Fabrics for Storage Area Networks

A Fresh Perspective

A G2M Research White Paper

White Paper Key Points

1. The choice of a SAN fabric and protocol has a significant impact on application performance, availability, and cost.
2. A dedicated storage fabric such as Fibre Channel provides the highest throughput and lowest latency/jitter, albeit at a higher CapEx compared to existing Ethernet LANs.
3. When a dedicated SAN isn't an option, Ethernet-based storage protocols offer lower CapEx and OpEx but at the expense of performance.
4. Where it is a close call, deploy higher speeds of the SAN fabric that you are currently using, as it represents the lowest risk and potential to disrupt your operations.

Fabrics for Storage Area Networks: Executive Summary

Amidst the surging demands on modern storage architectures, driven by the massive influx of data from AI, autonomous vehicles, surveillance systems, and the intricate web of regulatory and financial requirements, a fresh perspective is essential. As we navigate this dynamic landscape, it becomes increasingly pertinent to reevaluate strategies that address the trifecta of performance, cost efficiency, and robust support. Stay ahead in this ever-evolving scenario by redefining how you approach the challenges posed by the data deluge.

In the current landscape of Storage Area Networks (SANs), there is a wide array of architectures, fabrics, and protocols available for use. However, choosing the most suitable approach, fabric, and protocol to optimize application performance is far from straightforward, and the tradeoffs involved may not always be apparent. The decision-making process for selecting a storage architecture and related fabric and protocols is typically guided by four key factors:

1. **Scale of the SAN:** Determining the number of nodes that need access to the shared storage pool.
2. **Performance Requirements:** Evaluating throughput, which can be measured in terms of bandwidth or packets per second for a specific packet size, and latency time to complete a command).
3. **Consistency of Performance:** Considering the level of performance consistency required by applications before their availability is impacted.
4. **Network Type:** Deciding whether a dedicated SAN or a converged network is preferable, based on economic factors or other constraints.

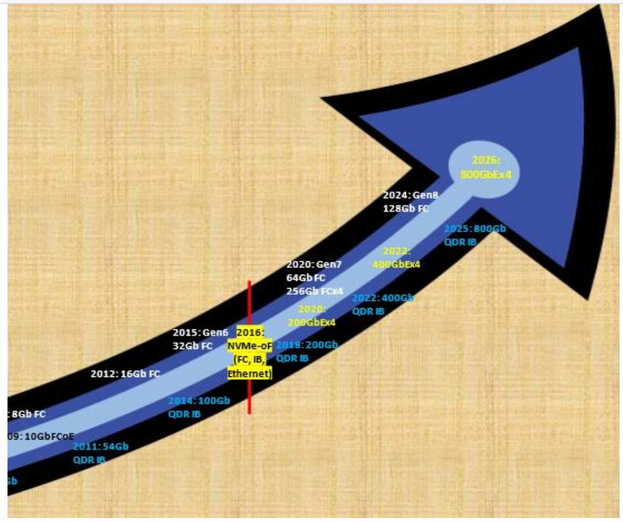
For applications that can afford a dedicated SAN and demand thousands of nodes sharing the same storage namespaces, along with significant throughput per initiator node, low jitter, and low latency, Fibre Channel (FC) has distinct advantages over Ethernet-based fabrics. While converged Ethernet networks (combining SAN and LAN functionality) are more cost-effective and currently offer clock speeds of up to 200Gb/second, they can suffer from network congestion, leading to reduced usable bandwidth and increased jitter. Deploying a dedicated Ethernet-based SAN can mitigate these issues but often comes with higher SAN costs.

Alternatively, for use cases where a dedicated SAN is not feasible or economically viable, Ethernet converged networks utilizing NVMe over Fabric protocols for the SAN generally offer superior performance and lower latency. However, careful attention must be paid to the impact of congestion on SAN performance. Understanding the delineation between these two scenarios and knowing how to address the "gray area" is crucial.

The Evolution of Storage Area Networks (SANs)

In the 1980s, Local Area Networks (LANs) built on Ethernet technology gained popularity in both workspaces and data centers. File servers were introduced to provide shared storage to workstations over the LAN. As organizations increased the number of file servers, the need for centralized high-performance storage systems became evident. These systems were initially connected to servers through LAN, but the growing storage traffic started to impact other LAN operations. To address this, ANSI standardized Fibre Channel (FC) in 1994, a serialized optical networking technology specifically designed for storage systems.
Today, SANs utilize block-level storage protocols to offer remote access to physical or virtual blocks on shared storage devices. Initially based on the Small Computer System Interface (SCSI) specification, several Ethernet-based SAN protocols were proposed as alternatives to FC. Protocols like iSCSI and Fibre Channel over Ethernet (FCoE) aimed to build converged LAN/SAN networks and reduce costs.

