KIOXIA

Frequently Asked Questions

What You Need To Know About PCIe® 4.0 NVMe™ SSDs

What is PCIe 4.0?

PCI (Peripheral Component Interconnect) Express is a high-speed serial computer expansion bus standard that enables a host CPU to communicate with external devices through a direct PCIe connection to the CPU, or through an expansion card. The PCI-SIG standards group releases new versions of the PCIe specification, and when they do, the bandwidth is typically doubled from the previous release. The latest PCIe 4.0 revision can move data at approximately 2 gigabytes per second (GB/s) per lane versus the PCIe 3.0 revision which moves data at almost 1 GB/s per lane (Table 1), doubling performance:

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)**
1.0	2003	2.5	0.25	0.5	1.0	2.0	4.0
2.0	2007	5.0	0.5	1.0	2.0	4.0	8.0
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
5.0 (approved)	2019	32.0	3.938	7.877	15.75	31.51	63.02
6.0 (planned)	2021	64.0	7.877	15.754	31.51	63.02	126.03

^GT/s = gigatransfers per second

Table 1: PCIe revisions and respective performance capabilities (Source: PCI-SIG)

What does PCIe 4.0 mean for NVMe SSDs?

Performance is the overwhelming benefit of PCIe 4.0 relating to NVMe SSDs. The PCIe 4.0 standard is designed to increase data transfer speeds from 8 GT/s (PCIe 3.0) to 16 GT/s (2 GB/s per lane), and with faster speed for data to move, enables SSDs/GPUs/NICs and other devices to deliver faster I/O performance than previous PCIe revisions.

Another important aspect of these SSDs is the relationship between the PCIe 4.0 interface and the NVMe protocol. SSDs connect to a host in the data center through the PCIe interface that complies with a set of instructions known as the NVMe protocol. The interface is defined by the physical pin definitions and power requirements associated with the hardware transportation layer, while the protocol is defined by commands, instructions and drivers that SSDs and operating systems must follow. The NVMe protocol is defined by NVM Express™, Inc., a non-profit corporation comprised of industry companies working together to create and evolve the specifications. The current version of NVMe is 1.4. The interface and protocol are separate developments driven by their respective standards groups, and evolve independently of each other (Table 2). This separation between church and state enables focused development efforts and innovative capabilities that will progress NVMe SSDs forward.

SSD Interface	SSD Protocol
SATA	ATA/ATAPI/AHCI
SAS	SCSI
PCIe	NVMe

Table 2: SSD interfaces and associated protocol

The selection of which interface to use may be dependent on bandwidth (the amount of data transferred in a period of time), latency (the delay before a data transfer begins following an instruction), and scalability (the ability to adapt to growing workload requirements).

What is the adoption rate of NVMe SSDs in data centers?

Enterprise and data center NVMe SSDs are projected to represent majority use across data centers worldwide when compared to SATA and SAS SSDs. The combined segment (in units) is expected to grow from 42.5% in 2019, to 75% by the end of 2021, and 91% by the end of 2023[1].

Why are there different classes of NVMe SSDs for data centers?

NVMe SSDs are currently segmented into two categories: (1) enterprise NVMe SSDs and (2) data center NVMe SSDs. When market research firms report market data, the NVMe SSD segment typically combines both enterprise and data center categories into one all-encompassing segment. However, enterprise SSDs and data center SSDs have very distinct differences (Table 3):

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

Data center NVMe SSDs are designed for scale out and hyperscale environments where high read performance, Quality of Service (QoS) and power efficiency are key metrics. These drives are optimized for read-intensive applications and typically targeted at cloud workloads. Customers expect to pay less for these SSDs than their more fully-featured enterprise counterparts.

Characteristic	Enterprise	Data Center
Applications	Server data storage; All-flash arrays; Cache layer	Server data storage; Cloud data storage; Boot
Power	12W – 25W	4W – 12W
Performance	700K - 1M IOPS	10K – 700K IOPS
Form Factors	2.5-inch (z=15mm);	2.5-inch (z=15mm);
Spec Conformance	U.3 (SFF-TA-1001) U.2 (SFF-8639 Module)	U.3 (SFF-TA-1001) U.2 (SFF-8639 Module)
Features	Dual-port; Up to 30.7 TB[2] capacity; Data protection	Single-port; Up to 15.3 TB capacity; High QoS

Table 3: Enterprise vs Data Center SSDs (source: KIOXIA America)

What applications can benefit from the different classes of NVMe SSDs?

