KIOXIA

Top 5 Reasons

Why Deploy PCIe® 4.0 NVMe™ SSDs In Your Data Center?

With the advent of the PCIe 4.0 specification revision, data can be moved twice as fast, enabling external devices, such as SSDs, GPUs and NICs, to deliver I/O even faster than before. The speed upgrade can benefit data-intensive and computational applications such as cloud computing, databases, data analytics, artificial intelligence, machine learning, container orchestration and media streaming, as well as emerging applications that enable storage resource disaggregation over common networks, delivering nearly the same high-performance, low-latency benefits as if the drives were locally attached. Key server and SSD vendors are developing solutions that conform to the PCIe 4.0 standard (and associated NVMe protocol) for availability at the end of Q1 2020. There are a number of reasons to deploy PCIe 4.0 NVMe SSDs in your data center, but the top five supported by KIOXIA include:

1. ***Twice the Performance***
   (improves SSD data transfers from 8 GT/s to 16 GT/s, or from ~1 GB/s to ~2 GB/s per lane)
2. ***Improves Energy Efficiency***
   (doubles SSD performance per watt reducing footprint and power requirements)
3. ***Supports Emerging U.3 Tri-mode Backplane Infrastructures***
   (conforms to the SFF-TA-1001 specification[1] and SFF-TA-1005 specification[2])
4. ***Choices in SSD Class***
   (features enterprise and data center SSD options to fit specific application requirements)
5. ***Supports Cloud Data Center Architectures based on NVMe-oF™***
   (disaggregates SSDs from compute nodes for availability as shared network-attached resources)

Twice the Performance

The PCIe interface utilizes data lanes to move data independently, and at maximum speed to keep today's fast CPUs continually fed with data. The latest PCIe 4.0 specification revision released by the PCI-SIG standards group is designed to increase data transfer speeds from 8 gigatransfers per second (GT/s) in PCIe 3.0 to 16 GT/s. In other words, the PCIe 4.0 interface can move data at approximately 2 gigabytes per second (GB/s) per lane versus almost 1GB/s per lane supported by PCIe 3.0, doubling performance while delivering a 4-lane (x4) bandwidth of approximately 8 GB/s (Table 1).

Specification			Throughput				
PCIe Revision	**Introduced by PCI-SIG**	**Transfer Rate (GT/s)^**	**x1 (GB/s)***	**x2 (GB/s)**	**x4 (GB/s)**	**x8 (GB/s)**	**x16 (GB/s)**
3.0	2010	8.0	0.9846	1.969	3.94	7.88	15.75
4.0	2017	16.0	1.969	3.938	7.88	15.75	31.51

^ in gigatransfers per second
** in gigabytes per second*

Table 1: PCIe 3.0 vs 4.0 revisions and their respective performance capabilities (Source: PCI-SIG)

The performance upgrades can be validated using KIOXIA's newly announced enterprise-class CM6 Series of PCIe 4.0 NVMe SSDs for comparison with the previous generation CM5 Series NVMe SSDs (based on PCIe 3.0). When tested[3] internally using 3,200 GB[4] capacity drives at three Drive Writes per Day[5] (3 DWPD) in a mixed-use environment, the performance comparisons are as follows (Table 2):

Read / Write Operation *(for Read/Write Latency, lower is better)*		CM6-V Series SSD (3,200 GB) PCIe 4.0	CM5-V Series SSD (3,200 GB) PCIe 3.0	CM6 Series SSD PCIe 4.0 Advantage
Sequential Read	*(128 KB; QD=32)*	6,900 MB/s^	3,350 MB/s	+105%
Sequential Write	*(128 KB; QD=32)*	4,200 MB/s	3,040 MB/s	+38%
Random Read	*(4 KB; QD=256)*	1,400K IOPS*	750K IOPS	+86%
Random Write	*(4 KB; QD=32*	350K IOPS	160K IOPS	+118%
Mixed 70%R/30%W	*(4 KB; QD=256)*	750K IOPS	360K IOPS	+108%

^ in megabytes per second
** in input/output operations per second*

Table 2: Performance comparisons between PCIe 4.0 and PCIe 3.0 using KIOXIA enterprise-class NVMe SSDs

From these results, the CM6 Series delivers significant performance gains over the previous CM5 Series, and reflect the performance increases associated with the PCIe 4.0 interface, doubling transfer rates and data rates over PCIe 3.0 based SSDs.

Improves Energy Efficiency

The increased performance gains outlined in Table 2 can also be utilized to measure improvements in energy efficiency. Performance per watt is one example where a computing system can be measured by the rate of computation for each watt of power it consumes. Table 2 can also be presented from a total cost of ownership (TCO) perspective using CM6 and CM5 Series SSDs to validate outcomes. Since both series' support power envelopes from 9 watts to 25 watts, and active power envelope of 25 W was used in this example to demonstrate performance per watt as follows:

Read / Write Operation *(3,200 GB capacity SSDs - 25W power)*		CM6 Series SSD PCIe 4.0	CM5 Series SSD PCIe 3.0	CM6 Series Power Advantage
Sequential Read	*(128 KB; QD=32)*	~276 MB/s per watt	~134 MB/s per watt	+105%
Sequential Write	*(128 KB; QD=32)*	~168 MB/s per watt	~121 MB/s per watt	+38%
Random Read	*(4 KB; QD=256)*	~56K IOPS per watt	~30K IOPS per watt	+86%
Random Write	*(4 KB; QD=32*	~14K IOPS per watt	~6.4K IOPS per watt	+118%
Mixed 70%R/30%W	*(4 KB; QD=256)*	~30K IOPS per watt	~14.4K IOPS per watt	+108%

Table 3: Performance comparisons between PCIe 4.0 and PCIe 3.0 using KIOXIA enterprise-class NVMe SSDs

The CM6 Series delivers significant power/performance efficiency advantages over the previous CM5 Series, and reflect the performance per watt increases associated with the PCIe 4.0 interface, improving energy efficiency over PCIe 3.0 while reducing footprint and power requirements.