A significant advancement in the storage interface arena is the Non-Volatile Memory Express (NVMe) interface. Standardized nearly a decade ago, NVMe offers significantly reduced latency compared to SCSI protocols. Various remote access versions of NVMe, known as NVMe over Fabric (NVMe-oF), have been developed, including NVMe over RDMA (NVMe/RDMA) for Ethernet, NVMe over TCP (NVMe/TCP) for Ethernet, and NVMe over Fibre Channel (FC-NVMe). The first standardized NVMe protocols emerged in 2016, and major storage vendors introduced arrays in mid-late 2017. While initial deployments of NVMe based storage arrays started in late 2018, SANs utilizing NVMe in its various flavors have resulted in its increasing usage in SANs today for a variety of use cases. As the storage landscape continues to evolve, NVMe over fabrics are poised to play a crucial role in delivering lower latency and higher performance, paving the way for more efficient and advanced storage solutions in the years to come.



Considerations When Selecting a Fabric for Storage Area Networks

When choosing a fabric for Storage Area Networks (SANs), several critical factors come into play, demanding careful evaluation to ensure optimal performance and efficiency. These factors include:

1. **Scalability Requirements:** Assess the scale of the SAN and determine the number of nodes necessitating access to the shared storage pool. Understanding the extent of expansion potential is crucial to accommodate future growth.
2. **Storage Throughput Needs for Each Application Instance:** Analyze the storage performance needs of each application instance. This encompasses throughput requirements (measured in GB/second) and packet sizes (large, typically 4KB per packet or more, or small, usually less than 1KB per packet). Additionally, consider measuring throughput in packets per second (PPS) by dividing GB/second by the packet size.
3. **Latency Needs for the Application:** Evaluate the desired latency at the application level, which refers to the time delay between issuing a read/write command by the application and its completion. Latency is typically measured in microseconds (us).
4. **Throughput and Latency Jitter Requirements:** Examine the requirements for throughput jitter, which represents the variance in throughput over time. Similarly, assess latency jitter, measuring the variation in latency over time.

Furthermore, other crucial criteria to consider include delivery reliability, cost (both Capital Expenditure and Operational Expenditure), and adoption rates.

Comparing the Various Fabric and Protocol Choices

When evaluating the different storage protocols, it's essential to recognize that the comparison largely hinges on the specific characteristics of the network components involved. However, certain general observations can be made. In dedicated Storage Area Networks (SANs), such as Fibre Channel (FC) networks, performance, latency, and jitter tend to be superior to Ethernet-based networks. Nevertheless, it is vital to acknowledge that there are scenarios where Ethernet SANs can be a viable alternative. For instance, 100GbE NVMe-oF networks dedicated solely to storage purposes can deliver higher throughput. However, their scalability beyond the rack level currently presents challenges.

Table 1 below provides a comprehensive comparison of these various protocols based on seven key criteria. Subsequent sections will delve into each set of fabrics and protocols, examining their primary use cases and the driving factors behind their success or limitations for those specific scenarios.

In this study, we will exclude protocols that have not attained widespread adoption or are not expected to do so. Examples of such protocols include NVMe/IB primarily for high performance computing and FCoE which never gained broad acceptance.

For more detailed insights into the fabric and protocol choices, refer to Table 1 and the following sections.

Fibre Channel: History, Evolution, and Adoption

Fibre Channel (FC) has remained the preferred fabric for Storage Area Networks (SANs) since its inception over two decades ago. More than 100 million FC ports have been shipped since 2001, and currently, an estimated 46 million FC ports are operational, with the majority using Gen 5 (16Gb FC) or Gen 6 (32Gb FC) technology (Gen 7 is in it's early stages of the product life cycle). While FC unit sales have remained relatively even or a slight decline since the early 2010s, we've seen a recent uptick in port shipments in the last 24 months due to data center infrastructure upgrades post covid. Nevertheless, FC remains the dominant fabric choice for leading storage array vendors like Dell/EMC, NetApp, and Pure Storage. FC switches are offered by Broadcom (Broadcom) and Cisco, while FC Host Bus Adapters (HBAs) are available from ATTO, Broadcom (Emulex), Marvell (QLogic), and other manufacturers. The primary protocol supported on FC fabrics is Fibre Channel Protocol (FCP), a SCSI-based protocol.