Enterprise-class NVMe SSDs are targeted to high-performance server applications, of which, databases, data analytics and compute-side artificial intelligence (AI)/ machine learning (ML) are good examples:

- Database applications benefit from low latency, dual-ports for multi-path and high-availability, and the highest available transactions per second.
- Data analytics searches benefit from high sequential and random read bandwidth.
- Compute-side AI and ML benefit from the ability to quickly feed data into DRAM and GPUs during the staging phase at very fast data transfers and perform near-immediate validation of data for corruption during the checkpoint phase.

Data center-class NVMe SSDs are targeted for general purpose server applications and scale out environments, of which, cloud compute, containers, content delivery networks (CDNs), databases and media streaming are good examples:

- Cloud compute applications benefit from high-performance and low-latency especially when compared to SATA deployments.
- Container orchestration benefit from high transfer rates and high queue depths in random read and write environments.
- Content delivery networks benefit from SSDs that excel at read-intensive workloads (typically 90% read and 10% write) prevalent in these applications.
- Database applications benefit from low-latency and high transactions per minute.
- Media streaming benefits from very high read bandwidth. Moving content into several systems as fast as possible so it can be streamed to many subscribers simultaneously is also a key benefit of these SSDs.

What emerging applications can benefit from PCIe 4.0 NVMe SSDs?

The ability to disaggregate and pool compute, storage and network resources independently, and provision the right amount of resources for each application workload, has become an emerging cloud application. This requires moving from a direct-attached storage architecture to a disaggregated shared storage model where NVMe over Fabrics (NVMe-oF™) has quickly become the network protocol of choice for cloud architectures. It provides several benefits in resource utilization while delivering nearly the same high-performance and low-latency benefits as if the NVMe SSDs were locally attached. When cloud storage volume management software operates with the NVMe-oF specification, NVMe SSDs can be pooled and accessed at much lower latencies for host sharing versus other SSD interfaces.

A second emerging application combines SAS, SATA and PCIe interfaces into a backplane managed by a Universal Backplane Management (UBM)-compliant system (the SFF-TA-1005 specification[3]). This enables SAS/SATA SSDs/HDDs and NVMe SSDs to be mixed and matched within one UBM-enabled backplane. The SFF-TA-1001 specification[4] (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 reduce TCO and storage deployment complexities. It also provides a viable replacement path between SATA, SAS and NVMe storage media while maintaining backwards compatibility with U.2 NVMe-based platforms. All of these benefits will help to increase PCIe 4.0 NVMe SSD adoption in servers.

Does KIOXIA offer PCIe 4.0-based SSDs?

KIOXIA was the first to publicly demonstrate PCIe 4.0 NVMe SSDs at Flash Memory Summit 2019 (with Broadcom® Inc.). The product demonstrations included the company's enterprise-class CM6 Series and data center CD6 Series SSDs. Both series' use KIOXIA's 96-layer BiCS FLASH™ 3D flash memory technology, are compliant with the NVMe 1.4 specification, and support 2.5-inch form factors. They offer capacities ranging from 960 GB[2] to 30,720 GB (@1 DWPD[5]) and 800 GB to 12,800 GB (@3 DWPD). Each series is backwards-compatible with PCIe 3.0 environments and support Sanitize Instant Erase[6] (SIE) and Self-Encrypting Drive (SED) with TCG-Opal/Ruby encryption[7] security options[8], and compliant with FIPS 140-2 (Level-2) support[9].

