# SAN Basics: The Transition to NVMe

by Mike Kieran, NetApp Technical Marketing Engineer
Michael Peppers, NetApp Technical Marketing Engineer

MAY 2018

NetApp®

# Introduction

Enterprise data storage is being transformed by new technology—again.

This time, the emergent technology is NVMe—short for nonvolatile memory express—a new storage access and transport protocol that delivers the fastest response times yet for business-critical enterprise applications.

Like the arrival of flash storage, NVMe is about to provide a major speed boost for the enterprise data storage systems. But this time, the transformative effect could be greater still, because NVMe isn't just a storage spec; the broader NVMe over Fabrics (NVMe-oF) protocol re-architects the entire data path, from server to storage system.

In this primer, we'll examine how NVMe technology can enable organizations to build storage area networks (SANs) with the performance, reliability, availability, and scalability they need to run their core enterprise applications.

You'll learn how NVMe technology:

- Defines a new high-speed data transfer protocol for solid-state drives (SSDs) to communicate with enterprise or client systems
- Provides higher IOPS and reduced latency from the host software stack, through the fabric to the storage array
- Takes advantage of the ever-increasing speed of the Fibre Channel transport and other networking technologies
- Can provide your organization with world-class performance for SANs running business-critical applications

# New Business Imperatives

It's hard to find a large organization today that isn't going through a "digital transformation"—examining every aspect of their business to see where data can be exploited to accelerate time to market, reduce error rates, and improve the customer experience.

In some cases, organizations are accomplishing this by updating existing applications, such as customer relationship management or enterprise resource planning. In others, they're deploying entirely new workloads, including mobile, video, automated bots, real-time analytics, artificial intelligence, and the Internet of Things.

Across all use cases, the volume of data that's being created, aggregated, shared, and stored is increasing at a rapid rate. More important, the **value** of that data is growing exponentially over time. Today, the world's most valuable resource is no longer oil, but data.

That's why NVMe arrives in the data center at the perfect time:

- It empowers IT organizations to provide their internal customers with real-time response for crucial business workloads.
- It enables organizations to reduce time to value by identifying and exploiting opportunities that materially impact the business.

Today, IT leaders are responsible for upgrading their existing infrastructure, delivering everything as a service, and spinning up new workloads, often with little in the way of new budget. That's why NVMe technology is a crucial technology in the modern data center. It reduces response times for critical applications by reducing latency—the time it takes to access application data on a SAN—from milliseconds to microseconds.

Reducing latency with NVMe technology can also add business value for many use cases, including:

- **Real-time analytics**. With billions of debit and credit card transactions being processed every day, retailers and credit card companies have a huge incentive to prevent fraud before it happens. NVMe SANs provide the technology for the near-instantaneous response necessary to avoid huge potential losses, in-store and online.
- **Software agents**. Intelligent software bots are increasingly providing high-quality customer engagement, technical support, and other interactive services, often faster than a human could. In the future, NVMe technology will enable systems to respond so quickly and accurately that users won't know whether they're talking with a human or a bot.
- **Internet of Things**. NVMe-based storage systems can rapidly ingest massive data sets, and they support huge numbers of simultaneous client connections. That makes them ideal for use both as edge data collection points, and as centralized storage pools that can host the resultant data lakes created by aggregating all the sensor data.

> "NVMe over Fibre Channel delivered 58% higher IOPS and 34% lower latency than SCSI FCP."
>
> —Demartek analyst report, *Performance Benefits of NVMe over Fibre Channel – A New, Parallel, Efficient Protocol*, May 2018

# NVMe and Solid-State Drives

When flash technology arrived in the data center, first as hybrid disk-flash arrays and then as all-flash arrays, it had a huge impact on the performance of enterprise applications. Flash storage has also changed performance expectations for end users. Users with flash-enabled PCs have become accustomed to ever-faster response times and expect the same experience for all of their business applications, even as those applications become more complex and extend from the data center to the cloud.

