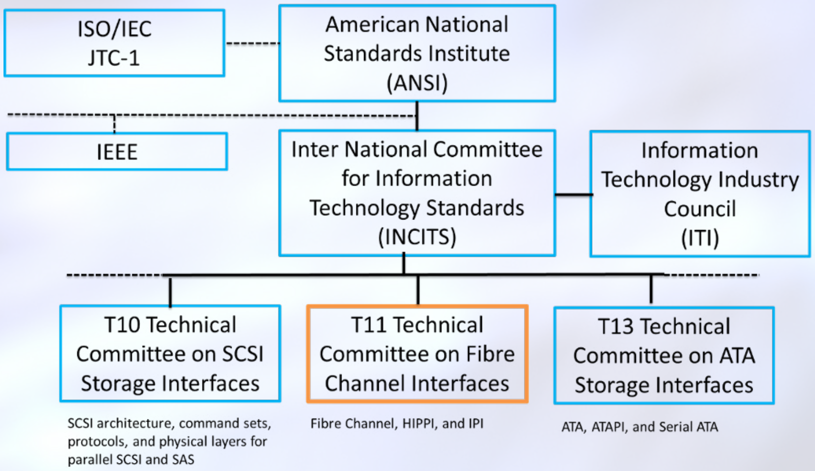
- Provide a common language to measure and evaluate performance,
- Make interoperability of components made by different companies possible, and
- Protect users by ensuring predictable behaviors, reliability, and performance.

What is INCITS?

INCITS -- the InterNational Committee for Information Technology Standards – colloquially pronounced as "Insights" -- is the central U.S. forum dedicated to creating technology standards for the next generation of innovation. INCITS members combine their expertise to create the building blocks for globally transformative technologies. The T11 Technical Committee is governed by the INCITS standards development platforms.

INCITS T11 is the parent committee of Task Groups T11.2 and T11.3. INCITS T11 coordinates the work of the Task Groups and retains overall responsibility for work area. INCITS T11 held its first meeting in February, 1994 and continues to hold FC standards development meetings six times a year.

INCITS T11 Technical Committee relies on the expertise of member engineers, entrepreneurs, developers, and other top-notch professionals to create consensus-driven, market-relevant FC standards that are at the heart of the FC SAN industry. Participation in the T11 Committee is open to anyone materially and directly affected by the work of these Task Groups. The diagram below provides context for how INCITS T11 is organized under the direction of the ISO American National Standards (ANS) process and governance.



The physical variants Task Group T11.2 is the Task Group within the T11 Technical Committee responsible for all projects and parts of projects dealing with physical variants i.e. media, connector, transmitter, and receiver requirements. T11.2 held its first meeting on August 5, 1997 and continues to develop FC technologies today. It recently standardized FC-PI-7 which includes 64GFC single link and parallel links 256GFC, and is developing the next generation of transport speedup.

T11.3 is the Task Group within the T11 Technical Committee responsible for all FC projects which define Fibre Channel Interconnection Schemes. T11.3 held its first meeting on April 23, 1998. The primary focus of T11.3 activities has been directed towards the FC family of transport standards. It should be noted that included in the FC family are mappings which allow protocols from the IP, HIPPI, and NVM Express™ (NVMe™) standards families to be transported across FC.

The three INCITS committees, T10, T11, and T13 are aligned with focus on T10 and SCSI, T11 and FC, and T13 and ATA. These three committees support each other and work jointly to embrace new technologies and advance current technologies. For example, T10 and T11 have actively contributed to, and added support for the NVMe specifications.

FCIA is a mutual benefit, non-profit international organization of manufacturers, system integrators, developers, vendors, industry professionals and end users and provides market direction to the INCITS T11 Task Groups and committees. FCIA members contribute to the marketing of FC solutions.

As an example of the INCITS T11 standardization process, in 2018 there was a need for higher-speed FC networks, faster uplinks, and faster Inter-Switch links. A proposal to do just this, called FC-PI-7, was developed through the work of the INCITS technical committees, T11, T11.2, and T11.3 with market guidance from the FCIA.

Two of the most important aspects of FC are backwards compatibility and "plug and play" to utilize existing infrastructure with new speeds. In FC terms, these are "must-haves," and so they were defined in the FC-PI-7 scope of work along with the requirement to use existing cable assemblies and realize reach goals of 100 meters. The technical requirements of backward compatibility with 32GFC and 16GFC versions of the FC-PI-7 standard were addressed by experts in specific disciplines.

The need for FC to operate over long and short distances is also critical, so the scope of work included a 10KM optical variant and an electrical variant for backplane and adapter applications. The FC-PI-7 standards for 64GFC enabled twice the throughput of 32GFC. Now, of course, you can't just have long links – you also need to have high-quality links as well.