The CM6-V Series features the following performance capabilities, tested at 3 DWPD endurance in a mixed-use environment (Table 4):

CM6-V Series:

SPECIFICATION	800 GB	1,600 GB	3,200 GB	6,400 GB	12,800 GB
Sequential Read *(128 KB; QD=32)*	6,900 MB/s	6,900 MB/s	6,900 MB/s	6,900 MB/s	6,900 MB/s
Sequential Write *(128 KB; QD=32)*	1,400 MB/s	2,800 MB/s	4,200 MB/s	4,000 MB/s	4,000 MB/s
Random Read *(4 KB; QD=256)*	880K IOPS	1,300K IOPS	1,400K IOPS	1,400K IOPS	1,400K IOPS
Random Write *(4 KB; QD=32)*	100K IOPS	215K IOPS	350K IOPS	325K IOPS	330K IOPS
70%R/30%W *(4 KB; QD=256)*	295K IOPS	525K IOPS	750K IOPS	710K IOPS	710K IOPS
Read Latency *(@ QD=1)*	90 µs	90 µs	90 µs	90 µs	90 µs
Write Latency *(@ QD=1)*	10 µs	10 µs	10 µs	10 µs	10 µs

Table 4: KIOXIA CM6-V Series enterprise-class PCIe 4.0 SSD performance[10]

The CD6-R Series features the following performance capabilities, tested at 1 DWPD endurance in a read-intensive environment (Table 5):

CD6-R Series:

SPECIFICATION	960 GB	1,920 GB	3,840 GB	7,680 GB	15,360 GB
Sequential Read *(128 KB; QD=32)*	5,800 MB/s	5,800 MB/s	6,200 MB/s	6,200 MB/s	5,500 MB/s
Sequential Write *(128 KB; QD=32)*	1,300 MB/s	1,150 MB/s	2,350 MB/s	4,000 MB/s	4,000 MB/s
Random Read *(4 KB; QD=256)*	700K IOPS	700K IOPS	1,000K IOPS	1,000K IOPS	750K IOPS
Random Write *(4 KB; QD=32)*	30K IOPS	30K IOPS	60K IOPS	85K IOPS	30K IOPS
70%R/30%W *(4 KB; QD=256)*	100K IOPS	100K IOPS	180K IOPS	270K IOPS	100K IOPS
Read Latency *(@ QD=1)*	90 µs	90 µs	90 µs	100 µs	120 µs
Write Latency *(@ QD=1)*	35 µs	35 µs	35 µs	35 µs	35 µs

Table 5: KIOXIA CD6 Series data center-class PCIe 4.0 SSD performance[11]

The main features of CM6 Series enterprise-class SSDs and CD6 Series data center-class SSDs:

Feature	CM6 Series	CD6 Series
Interface Support	PCIe 4.0 1x4 or 2x2	PCIe 4.0 1x4
Capacities	960 GB to 30,720 GB (1 DWPD) 800 GB to 12,800 GB (3 DWPD)	960 GB to 15,360 GB (1 DWPD) 800 GB to 12,800 GB (3 DWPD)
Controller	18-channel	16-channel
Configurable Power	9, 11, 14, 16, 18, 25 watts	9, 11, 14, 16, 18, 25 watts
Read/Write Latency	90 µR/10 µW	90 µR/10 µW
Single Port	Yes (1x4)	Yes (1x4)
Dual Port	Yes (2x2)	N/A
Namespaces	64	16
Sector Size	512/520/4096/4104/4160	512/4096

The following tables compare CM6/CD6 Series SSDs to the previous CM5/CD5 Series generation based on PCIe 3.0 (Table 6 and Table 7). The comparisons include supported capacity ranges.

CM6-V Series vs CM5-V Series[12]:

		CM5-V SERIES SSDs				CM6-V SERIES GAINS (using Table 4)				
SPECIFICATION	**Units**	**800 GB**	**1,600 GB**	**3,200 GB**	**6,400 GB**	**800 GB**	**1,600 GB**	**3,200 GB**	**6,400 GB**	**12,800 GB**
SR *(128 KB; QD=32)*	MB/s	3,250	3,250	3,350	3,350	+112%	+112%	+105%	+105%	N/A
SW *(128 KB; QD=32)*	MB/s	1,250	2,450	3,040	3,040	+12%	+14%	+38%	+31%	N/A
RR *(4 KB; QD=256)*	KIOPS	370	650	750	770	+137%	+100%	+86%	+81%	N/A
RW *(4 KB; QD=32)*	KIOPS	95	145	160	165	+5%	+48%	+118%	+96%	N/A
70%R/30%W *(4 KB; QD=256)*	KIOPS	235	310	360	380	+25%	+69%	+108%	+86%	N/A
Read Latency *(@ QD=1)*	µs	110	110	110	110	-18%	-18%	-18%	-18%	N/A
Write Latency *(@ QD=1)*	µs	30	30	30	30	-66%	-66%	-66%	-66%	N/A