However, the speed gains possible from SSDs are approaching a limit because of bottlenecks in other parts of the data pathway. Both of today's main networking protocols—Fibre Channel and Ethernet—use the SCSI command set for the storage protocol. SCSI was developed for mechanical media almost 40 years ago, and although it can handle the data flow to and from spinning disks, it can no longer keep up with new flash storage media.

That's because SCSI puts I/O requests into a single queue, containing a maximum of 256 commands. As I/O requests arrive from an application, they have to wait in line while other requests are completed. In the era of spinning disks, these data access delays didn't really matter, because

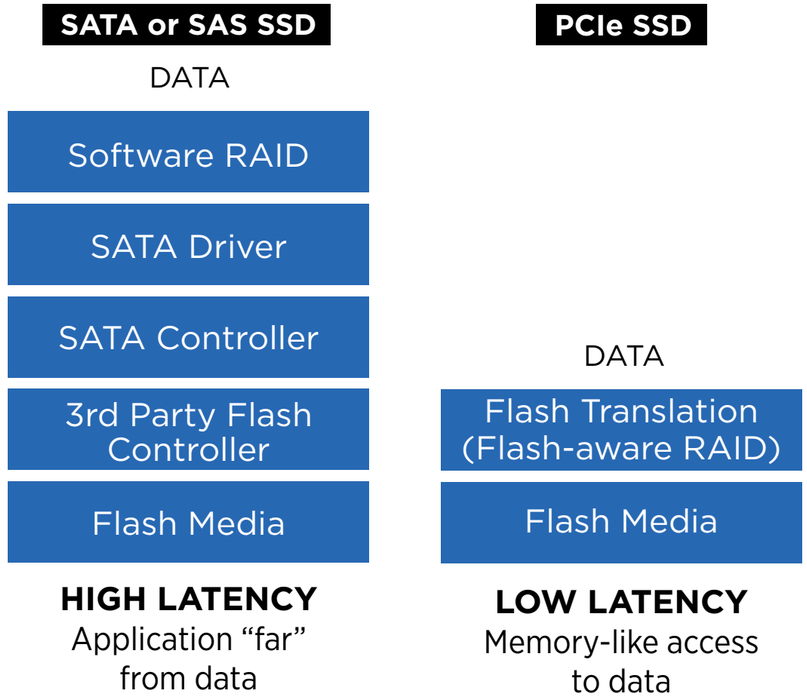


Figure 1: The NVMe software stack for SATA or SAS versus PCIe.

the drive's read-write heads were busy anyway, spinning around seeking tiny magnetic blocks on rapidly rotating platters. Flash and other solid-state media have no moving parts, and thus very low innate latency, so delays in the data path have become the new bottleneck, slowing data access.

Over the past decade, this situation has improved with the emergence of the Peripheral Component Interconnect Express (PCIe) bus, which is much faster than SATA and other legacy protocols. For instance, the PCIe bus supports up to 65,535 queues, each with a queue depth of 65,535 commands, and can connect storage directly to the CPU, providing memory-like access. Plus, PCIe requires a much smaller software stack than legacy protocols (see Figure 1).

Today, PCIe-based devices are used in all kinds of computing environments, which could pose a problem for SSDs, because every vendor's hardware requires the use of proprietary drivers. To eliminate compatibility issues, all the key industry players have agreed on a new modern data access standard for flash and other forms of solid-state memory running on a PCIe bus: the NVMe specification.

NVMe is actually two distinct but interwoven things:

- A protocol and instruction set optimized for solid-state storage devices
- A forward-looking set of open-source architectural standards for nonvolatile memory systems

