Preparing for Your Storage Future with NVMe™ SSDs

NVMe™ SSDs have moved into the server neighborhood delivering larger capacities, extended enterprise-class capabilities and, of course, the accelerated bandwidth and IOPS (input/output operations per second) performance that today's applications demand. They can connect directly to server CPUs through the PCIe® interface bus, bypassing the host bus adapters (HBAs) that SAS and SATA drives require for backplane connections. With the use of a streamlined SSD stack developed specifically for flash-based storage media, and lanes for data to flow, NVMe SSDs can move PCIe Gen3 data at up to one gigabyte per second (GB/s) per lane, sufficiently feeding today's fast CPUs with data, and replacing enterprise SATA SSDs in servers at a rapid rate.

The NVMe SSDs on the market today commonly feature enterprise-class reliability, availability and serviceability, and deliver the highest SSD performance, but at a premium cost. For some, the cost to move from enterprise SATA SSDs to enterprise NVMe may be prohibitive. Enter data center NVMe SSDs. This new category of NVMe SSDs is designed primarily for read-intensive applications in cloud / hyperscale data centers and focus on delivering predictable performance and quality of service. Compared to enterprise SATA SSDs, data center NVMe SSDs deliver significantly higher performance, larger capacities and more capabilities at a comparable price.

The market for enterprise and data center NVMe SSDs is on the rise and forecasted to represent about 42.5% of the total SSD use in the enterprise segment (in units) by the end of 2019[1]. This majority position is expected to grow to 75% by the end of 2021 and 91% by the end of 2023[1]. Conversely, enterprise SATA SSDs are expected to decline to about 45.5% of the total SSD consumption in the enterprise market by the end of 2019, continuing to fall to 17% by the end of 2021, and 5% by the end of 2023[1].

SATA performance has plateaued with no planned roadmap in its future. Enterprise customers are moving to NVMe SSDs to not only meet current application performance requirements, but also for cloud-based data center architectures. Industry innovations that include the PCIe 4.0 specification, tri-mode HBA / RAID controllers and U.3 universal backplanes will further enable NVMe SSDs to become the leading storage interface in the future.

Differences Between Data Center and Enterprise NVMe SSDs

Data center NVMe SSDs are designed to be cost-effective, high volume drives for cloud-based architectures. They have little to no custom firmware or specialized features than their enterprise counterparts. With fewer bells and whistles, and optimized performance, users expect to pay less versus full-featured enterprise drives. Enterprise NVMe SSDs deliver the highest industry performance and include enterprise-class features that customers are willing to pay a premium for such as dual port, end-to-end data protection, and wear / error reporting.

KIOXIA (formerly known as Toshiba Memory) offers both data center NVMe SSDs (CD5 Series) and enterprise NVMe SSDs (CM5 Series). These NVMe SSDs are developed using 64-layer BiCS FLASH™ 3D TLC technology and are available in 960GB[2] to 7,680GB capacities with 15,360GB capacity available with the CM5 Series. The performance differences[3] between the two NVMe SSD platforms are shown in Chart 1.

Read / Write Operation	Data Center NVMe SSD CD5 Series (1 DWPD*)	Enterprise NVMe SSD CM5 Series (1 DWPD*)	CM5 Series Advantages
Sequential Read (128K block size)	3,000MB/s	3,350MB/s	1.17x
Sequential Write (128K block size)	1,800MB/s	3,040MB/s	1.69x
Random Read (4K block size / 256 QD^)	500,000 IOPS	770,000 IOPS	1.54x
Random Write (4K block size / 256 QD^)	50,000 IOPS	80,000 IOPS	1.60x

Chart 1: Performance comparisons between data center NVMe SSDs and enterprise NVMe SSDs
**Drive Write(s) per Day*
^Queue Depth

Future Cloud Data Center Architectures

NVMe SSDs (both enterprise and data center) are expected to dominate SSD use in the enterprise segment in support of on-premises and cloud infrastructures. These architectures are general-purpose and have commonly utilized a pre-configured allocation of compute and storage resources to meet a whole slew of application workload demands. As data-intensive applications proliferate on a larger scale, these general-purpose infrastructures can limit performance, capacity and scalability. Disaggregation of compute, storage and network resources is becoming an accepted practice to best utilize the resources, requiring they scale independently of each other.

Moving from a general-purpose direct-attached storage (DAS) architecture to a disaggregated, shared storage model via the NVMe over Fabrics[4] (NVMe-oF™) specification presents several resource utilization benefits and delivers the high-performance and low-latency advantages of NVMe SSDs as if they were locally attached. Storage volume management software is available to better disaggregate NVMe SSD storage and share these resources in an orchestrated cloud environment.

