



White Paper

F-WP004

Benefits of PCIe vs SATA SSD

Table of Contents

Introduction	03
Why NVMe?	04
Benefits of using NVMe	05
Improving User Experience.....	06
NVMe Protocol.....	06
NVMe Management Interface	07
Summary.....	08

Introduction

As the cost of SSDs have come down over the past few years with improved lithography and 3D stacking technology, the penetration of SSDs has grown in consumer applications such as laptops and desktop PCs as well as into servers in the enterprise environment. With a significant boost in storage capacity and performance, the NAND Flash technology behind Solid State Drives offers much faster transfer speeds with increased system functionality.

Now that SSDs are a viable option for many environments, the decision of whether to stay with a SATA interface or the emerging PCIe NVMe protocols also comes into play. The Non-Volatile Memory Express (NVMe) protocol for NAND Flash storage solutions accelerates transfer speeds by overcoming legacy constraints in existing storage interfaces and protocols such as SATA, which addressed mechanical Hard Disk Drive Storage. Even with revisions over time, the current SATA interface bandwidth does not utilize the full potential of NAND Flash based SSDs.

Why NVMe?

The Non-Volatile Memory Express (NVMe) protocol is especially designed for NAND Flash storage solutions and accelerates data transfer speeds six and seven times the sequential read/write speeds of SATA. In addition to providing a leap in transfer speed performance, NVMe SSDs are smaller, lighter, and more efficient. Using an edge connector, the 22mm-wide M.2 form-factor reduces the overall size of internal storage devices and is easier to install without cables and bulky disk housing. (See Figure 1 and 2). Flexibility on the height and length dimensions allow customization for storage solutions as needed.