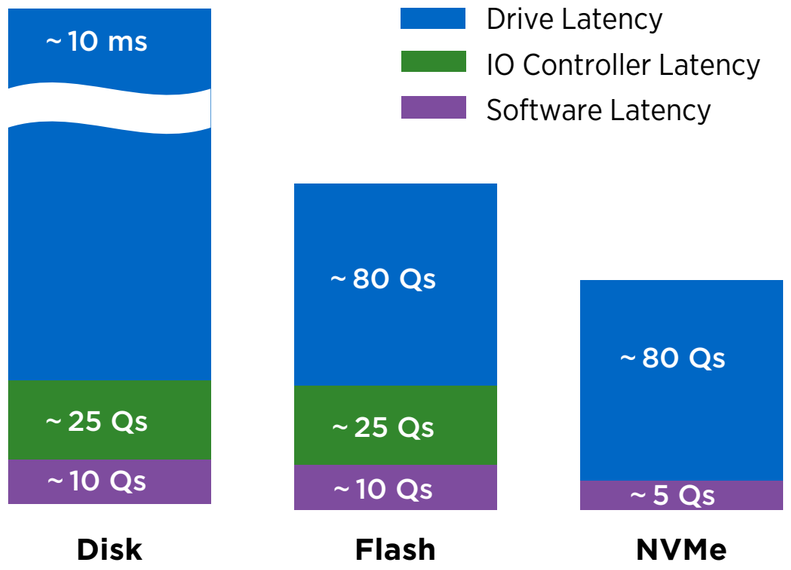


Figure 2: The impact of NVMe on media access latency.

The NVMe specification was designed from scratch to take advantage of nonvolatile memory in all kinds of computing environments, from mobile phones to webscale service providers. And thanks to its increased throughput and lower latency, NVMe is rapidly expanding its presence in the data center (see Figure 2).

Among the companies that have made massive investments in NVMe are NetApp, Broadcom, Intel, Samsung, Micron, Seagate, Cisco, Western Digital, Microsoft, and Toshiba. NetApp now uses NVMe storage across most of its hybrid-flash and all-flash storage systems to significantly improve performance.

In summary, NVMe is driving an architectural shift that makes communication with storage systems massively parallel. The result is greater bandwidth and lower latency connectivity between servers and storage devices.

And it's intended to be future-proof: NVMe supports today's NAND flash and 3D XPoint technologies and will work with currently in-development and as-yet-undiscovered persistent memory technologies as they emerge.

## Wait, There's More!

Many other factors enable NVMe to provide unparalleled performance in the data center, including:

- Interrupt handling
- Internal locking needed to serialize I/O requests
- Command streamlining
- Reduced context switches
- Lockless design
- Polling mode

# NVMe over Fabrics

So far, we've focused on how the NVMe protocol speeds up data storage. But NVMe is also a data transfer protocol—and a key component in the latest generation of high-speed data fabrics.

The NVMe specification is continually evolving; with the arrival of NVMe over Fabrics (NVMe-oF), NVMe's performance and latency benefits now extend across network fabrics such as Ethernet, Fibre Channel, and InfiniBand. NVMe-oF adds support for a front-end storage interface, greater distances within a data center over which NVMe devices and subsystems can communicate, and the ability to scale out to large numbers of NVMe devices.

These are crucial enhancements, as networks themselves continue to get faster and faster. For example, Fibre Channel has accelerated, first to Gen 5 at 16Gbps (gigabits per second) and now to Gen 6 (32Gbps), with Gen 7 on the horizon. Meanwhile, Ethernet has likewise increased to speeds of 100Gbps and beyond.

