FUJITSU

White Paper
Unleashing the Flash Beast – The Power of NVMe

Three new developments are on the way that will revolutionize storage technology: NVMe (Non-Volatile Memory Express), NVMe-oF(NVMe over Fabric) and Storage Class Memory (SCM). These innovations will take Flash technology to the next level and beyond. In the years ahead new storage systems will emerge that can manage huge data streams in parallel while guaranteeing minimum latency at the same time. This will make data-intensive workloads much easier to handle, and also give a boost to Artificial Intelligence, Big Data Analytics and the Internet of Things. This white paper examines these new developments and their impact on storage environments. It also provides you with key insights to guide your planning for the future.

Content

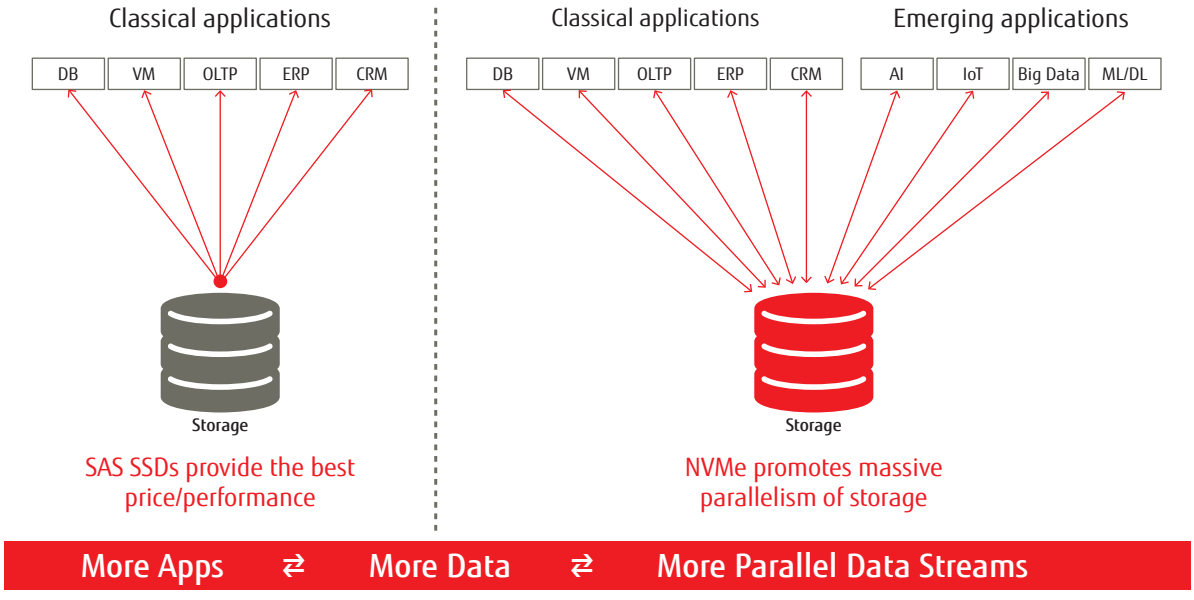
Introduction

Conquering the storage parallelism conundrum

The rise and convergence of new age technologies like Artificial Intelligence, Big Data, Analytics and Internet of Things have led to the advent of next-generation data-intensive workloads and ever-proliferating applications which are exponentially increasing the demand for accelerated data storage.

Today we have a lot of traditional workloads storing their data on slower HDDs (providing maximum storage consolidation) or the much faster SSDs (which offer the best price/performance), both of which reuse traditional interface protocols. With the ever-growing demand for additional data streams to the storage system, unlocking parallel access to storage became a challenge.

This led to the entire storage world coming together to come up with a new interface protocol that will revolutionize storage – and NVMe was the result. NVMe deploys flash storage on a PCIe bus and offers up to 64,000 parallel pathways from the CPU, thus overcoming serial limitations in storage I/O processing. The protocol capitalizes on the multiple parallel, low latency paths to flash storage to offer faster storage response times and higher throughput for speedier application performance.