Presently, FC's maximum network speed is 64 Gb/second for a single lane, with Inter-Switch Links (ISLs) utilizing four lanes capable of reaching 256Gb/second. In the upcoming years, FC speeds are projected to double to 128Gb/second on a single lane and up to 512Gb/second on quad-lane ISLs. The majority of FC SANs deployed today use 32/16GbFC.

Use Cases and Considerations:

As a dedicated storage network, FC SANs excel in most real-world use cases compared to Ethernet-based alternatives like iSCSI.

Fabric/ Protocol	Dedicated SAN?	Scalability	Ease of Expansion	Through -put	Latency	Jitter	Delivery Reliability	Cost	Adoption
FC/FCP	Yes	5	5	4	4	5	5	2	5 ↔
FC/NVMe-oFC	Yes	5	5	4	5	5	5	2	1 ↑
Ethernet/iSCSI	No	3	2	3	2	2	2	5	3 ↓
Ethernet/NVMe-oF	No	4	2-3	5	3-4	3	3	5	1 ↑

Table 1: Rating of SAN Protocols Across Various Criteria (*Ratings: 1=worst to 5=best*)

While FC SANs do come with higher CapEx, they offer significant benefits such as scalability and immunity to network congestion, thanks to a credit-based flow control mechanism. This enables the creation of FC SANs with thousands of nodes without compromising throughput, latency, or jitter. FC SANs also support multi-pathing, enhancing performance and reliability. These advantages simplify performance optimization and storage expansion, tasks that are more challenging in converged Ethernet networks due to the constant need for performance tuning as workloads change and storage is added to the SAN.

The capabilities of FC have led to widespread adoption in Fortune 500 (F500) enterprises and high-performance workflows. Use cases for FC SANs encompass:

- Large database clusters like Enterprise Resource Planning (ERP) and billing systems.
- Banking financial management systems.
- Airline and travel booking systems.
- Large retail point of sale systems.
- Persistent storage for significant in-memory database solutions like SAP HANA.
- Large media and entertainment workflow clusters, including raw content post-production and video editing systems.
- Oil and Gas analytics clusters.

Ethernet/iSCSI: History, Evolution, and Adoption

iSCSI was once considered a potential competitor to, or even a replacement for, FC in SANs. However, it never achieved the same level of adoption and deployment as FC due to being limited to the throughput until the emergence of 10/100 Gigabit Ethernet. Nevertheless, the majority of Ethernet-based SANs today rely on iSCSI, benefitting from its cost-effectiveness as iSCSI utilizes standard Ethernet network interface cards (NICs) and switches. The widespread support from major storage array manufacturers, thanks to its cost advantage and universal interoperability, has contributed to iSCSI's significant installed base.

Despite these advantages, iSCSI does have drawbacks, particularly concerning latency. Relative to FC, iSCSI's more complex protocol stack results in increased latency, and Ethernet's susceptibility to congestion further compounds the issue as seen in **Figure 1**

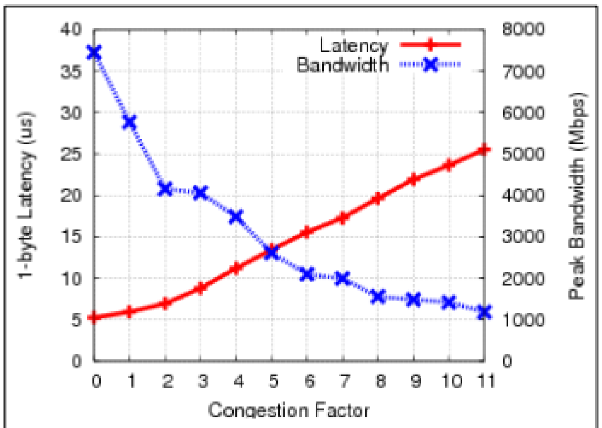


Figure 1: Ethernet Congestion vs Latency

According to a recent report from Crehan Research Inc. shipments of 25 Gigabit Ethernet (25GbE) adapters and controllers (NICs, surpassed 10GbE shipments during the first half of calendar-year 2022.

The high-end network speed for Ethernet is 100/200Gb/second in a single lane (100/200GbE), with quad-lane 100Gb Ethernet reaching 400Gb/second (400GbE). However, 100/200GbE is currently limited to applications where its high cost per port is justified by specific performance requirements. Managing network congestion in 100/200GbE environments adds complexity to deployment, management, and optimization, especially in SAN configurations (see **Figure 2**). It should also be noted the 400GbE adapters are available but represent a very small fraction of the market (cloud and telecom) due to infrastructure cost.