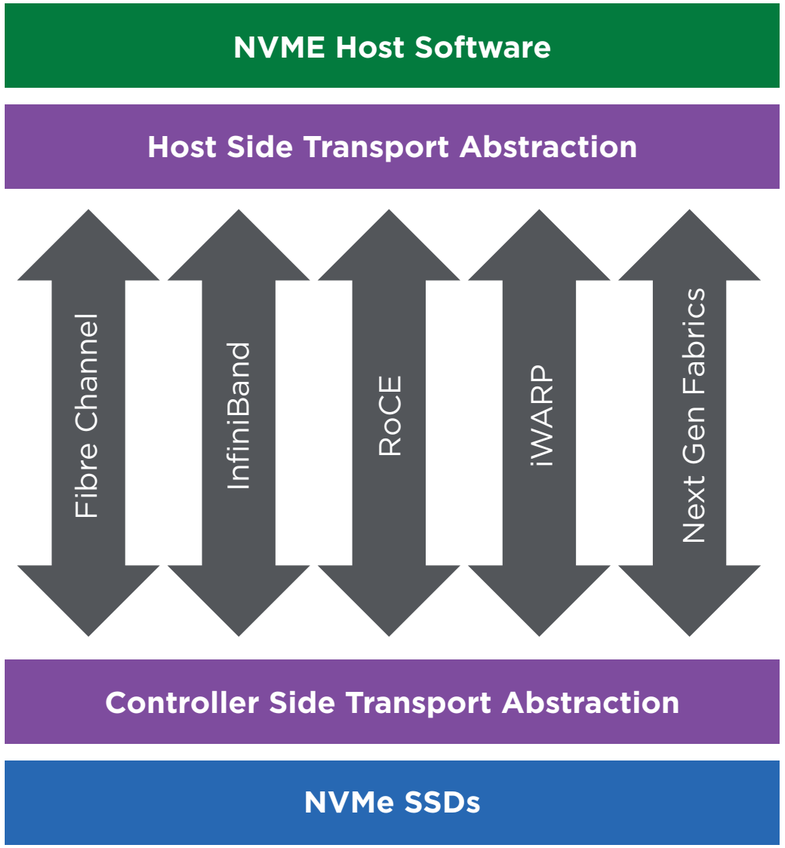


Figure 3: NVMe-oF supports multiple networking protocols.

## Removing the Interface Bottleneck

NVMe-oF is a relatively new technology, but its underlying components are well established in enterprise data centers. One important factor is that it's agnostic about the underlying transport mechanism. Although many enterprise IT shops are focused on using Fibre Channel (NVMe/FC), the NVMe-oF standard also supports remote direct memory access (RDMA) transport mechanisms, such as InfiniBand, RoCE, and iWARP (Figure 3).

## Transitioning to NVMe/FC

For many IT architects making the move to NVMe-oF, Fibre Channel remains the preferred networking technology because of its performance and reliability, plus its support for fabric-based zoning and name services. Better still, Fibre Channel is capable of transporting multiple higher-level protocols at the same time, such as NVMe/FC and FCP (SCSI on Fibre Channel).

For some organizations, the business advantages justify an end-to-end NVMe system, with multiple NVMe storage devices using NVMe-oF with either a Fibre Channel or RDMA interface. This would be blazingly fast, in terms of both IOPS and latency.

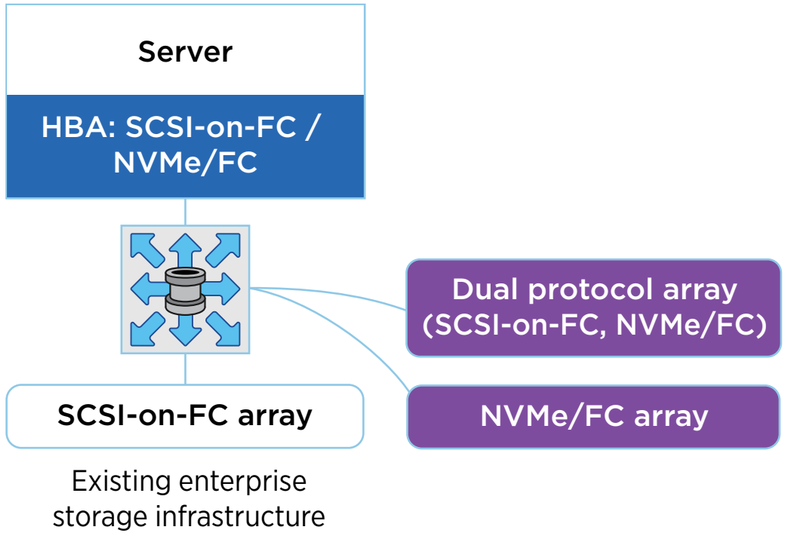


Figure 4: Dual-protocol infrastructure eases the transition from Ethernet to NVMe-oF.