NVMe: beginnings of a new storage revolution

Every couple of years an innovation comes along which disrupts the technology landscape. In the storage world it was the advent of Flash couple of years back, and now we are on the cusp of another disruption – only this time it's the confluence of three emerging storage technologies – NVMe (Non-Volatile Memory Express), NVMe-oF(NVMe over Fabric) and Storage Class Memory (SCM).

In the not so distant past, Flash SSDs revolutionized the storage world as it easily met the challenges of an insatiable demand for faster data storage. But fast forward a couple of years, and today there is a tremendous need for massive parallelism and minimal latency to meet the demands of emerging workloads, which can only be achieved by the adoption of new interface protocol designed from ground-up for Flash. In other words, SAS-based SSDs provide excellent price/performance and are now the mainstream choice for most storage buyers. But as the SAS protocol was built for disks, a new protocol is now needed to unleash the power of parallelized all-flash storage.

NVMe is the promising new optimized, high-performance interface that is designed for systems that use PCIe based SSDs. NVMe provides a standard access protocol for flash to leverage the PCIe bus to directly connect CPUs to attached storage, thereby reducing latency, increasing throughput and most importantly offering massive parallelization capabilities. So while mainstream Flash SSDs will continue to be available with SAS or SATA for price/performance scenarios, NVMe SSDs will be very attractive for usage scenarios that demand extreme performance. In fact, Gartner statistics project an attach rate of 69% for PCIe NVMe SSDs in servers and 26% for PCIe NVMe SSDs in storage arrays in the next two years.

Although still very nascent, NVMe over Fabric extends the latency, parallelism and performance advantages to the entire fabric (over Fibre Channel, RDMA and TCP, etc.). With NVMe-oF, the potential impact of NVMe would not just accelerate SSDs – it will also power computing, fabrics, storage, system management and more.

The transitions of interface and Media are developing in parallel. While on the interface side NVMe is all set to become more prevalent, on the media side Storage Class memory (SCM) is an emerging memory technology. SCM is created out of flash-based NAND, slightly slower than DRAM but persistent like traditional storage. It provides near DRAM performance at lower costs. As storage becomes more like memory, internal latencies of storage media will nosedive, thereby making protocol latency more significant. In the future, SCM-based arrays will leverage NVMe and NVMe-oF to offer accelerated and highly responsive storage with latencies measured in microseconds. While still evolving, all the technology advances in interfaces and media will reshape the storage industry in the days to come.

Putting the revolution in the right perspective

Today new breakthroughs are often hailed as "revolutionary" – but upon closer look, there is little that seems unusual or groundbreaking about them. But NVMe is different. The following brief comparison of SATA, SAS and NVMe in terms of performance and speed shows why.

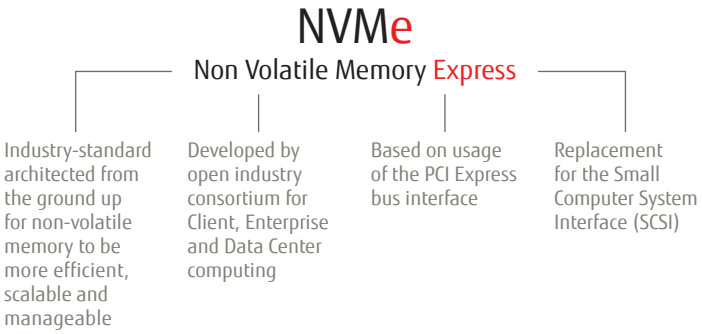
The SATA interface can handle a queue in which 32 commands can be processed.

The SAS interface can handle a queue in which some 256 commands can be processed.

NVMe can handle up to 64,000 queues – and is designed so that up to 64,000 commands can be processed in each queue.

What is NVMe?