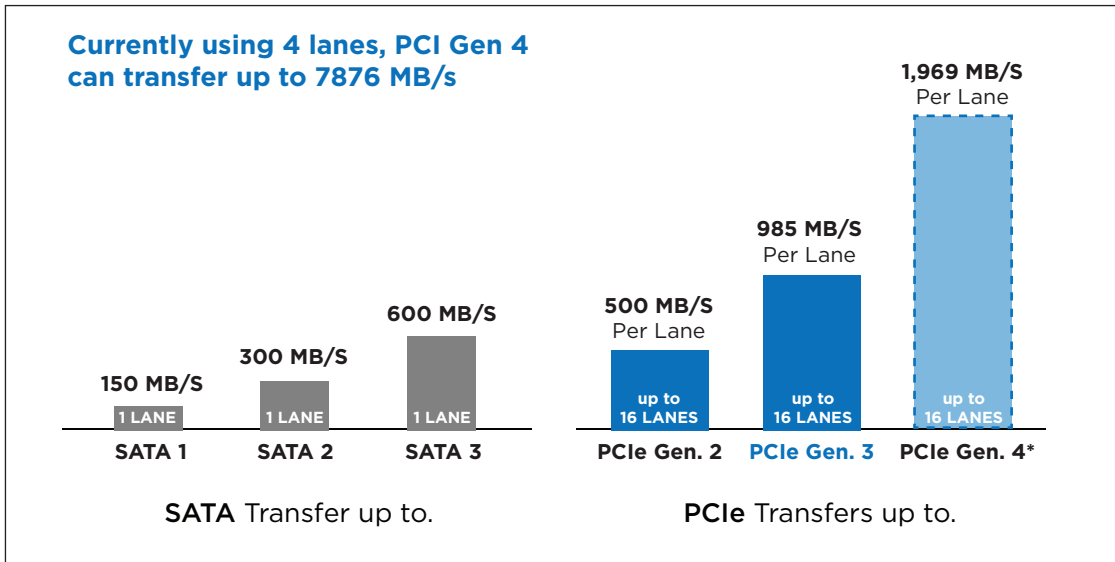


Figure 1

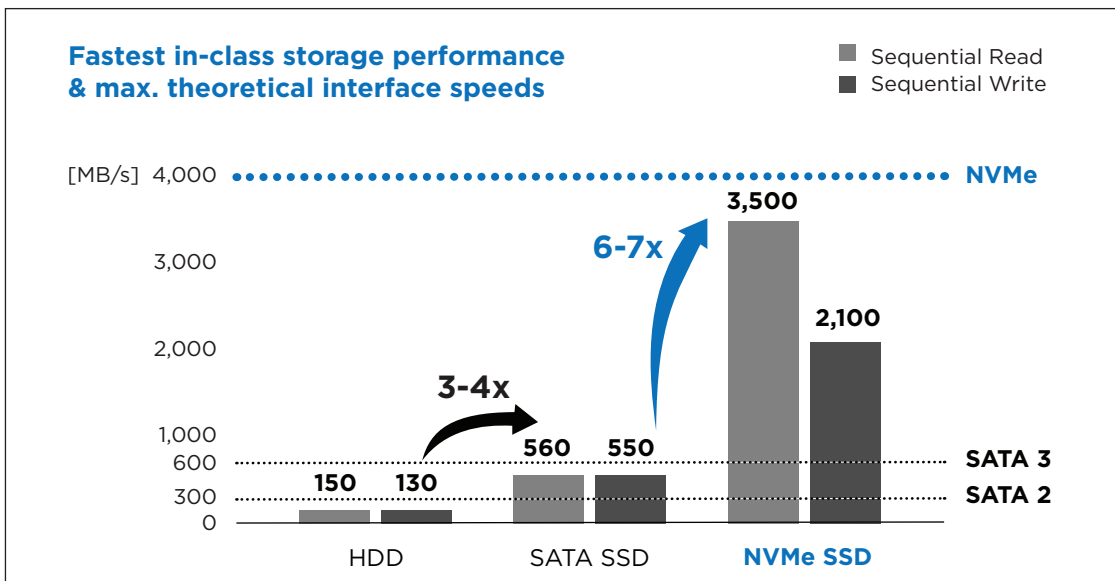


Figure 2

Benefits of using NVMe

NVMe supports a large number of deep queues and commands per queue. It also enables parallelism in multiple cores of the system, while supporting the Non-Uniform Memory Architecture (NUMA). In addition, lockless command submission and completion provides much better latency and bypasses all SCSI layers. Commands use simple PCIe register writes that allow much more flexibility in issuing 64-bit commands without any contention.

NVMe provides a Controller Memory Buffer (CMB) feature that allows a host to prepare commands in controller memory. The addition of the CMB feature means the controller no longer needs to fetch command buffers through PCIe reads. This helps give NVMe a more streamlined interface, focused on passing memory blocks instead of SCSI commands. This feature provides lower latency than other protocols. The NVMe arbitration mechanism improves flexibility in providing priority on a command basis for a better adherence to Service Level Agreements (SLAs). The protocol enables reservations in data center applications and all major operating systems also include native NVMe driver support for the ease of deployment. (See Figure 3)

NVMe also includes a Host Memory Buffer (HMB) that provides additional memory for its non-critical data structures. This allows DRAM-less NVMe controllers to use system memory which is well suited for client and mobile NVMe controllers.

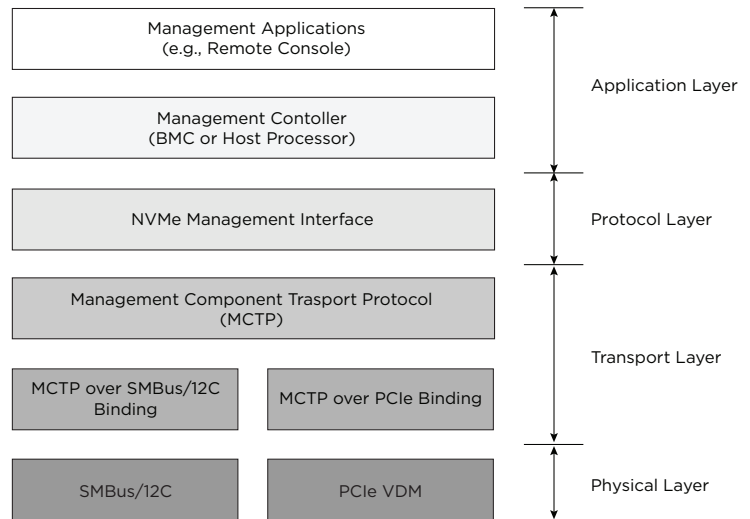


Figure 3: Four-Layer NVMe Management Hierarchy
(Source: NVMe Express Org)

NVMe's lower I/O latency is achieved by reducing the block I/O stack overhead in the operating system, I/O process/resource locks and resource contentions. A version of a planned hot-swap function is available for 2.5-inch form-factor devices. All modern processors have implemented NVMe PCIe 3.0 directly on the CPU. There is no need for a Host Bus Adapter (HBA) or an I/O controller. These absences reduce latencies and potential for errors. Reduced power consumption in 2.5-inch and M.2 form-factor NVMe PCIe Gen3 x4 devices can deliver 3000MB/s at 0.55W, more than five times the performance for the same power consumption, as a SATA device delivering 600MB/s at 0.6W.

An important point to take into consideration is that PCIe NVMe SSDs consume more power and are more expensive than a standard 2.5" SATA SSD. NVMe based SSDs are optimal for in system storage of data (such as configuration details, state recovery etc), but less cost effective for large volumes of data.

Improving User Experience

NVMe has many benefits compared to SATA or SCSI flash storage. Direct connection to the CPU provides lower latency compared to a connection via I/O controllers, multiplexers or storage networks. The scalable performance of up to 1GB/s per lane is achievable using PCIe 3.0. Currently, the NVMe standard protocol supports up to 40 PCIe single-lane devices on a CPU system. Performance will continue to improve as designs are in the works to support up to 48 devices in the future.