Supports Emerging U.3 Tri-mode Backplane Infrastructures

PCIe 4.0 NVMe SSDs are also helping to drive emerging U.3 tri-mode backplane infrastructures that combine SAS, SATA and PCIe interfaces into one backplane managed by a Universal Backplane Management (UBM)-compliant system (the SFF-TA-1005 specification[2]). As such, SAS/SATA SSDs and HDDs, and PCIe 4.0 NVMe SSDs, can be mixed and matched within this UBM-enabled backplane, and are compatible with U.2 NVMe platforms. The SFF-TA-1001 specification[1] (also known as U.3) defines the links between the drive and backplane connectors so that PCIe, SAS or SATA interfaces, and their respective protocols, can be supported from one interface slot. The ability to add, replace or interchange SSDs within one universal tri-mode backplane configuration helps to reduce TCO and storage deployment complexities, while providing a viable replacement path between SATA, SAS and PCIe/NVMe storage media. Its backwards compatibility with U.2 will help increase U.3 adoption in servers.

Choices in SSD Class and Key Applications

The PCIe 4.0 NVMe SSD market includes both enterprise and data center types, and reported as one category by market research firms:

- **Enterprise-class NVMe SSDs** are designed to run 24 hours/7 days per week in data center servers and storage and deliver the highest performance any class of SSD can currently achieve. They include such features as dual-port, larger capacities, read-intensive and mixed-use endurances, and data protection (data integrity checking, high reliability, media wear reporting and error reporting).

They are targeted for high-performance server applications, such as:

- ***a. Databases:*** where low latency, dual-ports (for multi-path and high-availability), and the highest available transactions per minute are key capabilities
- ***b. Data analytics:*** where high sequential and random read bandwidth benefit analysis and searches
- ***c. Compute-side AI and ML:*** where fast data transfers into DRAM and GPUs are required during the staging phase as well as immediately validating data for corruption during the checkpoint phase are key benefits

- **Data center NVMe SSDs** are designed for scale out and hyperscale environments where read performance, Quality of Service (QoS) and power efficiency are key metrics. These SSDs are optimized for general purpose server applications and scale out environments, such as:
  - ***a. Cloud compute*** where high-performance and low-latency are primary benefits, especially versus SATA
  - ***b. Container orchestration*** where high transfer rates and IOPS performance, in combination with high queue depths in random read and write environments, are key capabilities
  - ***c. Content delivery networks*** where high read-intensive performance (typically at 95%R/5%W) are key benefits
  - ***d. Databases*** where low latency and high transactions per minute are key features
  - ***e. Media streaming*** where high read bandwidth can move content into several systems quickly for streaming to many subscribers simultaneously is a key benefit

KIOXIA was first to publicly demonstrate PCIe 4.0-based NVMe SSDs at Flash Memory Summit 2019, and included the enterprise-class CM6 Series and the data center-class CD6 Series. Both series' are developed using 96-layer BiCS FLASH™ 3D flash memory technology, are NVMe 1.4-compliant, and support 2.5-inch form factors. They offer capacities ranging from 960 GB to 30,720 GB (1 DWPD) for read-intensive applications and 800 GB to 12,800 GB (3 DWPD) for mixed-use applications.

Supports Cloud Data Center Architectures based on NVMe-oF

PCIe 4.0 NVMe SSDs can be accessed at lower latencies than other SSD interfaces for host sharing over a network through the NVMe over Fabrics (NVMe-oF™) specification and associated cloud storage volume management software, such as KumoScale™ from KIOXIA. NVMe-oF is quickly becoming the network protocol of choice for cloud architectures given its ability to disaggregate and pool compute, storage and network resources independently. It can provision the right amount of resources for each application workload at low latency and at high bandwidth / IOPS performance as if the PCIe 4.0 NVMe SSDs were locally attached.

Notes:

[1] The SFF-TA-1001 Universal x4 Link Definition specification for SFF-8639 defines the links between the drive and backplane connectors so that PCIe, SAS and SATA interfaces, and their respective protocols, can be supported from one interface slot. The specification is available at: http://www.snia.org/sff/specifications.

[2] The SFF-TA-1005 Universal Backplane Management (UBM) specification defines the universal management that enables SAS/SATA SSDs/HDDs and PCIe NVMe SSDs to be mixed and matched within one UBM-enabled backplane. The specification is available at: http://www.snia.org/sff/specifications.

[3] An Online Transaction Processing (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 KIOXIA CM6 Series and CM5 Series of enterprise-class NVMe SSD products, tested at 3 DWPD (Drive Writes per Day) in a mixed use environment, and configured with 3,200GB capacities.

[4] Definition of capacity: Definition of capacity - KIOXIA Corporation defines a megabyte (MB) as 1,000,000 bytes, 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 1 Gbit = 2^{30} bits = 1,073,741,824 bits, 1 GB = 2^{30} bytes = 1,073,741,824 bytes and 1 TB = 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, and/or pre-installed software applications, or media content. Actual formatted capacity may vary.

[5] Drive Write(s) per Day: One full drive write per day means the drive can be written and re-written to full capacity once a day, every day, under the specified workload for the specified lifetime. Actual results may vary due to system configuration, usage, and other factors.