Use Cases and Considerations:

iSCSI has found its place in use cases where low cost for shared storage is the primary consideration and where performance and data delivery reliability are not critical factors.

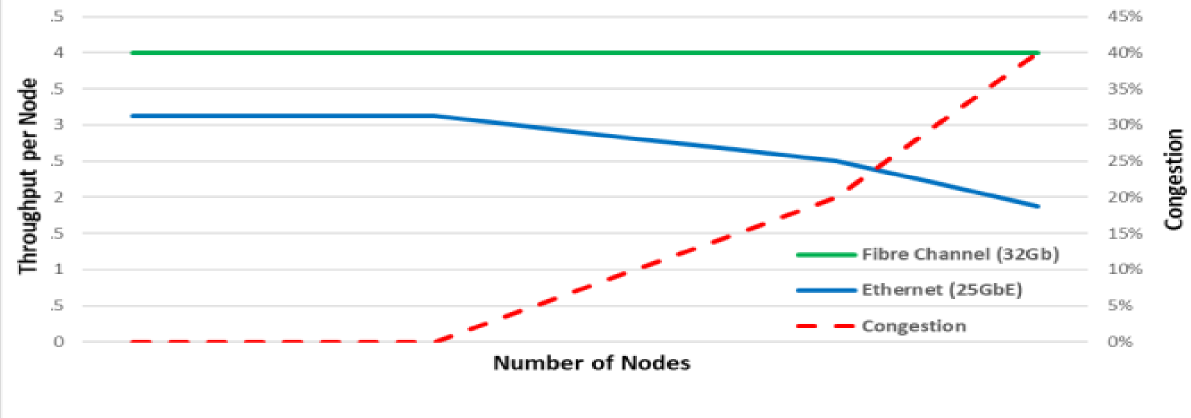


Figure 2: Per-Node Performance With Congestion

Historically, this includes small to medium businesses (SMBs) and low-performance workflows and applications that require shared storage, such as file sharing, email, or web servers. Nevertheless, iSCSI's notable challenge lies in its high latency, with typical throughputs reaching only around 40%-50% of line rate. Throughput jitter and latency jitter are also significantly higher in iSCSI networks compared to other SANs, particularly for more complex network topologies. The potential for out-of-order packet delivery and packet loss contributes to these challenges.

While Ethernet networks theoretically offer unlimited scalability, iSCSI over Ethernet networks start experiencing congestion and "noisy neighbor" issues well before reaching even half of their theoretical throughput. This negatively impacts actual network throughput by causing retries that reduce network capacity, even in 100/200GbE networks, and hinders network reliability for storage applications. Ethernet's susceptibility to congestion also adversely affects its throughput.

NVMe over Fabrics: Protocols and Considerations

NVMe over Fabrics (NVMe-oF) is an extension of the NVMe protocol that allows flash devices to be accessed over a network, such as Ethernet or Fibre Channel. This enables faster and more efficient connectivity between storage and servers, as well as lower CPU utilization.

NVMe over Fabrics (NVMe-oF) introduces a variety of protocols, each designed to enhance performance and minimize latency compared to traditional SCSI-based protocols in dedicated Storage Area Networks (SANs). Among these protocols are NVMe over RDMA (NVMe/RDMA), NVMe over TCP (NVMe/TCP), and NVMe over Fibre Channel (FC-NVMe).

NVMe/RDMA utilizes a flow control methodology similar to FC's credit-based flow control, ensuring in-order delivery of packets without loss. By eliminating the SCSI protocol, NVMe/RDMA achieves reduced latency while capitalizing on the advantages inherent in a dedicated SAN.

FC-NVMe operates on standard FC networks (existing infrastructure including switches, HBAs, cabling, etc.) and enjoys all the benefits of a dedicated SAN while benefiting from the elimination of the SCSI protocol, further reducing latency.

On the other hand, NVMe/TCP adopts a different approach by utilizing standard TCP flow control, allowing for interoperability with L2 Ethernet switches and management applications. Though TCP-based NVMe adapters tend to exhibit approximately 10% higher latency than NVMe/RDMA in laboratory conditions, the difference is minimal under realistic conditions.