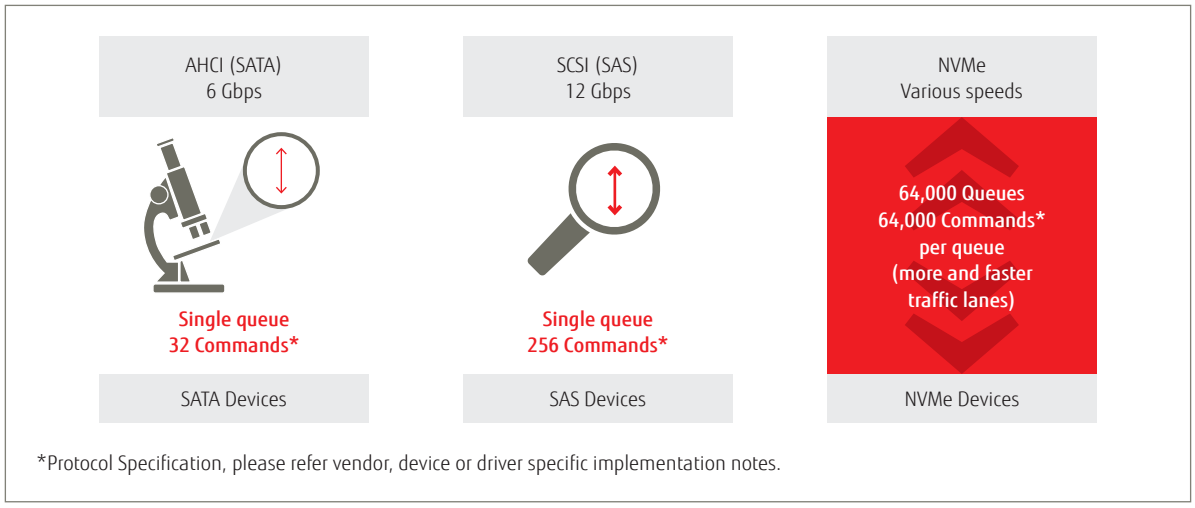
NVM Express Inc. – an organization comprising more than 60 technology vendors – defines NVMe as "an optimized, high-performance, scalable host controller interface designed for Enterprise and Client PCIe SSDs. NVM Express revolutionizes storage by delivering faster access to data and lowering power consumption. This reduces the Total Cost of Ownership (TCO) for enterprises and extends battery life for mobile clients. NVM Express streamlines the legacy storage stack to reduce latency significantly, delivers higher Input/output Operations per Second (IOPS) for lower TCO."



NVMe over PCIe: Massively parallelized storage

NVMe is a high-performance interface for storage – specifically for solid-state storage (not hard drives). It uses the PCI Express (PCIe) bus as the interface, hence the word "express". NVMe is designed to be faster and far more optimized than its predecessors, SAS and SATA. It is a higher-performance, lower-latency alternative to the Small Computer System Interface (SCSI) command set for transferring data between CPUs and local or remotely attached disk drives. SCSI was created during the days of slower storage media, such as hard disk drives (HDDs) and tape, and it thus inherited the innate disadvantages of the complexity, multiple interface layers and lack of parallelism.

PCIe has a number of advantages, including lower latencies, scalable performance/bandwidth, better I/O performance and low power consumption, thanks to its direct connection to the CPU. Currently, PCIe 3.0 can support up to 16 lanes, with each lane supporting1Gb/s of throughput per lane. Besides the noticeable performance improvement, this enables NVMe to support massive concurrence of parallel requests. For example: NVMe can support up to 64,000 queues and 64,000 commands per queue, while SATA can queue only 32 commands in a single queue, and SAS can queue about 256 commands in a single queue.



How NVMe revolutionizes storage technology

The Flash performance booster and latency tamer

NVMe also corrects the weakest link in the development of all-flash storage systems – the interface protocols. NVMe has a streamlined and simple command set that uses less than half the number of CPU instructions needed by SAS and SATA to process an I/O request. This leads to far higher IOPS per CPU instruction cycle and lower I/O latency. Chart 1 highlights the increase in efficiency and reduction in CPU load while using NVMe as compared to SAS and SATA.

NVMe eliminates performance bottlenecks by introducing a more streamlined stack, with a new set of commands and a massive number of I/O queues for direct-connected storage.

NVMe also supports a slew of enterprise features to meet the requirements of highly demanding and compute-intensive applications such as power and thermal management, with no need for customer device drivers, I/O multipath and namespaces for redundancy and load balancing purposes.

NVMe uses less than half the number of CPU instructions to process an I/O request than SAS or SATA!