To that end, then, FC-PI-7 also included a corrected bit-error-rate (BER) target of 1e-15 and utilized advanced bit error recovery achieved through use of forward error correction (FEC). FEC is a method of obtaining error control in data transmission in which the source (transmitter) sends redundant data and the destination (receiver) recognizes only the portion of the data that contains no apparent errors.

Once the FC-PI-7 standard writing and reviewing was complete, inclusive of a public review, the standard was forwarded and then approved by the INCITS Executive Board for ANSI/ISO publication. The Executive Board is the INCITS consensus body and has the ultimate responsibility for standards developed and approved by INCITS. When the process was complete, this 5th revision of the FC physical layer standard was published, having evolved to 6400MB/s per port from 1/2/4GFC to 64GFC, or 100MB/s.

A second example is the update and enhancement of an INCITS standard which was ratified in 2017, FC-NVMe. The FC-NVMe INCITS standard defines the mapping of NVMe™ over Fabrics (NVMe-oF™) to the FC interface. FC-NVMe-2 is also a ratified standard, having been completed by the T11-3 Technical Committee and approved by the INCITS Executive Board.

Published INCITS T11 FC standards are maintained on a periodic review cycle, and may be amended. This is a major benefit to those who have materially invested in FC SAN standards and technologies and plan to continue those investments with standards-based technology refreshes.

For more information on how FC standards are made, read my blog series here where I discuss these topics in more detail.

Fibre Channel and the Autonomous SAN

By Howard Johnson, FCIA Member, Technology Architect, Broadcom

Automation, machine learning, DevOps and digital transformation are all the buzz of the modern IT infrastructure. Heralding advancements in capabilities and appearing seemingly everywhere with the promise of simplifying network operations; but what are these advancements really all about? To get an idea of "how stuff works," let's take a look at a new technology that helps these emerging technologies integrate into Fibre Channel.

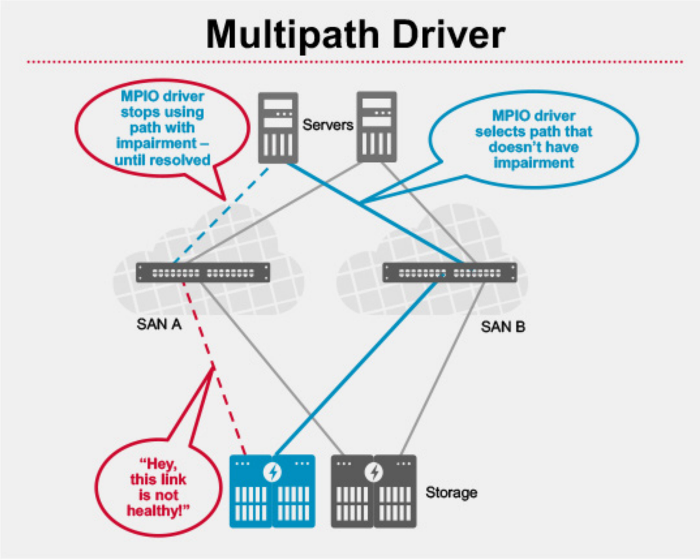
Storage networks are the mainstay of major industries such as finance, healthcare, manufacturing, insurance, government, and many more. As such, Fibre Channel Storage Area Networks (SANs) have stringent reliability and availability requirements that distinguish Fibre Channel solutions from alternative technologies. The impact of an outage in mission critical industries is calculated to be about $100 every second. Imagine going "out-to-lunch" for an hour and returning to a parking meter that read over $300,000 – I'd say it's time to skip lunch!

With SANs at the core of every major business segment, the INCITS T11 standards body that creates the Fibre Channel industry standards, is continually pursuing improvements in reliability and availability. Supporting this focus, Fibre Channel architecture teams have embarked on a method of enhancing the ability of the SAN infrastructure to heal itself, such as identifying when a simple oversubscription condition exists (for example, a 32GFC storage device overdriving a 16GFC server). The first phase of this capability is embodied in a new element of the architecture known as Fabric Notifications.

What's the Buzz?

Fabric Notifications effectively creates a mechanism to surface events occurring in the fabric to the end devices that need the information to make resiliency-based decisions. For example, today when a multi-path solution determines that an IO has been impacted, it cannot tell if the impact was due to a logic error or a physical error. Logic errors are often recovered through retry or logical reset operations that are relatively non-disruptive to the system; however, physical errors often require the administrator to intervene to complete the repair. When the physical error is persistently intermittent, the task of the administrator becomes especially difficult to perform.