NVMe's multiprotocol support simplifies the buying decisions you'll face in making such an important migration. The key is to take advantage of your organization's existing infrastructure. For many organizations, the transition

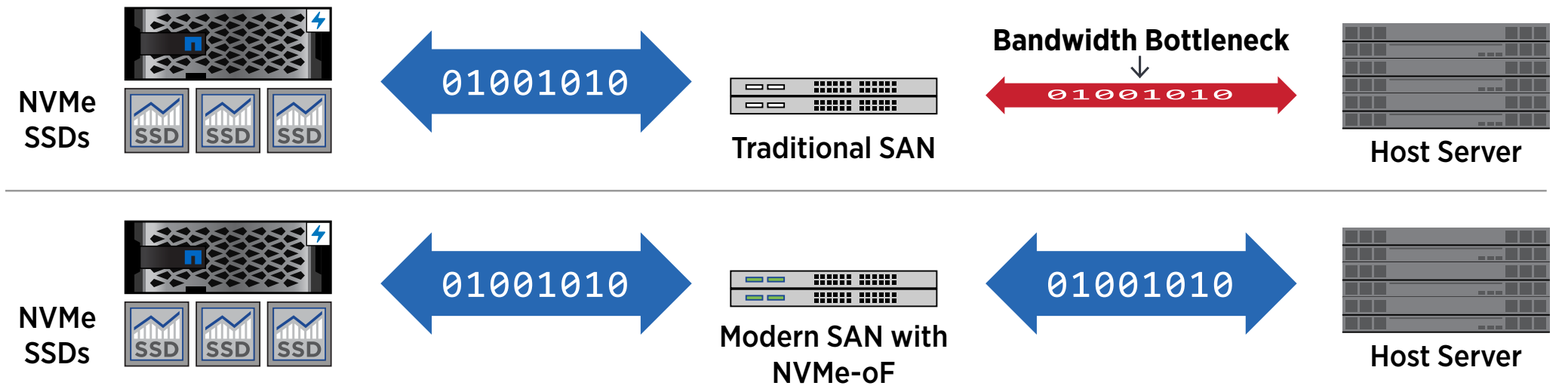


Figure 5: The complete NVMe/FC solution provides a low-latency data path from the server right through to the storage system.

to NVMe-oF will take a couple of years, and that's not a problem.

Instead of replacing one technology with another, a better solution is a dual-protocol Fibre Channel fabric, running concurrent FCP and NVMe/FC traffic. This provides a stable and cost-effective upgrade path, removing uncertainty and risk during a crucial technology transition.

## NVMe/FC Solution Stack

While NVMe is a disruptive technology, you can introduce it nondisruptively into your organization.

As shown in Figure 5, an end-to-end NVMe/FC solution stack has three main components–the server, the storage system, and the fabric infrastructure that connects them. Although it's possible to implement this type of SAN infrastructure with 16Gb Fibre Channel switches

and directors, for best results an end-to-end 32Gb Fibre Channel infrastructure is recommended.

The complete NVMe-oF data path required to achieve maximum performance improvements is available **today** in NetApp solutions incorporating state-of-the-art networking infrastructure from technology partners such as Brocade and Broadcom. Plus, some customers with recent storage systems and Fibre Channel hardware bus adapters will be able to upgrade seamlessly, just by updating drivers and other software.

A sample configuration listing the server hardware and software, fabric infrastructure, and storage systems is shown in Figure 6.

NetApp and Brocade have collaborated to create a concise e-book, NVMe over Fibre Channel for Dummies, which you can download free of charge.

For more detailed information, download this Technical Report to learn how to design and implement an end-to-end NVMe/FC solution.

| | ONTAP 9.4 Stack | Source |
|---|---|---|
| **Server** | | |
| *Server OS* | SLES 12 SP3 | SUSE Linux |
| *Server side HBA* | Emulex LPe32004 32G FC HBA | Broadcom |
| *Server side FC/ NVMe driver* | Broadcom latest drivers | Broadcom |
| | SUSE Linux Enterprise 12 Inbox driver | SLES |
| **Fabric Infrastructure** | | |
| *Switch* | Gen5: Brocade 6505/6510/6520/8500 | Broadcom |
| | Gen6: Brocade G610/G620/G630/X6 | Broadcom |
| | Cisco MDS 9132T | Cisco |
| *Switch firmware level* | Firmware 8.1.1–8.2.x (G610/G620) | Broadcom |
| | Firmware 8.2.x (G630) | Broadcom |
| | Broadcom firmware downloads | Broadcom |
| | Firmware NX-OS 8.2(1), 8.2.(2) | Cisco |
| *Switch Management Software* | Brocade Network Advisor (BNA) 14.4.0 | Broadcom |
| **Storage and ONTAP** | | |
| *Controllers* | A300, A700, A700s, A800 | NetApp |
| *FC/NVMe Protocol license* | ONTAP 9.4 | NetApp |

Figure 6: Complete NVMe/FC solution stack that you can implement today.