Replacing Enterprise SATA with NVMe SSDs

Today's servers feature SAS HBAs or RAID cards that enable SAS and SATA SSDs/HDDs to be used in the same drive bay, creating a flexible storage environment. Though SATA drives can be easily swapped with SAS drives in a SAS-enabled server, there has been no support for NVMe SSDs, as they do not connect directly to a SAS RAID controller or a SAS/SATA backplane. With this limitation, many of today's data centers still use SAS-based hardware RAID to achieve enterprise-class fault tolerance and performance. Moving to an NVMe-based SSD configuration typically requires the purchase of NVMe-enabled servers from such leading companies as Dell EMC™ and Hewlett Packard Enterprise®.

Server evolution is moving toward universal backplanes that combine SAS, SATA and NVMe protocol support into a single backplane and can be managed by a Universal Backplane Management (UBM)-compliant system (SFF-TA-1005 specification[5]). This enables SAS/SATA SSDs/HDDs and NVMe SSDs to be mixed and matched within a UBM-enabled backplane (Figure 1). The SFF-TA-1001 specification[6] (also known as U.3) defines the links between SFF-8639 drive and backplane connectors that may accommodate drives supporting NVMe, SAS or SATA interface protocols from one PCIe slot. As such, U.3 is a tri-mode standard that is built on the U.2 specification using the same SFF-8639 connector. All drives and backplane designs must be compliant to the U.3 specification for full interoperability and functional with a UBM framework.

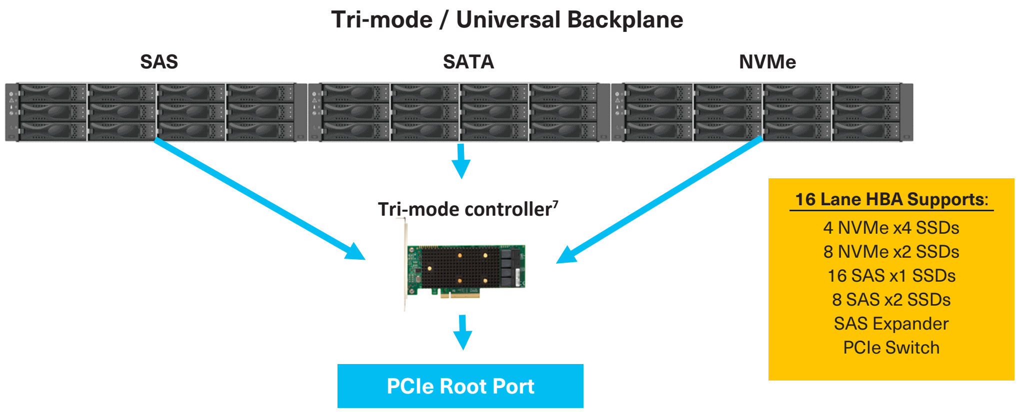

Figure 1: The U.3 tri-mode universal storage configuration for SAS, SATA and NVMe protocols

In the U.3 configuration, the tri-mode controller establishes connectivity between the host server and the drive backplane, supporting SAS/SATA and NVMe storage protocols. The link definition and design for a multi-protocol accepting device connector based on SFF-8639 technology are included in the SFF-TA-1001 link definition specification. The SFF-8639 connector provides for a PCIe x4 connection and can also accommodate separate SAS/SATA host connections. All drives are managed by the UBM framework, which is identical across all drive types regardless of the interface protocol (SAS, SATA or NVMe) or the storage media (HDDs or SSDs).

The ability to add, replace or interchange SSDs within one server and universal tri-mode backplane configuration will help to reduce TCO and storage deployment complexities. The universal backplane also provides a viable replacement path between SATA, SAS and NVMe storage media and maintains backwards compatibility with U.2 NVMe-based platforms, all of which will help to increase adoption of NVMe SSDs in servers.

NVMe SSDs will be Faster with PCIe 4.0

The fourth generation of the PCIe interface was announced at Computex 2019, doubling the bandwidth per link. Theoretically, a PCIe 4.0 x16 slot can deliver up to 32 gigabytes per second (GB/s) of data traveling in each direction versus up to 16GB/s from a PCIe Gen3 x8 slot. In addition to a lot more speed, PCIe 4.0 also has better signal reliability and integrity for improved performance. It is also backward- and forward-compatible where PCIe Gen3 devices work with PCIe 4.0 motherboards, and PCIe 4.0 devices work with PCIe Gen3 motherboards.

The first wave of PCIe 4.0-based enterprise and data center NVMe SSDs were announced at Flash Memory Summit 2019 by KIOXIA (formerly known as Toshiba Memory). It included the CD6 Series data center NVMe SSDs and the CM6 Series enterprise NVMe SSDs. Both platforms feature 96-layer BiCS FLASH™ 3D TLC flash memory in 960GB to 15,360GB capacities. The CM6 Series also includes a 30,720GB capacity at the high-end. Both series' support a 1x4 configuration while the CM6 Series enables 2x2 configurations where dual ports are used for redundancy, fault tolerance and high availability.