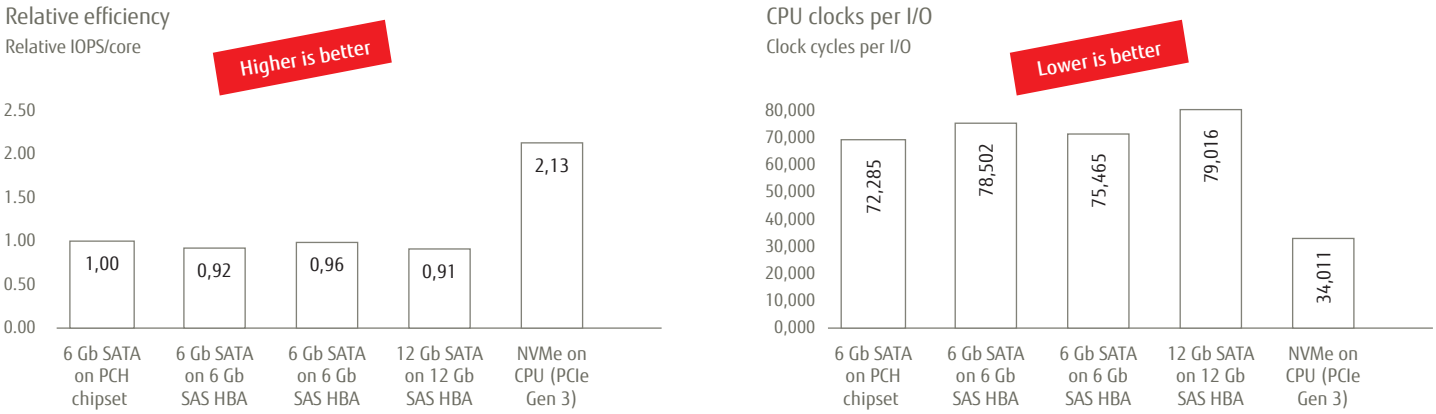


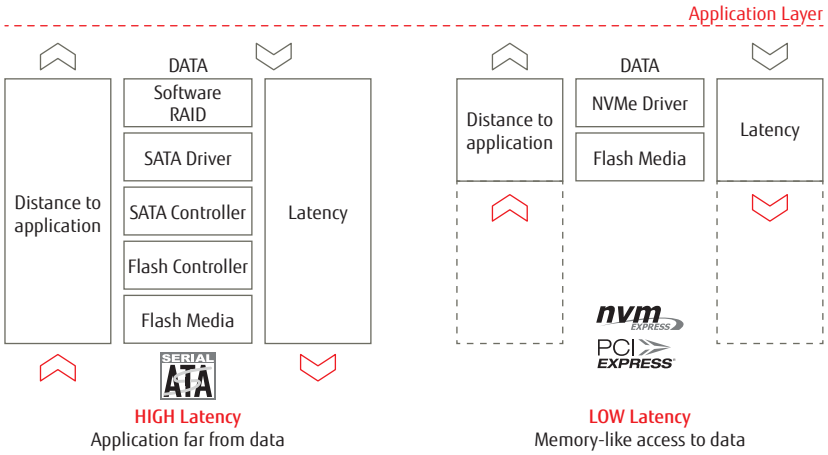
Chart 1 – Source: NVM Express Org Whitepaper

NVMe: simplified storage architecture

The difference between SATA and NVMe centers around how storage is tied up and handled. While SATA was designed to interface with mechanical hard disk drives, it is often used as a typical way to connect to SSDs – and is becoming increasingly inefficient. Basically, you are running SSDs using the same protocols as HDDs!

NVMe, on the other hand, is a specification designed to exploit the inherent advantages of nonvolatile solid state storage. It uses the PCI Express bus to connect the CPU with local or remotely attached storage, thereby eliminating the multiple layers of SCSI command stack and other bottlenecks that plague traditional HDD interfaces. Its efficient I/O architecture and streamlined software stack optimizes application's access to Flash and reduces both latency and CPU overhead to deliver extreme performance.