Despite the advantages offered by NVMe-oF protocols, the primary challenge lies in the relative immaturity of the larger storage ecosystem, particularly the availability of storage arrays that fully support NVMe-oF. This aspect demands consideration when implementing NVMe-oF solutions.

Additionally, it is crucial to be mindful of network congestion issues, especially in NVMe/RDMA networks. Under congestion conditions in lossless networks like DCB networks, initiators are expected to limit the rate of packet injection into the network, essentially "capping" the performance of application instances in NVMe/RDMA networks. This effect can influence both network scale and throughput/latency jitter.

As NVMe-oF protocols continue to evolve and gain traction, addressing these considerations will be essential in harnessing their full potential for enhanced storage performance in modern data centers. Most SANs and arrays support NVMe-oF protocols as an option, simplifying later adoption if desired.

Decision-Making Outside the Clear FC/Ethernet Use Cases

When your application doesn't neatly align with either the Fibre Channel (FC) or Ethernet use cases, determining the appropriate Storage Area Network (SAN) fabric becomes a critical decision. If a dedicated SAN is an option, **Table 2** offers valuable decision criteria to consider when choosing a SAN fabric and protocol for both greenfield deployments and enhancements to existing SANs.

For new SAN deployments, it is prudent to begin with what you are familiar with. For instance, deploying an FC SAN in an organization that has no prior experience with FC could be a high-risk undertaking. The same caution applies to large-scale NVMe/RDMA deployments. However, this doesn't mean that such decisions are inherently unsuitable; rather, they should only be pursued with the necessary design and deployment support, such as professional services, from your storage/server vendor.

Considering factors like application requirements, organization expertise, and potential risks will guide the decision-making process. When in doubt, consulting with experienced professionals and leveraging vendor expertise can be invaluable in ensuring the success of your SAN implementation.

Network Type		HDD or SAS/SATA Flash Storage Systems	NVM over Fabrics Storage Systems
Deploying on an Existing Network	1/10GbE	Ethernet/iSCSI	Ethernet/NVMe/TCP
	16/32GbFC	FC/FCP	FC/FC-NVMe
Deploying on a New Network	Large Workgroup (>10)	32/64GbFC/FCP	32/64GbFC/FC-NVMe
	Small Workgroup (< 6)	10/25GbE iSCSI	25GbE/NVMe/RDMA
	Single Rack	Ethernet/iSCSI	NVMe/RDMA
	Multiple Racks, Replacing 10Gb iSCSI	25/100GbE/iSCSI	25/100GbE/NVMe/TCP
	Multiple Racks Replacing 8/16Gb FC	32/64GbFC/FCP	32/64Gb FC-NVMe 25/100GbE/NVMe/TCP

Table 2: Decision-Making Criteria for Selecting a SAN

About G2M Research (www.g2minc.com)

G2M Research (a re-grate-it, inc. brand) is a highly specialized industry research and marketing firm that enables technology companies to systematically develop relevant products and marketing content, drive revenue growth, and optimize media and analyst communications. G2M Research provides targeted industry reviews and reports that provide market sizing, ecosystem mapping and market analysis, delivering a comprehensive view of a given technology. G2M Communications helps you build market valuation and brand preference through competitive analysis, thought leadership strategies, technical and sales enablement content development, corporate communications programs, targeted multi-channel campaigns and creative direction and execution.

About ATTO Technology (www.atto.com)

For 35 years ATTO Technology, Inc. has been a global leader across IT and media & entertainment, specializing in network and storage connectivity and infrastructure solutions for the most data-intensive computing environments. ATTO works with customers and partners to deliver end-to-end solutions to better store, manage and deliver data, often as an extension of their design teams. ATTO manufactures host adapters, SmartNICs, storage appliances and controllers, intelligent bridges, Thunderbolt™ adapters, and software. ATTO solutions provide the highest level of storage connectivity performance for Fibre Channel, SAS, SATA, iSCSI, Ethernet, NVMe and Thunderbolt. ATTO is the Power Behind the Storage.

References

1 – FCIA, “Fibre Channel’s Growth Trend”, Mark Jones.

2 – NETWORK Computing, “The Sinking Fibre Channel SAN”, Jim O’Reilly.

3 – NETWORK Computing, “Fibre Channel Really is Dead”, Kevin Deierling.

4 – BusinessWire, “Technavio’s Top Three Trends to Impact the Fibre Channel Adapter Market” .

5 – Dell Oro Group Blog, “Storage Systems Market in Transition”.

6 - SearchStorage, “Essential Guide: How to choose the best hardware for virtualization; iSCSI”, Margaret Rouse