Table 6: KIOXIA CM5 Series versus CM6 Series enterprise-class SSD performance (For Read and Write Latency, the smaller the number, the better the performance)

CD5-R Series vs CD6-R Series[13]:

		CD5-R SERIES SSDs				CD6-R SERIES GAINS (using Table 5)				
SPECIFICATION	**Units**	**960 GB**	**1,920 GB**	**3,840 GB**	**7,680 GB**	**960 GB**	**1,920 GB**	**3,840 GB**	**7,680 GB**	**15,360 GB**
SR *(128 KB; QD=32)*	MB/s	3,140	3,140	3,140	3,140	+84%	+84%	+97%	+97%	N/A
SW *(128 KB; QD=32)*	MB/s	880	780	1,520	1,980	+47%	+47%	+54%	+102%	N/A
RR *(4 KB; QD=256)*	KIOPS	305	270	465	550	+129%	+159%	+115%	+81%	N/A
RW *(4 KB; QD=32)*	KIOPS	20	20	40	50	+50%	+50%	+50%	+70%	N/A
70%R/30%W *(4 KB; QD=256)*	KIOPS	50	65	115	115	+100%	+53%	+56%	+134%	N/A
Read Latency *(@ QD=1)*	µs	110	120	120	135	-18%	-25%	-25%	-25%	N/A
Write Latency *(@ QD=1)*	µs	50	50	30	30	-30%	-30%	+16%	+16%	N/A

**For Random Read, the maximum queue depth for the CD5 Series is 128 – the CD6 Series is 256*

Table 7: KIOXIA CD5 Series versus CD6 Series data center-class SSD performance (For Read and Write Latency, the smaller the number, the better the performance)

The CM6 Series and CD6 Series deliver significant performance gains over the previous generation CM5 Series and CD5 Series, and reflect the performance increases associated with the PCIe 4.0 interface doubling in both transfer rate and data rate over the previous PCIe 3.0 revision.

When will KIOXIA-branded PCIe 4.0-based solutions become available in the market?

KIOXIA CM6 Series enterprise-class SSDs and CD6 Series data center-class SSDs are sampling as of February 2020.

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 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 = 230 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.

[3] The SFF-TA-1005 Universal Backplane Management (UBM) specification is available at: http://www.snia.org/sff/specifications.

[4] The SFF-TA-1001 Universal x4 Link Definition specification for SFF-8639 is available at: http://www.snia.org/sff/specifications.

[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.

[6] The Sanitize Instant Erase (SIE), Self-Encrypting Drive (SED), FIPS (Federal Information Processing Standards) optional models are available. SIE option supports Crypto Erase, which is a standardized feature defined by NVM Express Inc.

[7] SED supports TCG Opal and Ruby SSCs. It has a few unsupported TCG Opal features. For more details, please make inquiries through "Contact us" in each region's website, https://business.kioxia.com/

[8] Optional security feature compliant drives are not available in all countries due to export and local regulations.

[9] FIPS drives are designed to comply with FIPS 140-2 Level 2, which define security requirements for cryptographic module by NIST (National Institute of Standards and Technology). For the latest validation status of each model, please contact us in each region's website, https://business.kioxia.com/

[10] 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 of enterprise-class NVMe SSD products, tested at 3 DWPD (Drive Writes per Day) in a mixed-use environment, and configured with all supported capacities.

[11] 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 KIOXIA CD6 Series of data center NVMe SSD products, tested at 1 DWPD in a read-intensive environment, and configured with all supported capacities.

[12] CM5 Series performance specifications available from KIOXIA America, Inc. published marketing assets.

[13] CD5 Series performance specifications available from KIOXIA America, Inc. published marketing assets.