Simplified protocol stack



Source: NVM Express Org

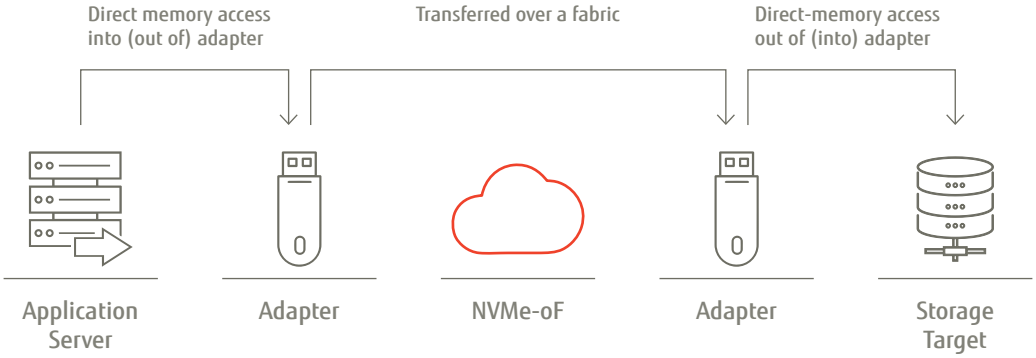
What is NVMe over Fabric?

NVMe over Fabric: next-generation storage fabric for the data center of the future

NVM Express Inc. defines NVMe over Fabric (NVMe-oF) as a "common architecture that supports a range of storage networking fabrics for NVMe block storage protocol over a storage networking fabric. This includes enabling a front-side interface into storage systems, scaling out to large numbers of NVMe devices and extending the distance within a data center over which NVMe devices and NVMe subsystems can be accessed."

While NVMe ultra-high-speed connection interfaces for memory-speed devices primarily exist "within the box," NVMe over Fabric (NVMe-oF) extends the very same capabilities across the entire data center. NVMe-oF specifications will operate the NVMe storage interface protocol across storage networks like RDMA over Converged Ethernet (RoCE), InfiniBand, iWarp and other network fabrics using low-latency RDMA interfaces. With NVMe-oF, the potential impact of NVMe would no longer be limited to just accelerate SSDs – it will power computing, fabrics, storage, system management and more. In short, it will impact the data center strategies of tomorrow.

How NVMe over Fabric works



Source: NVM Express Org

The impact of NVMe on real-world scenarios

The use of NVMe will be appealing for all usage scenarios which are relevant for SAS all-flash today, in addition to the new demanding scenarios of big data analytics, machine learning etc. which are data-driven and need incredibly responsive storage along with sustained high performance.

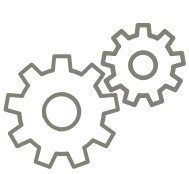
- **Big data analytics:** NVMe SSDs are well-suited for organizations that use analytics platforms to collect, process and analyze vast volumes of data to make agile business decisions – e.g., in aviation, connected cars and online retail.
- **VDI:** NVMe SSDs provide an exponential increase in IOPs, helping VDI platforms to easily overcome high I/O demands and boot storms that are typical in VDI environments.
- **Enterprise applications:** NVMe SSDs eliminate I/O bottlenecks while boosting the performance of business applications from Microsoft, SAP, Oracle and Splunk, among others.
- **Machine learning/deep learning:** Deep learning and machine learning use algorithms to determine a predictable behavior by analyzing seemingly disparate types of data – for example, fraud prevention, speech recognition, image classification, etc. These systems have to potentially retrieve and analyze millions of small files to provide real-time answers to inquiries. NVMe provides massively parallel access to storage with extremely low latencies to accomplish this.
- **HPC:** NVMe allows massive parallelization while processing I/O queues which – coupled with a reduced driver stack and direct CPU access via PCIe bus – translates into excellent speed and performance for sequential and random workloads. NVMe also supports storage consolidation via scale-out architectures.

What's next: adoption guidance

Essential guidance: Invest in SAS all-flash now, keep an eye on NVMe for specific workloads

Fujitsu believes NVMe and NVMe-oF can have a transformational effect on the entire storage industry. As is the case with any disruptive technology, there can be a lot of uncertainty regarding their adoption. As the storage industry collectively continues to adopt new standards and make these technologies mainstream, we have provided some guidance for organizations to make their NVMe journey easier.