As performance continues to be the consistent theme of PCIe 4.0, the CD6 Series and CM6 Series offer performance and power improvements of up to 30% when compared to previous generation PCIe Gen3-based SSDs, and deliver maximum data transfer rates of up to 6,700MB/s[8]. This equates to a performance increase over SATA speeds by a factor of up to 12x (Figure 3). Support for U.3 operations and tri-mode backplanes enable these high-performance NVMe SSDs to plug directly into any U.3-enabled SSD socket.

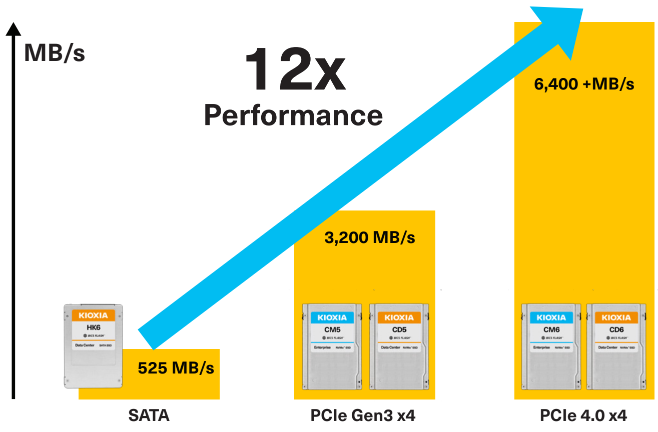

Figure 3: NVMe SSDs based on PCIe 4.0 deliver up to 12x performance gains over SATA SSDs

Summary

The server and storage market category for NVMe SSDs now includes enterprise and data center models that are penetrating server platforms at an increasing pace. Utilizing a streamlined SSD stack developed specifically for flash-based media, coupled with the ability to move data at up to 1GB/s per lane (for PCIe Gen3), NVMe SSDs can better feed server CPUs with data versus SATA SSD performance that has plateaued and has no planned roadmap in its future. Customers are moving to NVMe SSDs for their enterprise performance requirements and cloud-based data center architectures that will soon be enabled by faster, more efficient infrastructures, such as PCIe 4.0, that deliver more bandwidth for data to travel. New U.3 tri-mode architectures will soon enable NVMe SSDs to plug into any SSD socket for use, interchange, or replacement within U.3-based servers.

The next-generation of performance-centric and latency-sensitive applications require IT and DevOps organizations to rethink their enterprise and data center storage strategies, and at the center of this requirement is heightened demand for NVMe SSD storage. With NVMe SSD adoption in data centers expected to grow to 91% by the end of 2023[1], replacing SATA in servers, planning for NVMe SSD storage will be key to support a broad range of cloud-based and enterprise applications.

Notes:

[1] Source IDC: Worldwide Solid State Drive Forecast Update, 2019-2023, Market Forecast Table 12, Jeff Janukowicz, December 2019, IDC #44492119.

[2] Definition of capacity: KIOXIA Corporation defines a gigabyte (GB) as 1,000,000,000 bytes and a terabyte (TB) as 1,000,000,000,000 bytes. A computer operating system, however, reports storage capacity using powers of 2 for the definition of 1GB = 2^{30} bytes = 1,073,741,824 bytes, 1TB = 2^{40} bytes = 1,099,511,627,776 bytes and therefore shows less storage capacity. Available storage capacity (including examples of various media files) will vary based on file size, formatting, settings, software and operating system, or media content. Actual formatted capacity may vary.

[3] An OLTP application was used for measurement of server-side performance to provide the data locality benefits of direct-attached storage (high-performance / low-latency). The results showcase SSD interface bandwidth and performance and how many operations/transactions that a server's CPU can process. The performance measurements were derived from existing KIOXIA SSD products configured with 960GB capacities that included the CD5 Series data center NVMe SSDs and CM5 Series enterprise NVMe SSDs.

[4] The NVMe-oF specification defines a common architecture for NVMe block storage over a storage networking fabric. It enables storage disaggregation and the ability to pool resources and provision the right amount of storage for each application workload.

[5] The SFF-TA-1005 Universal Backplane Management (UBM) was developed by the SSD Small Form Factor Working Group of SNIA. The specification is available at http://www.snia.org/sff/specifications.

[6] The SFF-TA-1001 Universal x4 Link Definition for SFF-8639 was developed by the SSD Small Form Factor Working Group of the Storage Networking Industry Association (SNIA), and ratified in October 2017. The specification is available at http://www.snia.org/sff/specifications.

[7] Source: Broadcom Inc. – Broadcom 9400-16i tri-mode storage adapter.

[8] Based on KIOXIA testing and specifications. Target specifications are subject to change.