# Databases, SANs, and NVMe

The increasing popularity of NVMe is being shaped by workloads that are crucial to organizations today. These workloads range from long-established productivity standards to cutting-edge new apps.

Many of these business-critical workloads run either exclusively or preferentially on SAN systems:

- SANs are built on top of the fastest protocols available (that is, Fibre Channel and iSCSI).
- SANs are optimal for many database workloads, which are crucial business requirements.
- SANs are block-level systems that attach to the host like an external hard drive, a key architectural advantage in many environments.

In fact, most business-critical apps are built on top of databases, especially Oracle Database, Microsoft SQL Server, and SAP HANA.

SANs are ideal for database workloads because they're designed to exceed an enterprise's most rigorous key performance indicators,which typically center around:

- Performance
- Availability
- Ease of use
- Future-proof
- Return on investment

## Enterprise Database Workloads

To optimize performance, database administrators constantly monitor their systems to make sure each hardware component is being fully utilized. This is especially important with database applications, where server sprawl or I/O roadblocks can significantly degrade response time.

For example, a single database server connected to a SATA-based storage system inevitably runs into performance issues as I/O requests must wait to be processed. To solve this problem, IT architects sometimes spin up another server and split the workloads between them, or dedicate one server to user interaction and another to back-end processing. But this is an expensive alternative—a single enterprise database license can cost tens of thousands of dollars per year, not to mention the operational expenses. By the time you add in the costs of power and cooling, NVMe clearly lets you get more done, with less hardware, at lower cost.

## In-Memory Databases

NVMe is ideal for SAP HANA, Apache Spark, and other in-memory database applications that rely on main memory for data storage. These workloads operate with huge datasets that can be larger than the combined cluster memory, which puts enormous stress on the storage system.

In SATA environments, even with RAID protocols, a significant risk exists that a hardware failure will result in a node being temporarily degraded, or even worse, out of service. With NVMe, however, the bus is fast enough to keep up with the CPU, even with today's incredibly fast multicore processors.

# Business Outcomes

Technology research and development continue at a relentless pace. So as NVMe and other disruptive technology innovations come to market, one of the big challenges facing IT leaders is how to make well-grounded investments in the future.

For organizations that are on the path of digital transformation, the key priorities include:

- Using flash and the hybrid cloud to migrate to an emergent modern data center architecture
- Exploiting the strategic value of data, through advanced analytics and artificial intelligence
- Future-proofing core workloads by running them on a high-speed modern SAN architecture

Today, NetApp has the technology, products, and expertise to provide enterprises around the globe with an unparalleled SAN value proposition. In part, it's the product—blazing fast storage arrays and data fabrics that provide the nonstop availability that organizations need.

It's also the ecosystem—systems integrators and value-added resellers who build complete solutions around the products, configuring them to their customers' industry-specific needs.

Beyond that, it's the platform that customers love: NetApp ONTAP® provides world-class data management tools for the modern data center, custom configurations for core enterprise workloads, and the unique NetApp Active IQ® automated support system.

Known for decades as the dominant brand in file-based data storage, NetApp has taken the lead over the past few years in SAN technology as well, with sales of AFF SAN systems more than doubling year over year. In fact, most NetApp AFF customers are implementing SANs. A recent report from market research firm IDC confirms that NetApp is now the fastest growing among the top five SAN vendors, with significant momentum moving forward into the NVMe era.

LEARN MORE

netapp.com/us/products/storage-systems/storage-area-network.aspx

*Source: IDC, WW Quarterly Enterprise Storage Systems Tracker – 2017 Q4, Mar 2018.



NetApp®