- NVMe memory technologies will become available in storage systems from several vendors this year.
  - **Guidance:** To truly benefit from NVMe technology, a storage system needs to be designed from the ground up and be optimized end-to-end. It stands to reason that some early designs may blend current technologies with new ones, thereby compromising the advantages of NVMe. Go only for fully developed NVMe systems and conduct technical evaluations to ensure that the desired performance improvements are realized. There will be a lot of news about NVMe over Fabrics as the next big thing, but as this standard has only been released recently, do not expect mature designs over the next two years.
- This year will be a year of early adopters; current SAS-based all-flash storage will become mainstream.
  - **Guidance:** Current SAS-based all-flash systems are fully mature and increasingly improve the price per performance and price per capacity ratio. Their performance characteristics will meet the demands of most application scenarios. Consequently, their adoption will further grow in the double-digit range. NVMe will not replace these designs, but will be offered as an additional storage tier for usage scenarios with the highest demands on storage response time. Look for vendors who can offer SAS- and NVMe-based all-flash storage with full management compatibility to avoid operational silos.
- Drivers and optimizations for operating systems, hypervisors and adapters are currently being developed and will take more time for mainstream acceptance of NVMe storage.
  - **Guidance:** Expect a long-term co-existence of SAS- and NVMe-based all-flash storage systems until NVMe becomes the new mainstream. SAS-based SSDs will continue to be the choice for the price/performance storage tier while NVMe SSDs will be the choice for the highest performance storage tier. It follows that the performance and latency benefits will appeal initially to very high-end users who demand extreme throughput and IOPS for future workloads, along with lowest possible latency.



SATA/SAS HDDs » **SAS SSDs** » **NVMe SSDs**

The mature and proven

- Focus on maximum storage consolidation
- Balancing between lowest cost of capacity and price-performance in one system
- Use NVMe cache to optimize SATA disks usage

The new mainstream

- Focus on excellent price-performance
- Accelerated storage performance
- Improved storage economics

The emerging

- Focus on extreme performance
- Massive parallelized storage
- Maximum consolidation of existing and future application data

Fujitsu and NVMe: Powering the storage revolution

NVMe offers exciting opportunities to harness the real power of Flash and Fujitsu will introduce NVMe in its storage portfolio. Fujitsu's future storage system based on the NVMe protocol and PCIe bus technology will enable customers to manage massive parallel data access, without the need to increase the number of their storage systems. Fujitsu also plans to offer comprehensive management compatibility between current hybrid, all-flash storage systems and future NVMe storage systems, ensuring operational efficiencies for running differentiated storage tiers in parallel.

- Fujitsu currently leverages NVMe technology as secondary cache memory in the newly launched ETERNUS DX 8900 S4 hybrid array and also in PRIMERGY Servers, offering reliable high performance for latency-sensitive environments.
- Fujitsu will introduce future purpose-built NVMe storage products, designed to complement the existing portfolio of ETERNUS DX hybrid and ETERNUS AF all-flash systems. The future storage systems will support NVMe storage interface protocol to deliver massively parallel data access at unprecedented speeds.
- In the near term, Fujitsu foresees increased adoption of our class-leading SAS-based ETERNUS AF (All-Flash) Portfolio, which offers exceptional price/performance with ultra-low response time, while also fulfilling the most demanding workload requirements efficiently.

The possibilities are infinite! Please contact your Fujitsu representative to learn more about storage from Fujitsu that can be tailored to meet the demands of your current and emerging workloads.

About Fujitsu

Fujitsu is the leading Japanese information and communication technology (ICT) company offering a full range of technology products, solutions and services. Approximately 155,000 Fujitsu people support customers in more than 100 countries. We use our experience and the power of ICT to shape the future of society with our customers.

Fujitsu's storage portfolio balances storage capacity, performance and costs for the complete lifecycle of data – from production, business analytics and big data to backup and long-term archiving.

For more information, please visit
http://www.fujitsu.com/global/products/computing/storage/all-flash-arrays/

Published by
Fujitsu Limited