With Fabric Notifications, the fabric, or end device, detects the occurrence of the intermittent physical issue. It monitors the issue to see if persists, and if it does, it generates a message to be sent to the devices that are affected by the event. Instantly, the multi-path solution knows the location and nature of the physical error and can "route around" it by utilizing good, alternate paths. A clear benefit of this feature is that the administrator is not involved in the identification, isolation, and recovery of the error.



Tell Me More, Tell Me More

Let's break down the elements of Fabric Notifications to see how it actually works. There are two main functions provided with this architecture – signals and notifications. The signals are physical layer operations that allow the two ends of a link to indicate when transmission resources are being consumed to the point of causing problems in the fabric or internal to the device. The notifications are logical layer operations that allow the fabric or devices to notify each other of events they have detected and how those events might affect their logically connected peers.

These mechanisms are fully controlled by the end devices through an exchange of capabilities and the registration of operations. This approach ensures that only the devices that are capable of managing and responding to these new messages ever receive them. Hence, a device that wants to enable the Congestion Signal mechanism, exchanges that capability with the fabric port and the fabric port does the same. Upon successful completion of the exchange of capabilities, the two sides enable their logic to generate the signals based on their specific implementations. Likewise, a device that wants to enable the Notification mechanism registers to receive the notification Extended Link Service (ELS) and includes a list of the specific sub-functions they want to receive. In this way, the end device completely manages its participation level as well as controls the degree of participation.

When the exchange and registration process is completed , the devices are ready to receive signals and notifications from the fabric that will help them effectively evaluate and isolate errors. The devices can now make intelligent decisions for automated corrective actions (that is, they can engage in machine learning techniques that improve reliability, availability, and resiliency). Essentially, Fabric Notifications adds more information into the system and distributes that information to the participating devices. The devices can then make more informed decisions and respond intelligently to the events in the system. Furthermore, this mechanism allows solutions to manage their level of participation and only react to the events they care about – and only to the degree to which it helps, without impacting their operations.

What's a Device to Do?

The beauty of the Fibre Channel Fabric Notifications architecture is its simplicity and control. Since the participating devices control the level and degree of participation, solutions can evolve over time and adjust participation based on experience in the enterprise.

Basic solutions may choose to simply register to receive all notifications and just log them. For a SAN administrator, this simple action reduces problem determination and isolation considerably because the fabric and device logs describe exactly where the event occurred in the fabric and which devices are affected.

A slightly more advanced solution may choose to handle congestion conditions in the fabric and register for Peer Congestion Notifications. The device may then take the simple action of speed matching for flows to a device that exhibits congested behavior (for example, our 32GFC storage to 16GFC server connection automatically adjusts the flow to be 16GFC to 16GFC).

In our multi-path example, a solution may register for Link Integrity Notifications to allow the device to avoid paths that are automatically identified by the fabric to be intermittently impaired without requiring the SAN administrator to intervene.

Fibre Channel Autonomous SANs

The position of the Fibre Channel switch in the data path allows it to have visibility into the surrounding components and gather intelligence not only on the infrastructure, but the entire storage network, including the attached devices. This intelligence is exchanged amongst the switches of a fabric to create the vision of an autonomous SAN that is comprised of self-learning, self-optimizing, and self-healing activities. Fabric Notifications enable these characteristics and provide the foundation for delivering autonomous infrastructures. These capabilities eliminate countless wasted cycles of manual analysis, which becomes very challenging with the scale of modern data centers. Furthermore, Fabric Notifications reduce the period of time that applications perform at sub-par levels resulting in lost revenue or degraded service.

The Fibre Channel industry's Fabric Notifications architecture is the technology that supports Autonomous SAN technology with the capability to automatically identify and resolve issues. The fabric identifies data traffic congestion and facilitates automatic failover or traffic adjustment by notifying the devices. The devices in turn, take automatic corrective actions to mitigate the impact of the congestion or persistent, intermittent failures.

Imagine an environment where the network mitigates issues automatically and flags the troubled component for resolution later. It's technologies like Fibre Channel Fabric Notifications that enable infrastructures to manage themselves and brings us closer to truly autonomous SANs.

---

[i] *The capabilities exchange occurs when the N_Port sends the Exchange Diagnostic Capabilities (EDC) ELS command with a Congestion Signal Capabilities (CSD) descriptor in the payload to enable Congestion Signal processing with the F_Port to which it is attached. The registration process occurs when the N_Port sends the Register Diagnostic Functions (RDF) ELS with the Fabric Performance Impact Notification (FPIN) descriptor that contains a list of descriptor tags corresponding to the specific FPIN sub-functions it wants to receive. The sub-functions include descriptors that enable Link Integrity Notification (FPIN-LI), Delivery Notification (FPIN-DN), Peer Congestion Notification (FPIN-PN), and Congestion Notification (FPIN-CN) sub-functions.*


FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION
www.fibrechannel.org