NVMe also provides end-to-end data protection using the Data Integrity Field/Data Integrity extension (DIF/DIX). Using Enterprise, Opal™ and Opal light protocols from the TCG (Trusted Computing Group) provides this increased security. Opal specifies minimum acceptable core capabilities of a storage device tailored for the PC client and value Enterprise markets requirements. The Opal specifications provide a comprehensive architecture for putting storage devices under policy control as determined by the trusted platform host, the capabilities of the storage device to conform to the policies of the trusted platform and the life-cycle state of the storage device as a trusted peripheral. Opal protects the confidentiality of stored user data against unauthorized access once it leaves the owner's control by de-authentication following a power cycle.

NVMe Protocol

To optimize storage and retrieval, NVMe uses up to 64K commands per queue on up to 64K I/O queues for parallel operation. It has a paired submission-and-completion queue mechanism in host memory. Host software places commands into the submission queue and the NVMe controller places command completions into an associated completion queue. Multiple submission queues may report completions on a single completion queue, provided the controller supports arbitration with different priorities. Message-Signaled Interrupts Extended (MSI-X) and interrupt steering is supported as well. Optionally, support for many Enterprise capabilities like end-to-end data protection (compatible with T10 DIF and DIX standards), enhanced error reporting, autonomous power state transitions for clients and hinting are included. It uses a simple streamlined, minimum set of only 13 commands, as listed in Table 1 below.

Only 10 admin commands are required

ADMIN COMMANDS
Create I/O Submission Queue
Delete I/O Submission Queue
Create I/O Completion Queue
Delete I/O Completion Queue
Get Log Page
Identify
About
Set Features
Get Features
Asynchronous Event Requests
<i>Firmware Activate (Optional)</i>
<i>Firmware Image Download (Optional)</i>
<i>Format NVM (Optional)</i>
<i>Security Send (Optional)</i>
<i>Security Receive (Optional)</i>

Only 3 I/O commands are required

NVM I/O COMMANDS
Read
Write
Flush
<i>Write Uncorrectable (Optional)</i>
<i>Compare (Optional)</i>
<i>Dataset Management (Optional)</i>
<i>Write Zeros (Optional)</i>
<i>Reservation Register (Optional)</i>
<i>Reservation Report (Optional)</i>
<i>Reservation Acquire (Optional)</i>
<i>Reservation Release (Optional)</i>

Table 1 - Simple Command Set – Optimized for NVMe (Source: NVMe Express Org)

NVMe Management Interface

The NVMe Management Interface (NVMe-MI) defines an out-of-band management mechanism that is independent of physical transport and protocol. It standardizes the out-of-band management to discover and configure NVMe devices and it maps the management interface to one or more out-of-band physical interfaces, such as I2C (SMBus) or PCIe VDM.

The NVMe-MI carries the commands needed for systems management, which includes the following elements:

- Inventory
- Configuration
- Health Status Monitoring
- Change Management

Single Root I/O Virtualization (SR-IOV) virtualizes one PCIe device so it appears to be many devices. This procedure eliminates I/O bottlenecks in virtualized server environments and improves data rates up to full PCIe bandwidth.

There is a programming interface providing out-of-band management of NVMe Field Replaceable Units (FRU). An FRU is a circuit board, part or assembly that is easily removed and replaced from a computer or other piece of equipment. A user or a technician can replace an FRU without powering down or having to send the entire product or system to a repair facility.

Some major NVMe protocol options are:

- Up to 64K I/O queues, with up to 64K commands per queue
- Priority associated with each I/O queue with well-defined arbitration mechanism
- All information for a 4KB read request is in the 64B command itself
- Efficient small random I/O operation
- Efficient and streamlined command set
- MSI/MSI-X and interrupt aggregation
- No un-cacheable/MMIO register reads required
- A maximum of one MMIO register write is necessary
- Multiple namespaces
- Efficient I/O virtualization architectures like SR-IOV
- Multipath I/O, including reservations
- Robust error reporting and management capabilities
- End-to-end data protection (DIF/DIX)

Summary

NVMe enables customers to utilize the full performance and latency potential of flash storage, like higher I/O performance, by efficiently supporting more processor cores, lanes per device, I/O threads and I/O queues. NVMe eliminates the overhead required for processing SCSI and ATA I/O commands. NVMe implements simplified command processing because all the commands are the same size and in the same location or position. NVMe is less complex, more efficient, more serviceable and easier to use when compared to legacy systems.

SMART Modular's MP3000 PCIe NVMe SSD product line offers multiple form factors like M.2 2280, M.2 22110 and the E1.S. These SSD's are available in commercial and industrial temperature tested variations which are equipped with power loss protection. SMART also offers TCG Opal encryption as a standard on all of the MP3000 product line.



For more information, please visit: www.smartm.com

Headquarters/North America

T: (+1) 800-956-7627 • T: (+1) 510-623-1231
F: (+1) 510-623-1434 • E: info@smartm.com

Latin America

T: (+55) 11 4417-7200 • E: sales.br@smartm.com

Asia/Pacific

T: (+65) 6678-7670 • E: sales.asia@smartm.com

EMEA

T: (+44) 0 7826-064-745 • E: sales.euro@smartm.com

Customer Service

T: (+1) 978-303-8500 • E: customers@smartm.com