KIOXIA

Application Brief

MySQL Performance Gains with PCIe® 4.0 vs. PCIe 3.0 SSDs

The current PCIe 4.0 interface is now available in servers that enables twice the bandwidth versus PCIe 3.0 so peripheral devices such as SSDs, GPUs and NICs can access data faster than ever before. The speed upgrade is well-suited for data-intensive and computational applications such as cloud computing, databases, data analytics, artificial intelligence, machine learning, container orchestration, media streaming, to name a few, and enables today's fast CPUs to be continually fed with data.

To validate application performance and productivity gains that can be achieved with the PCIe 4.0 interface, KIOXIA Corporation, a leader in the PCIe 4.0 SSDs for server and storage segments, compares the speed advantages of a database application running on the same server with PCIe 3.0 SSD technology. A MySQL relational database management system (RDBMS) was used for this comparison as it is extremely versatile in both hyperscale and enterprise environments. It supports key applications such as webserver, online transactional processing (OLTP), e-commerce and data warehousing, and is the most widely deployed open source database in the world (ranked number two overall[1]).

Performance tests were performed between PCIe 4.0 and PCIe 3.0 SSDs that measured transactions per minute (TPM) and average read/write latency of a mainstream AMD QCT S5K server platform running MySQL v8.0. A KIOXIA CD6-R Series PCIe 4.0 NVMe™ data center SSD and a leading and currently available PCIe 3.0 data center SSD represented the storage media used in the comparison. A description of the benchmark tests (criteria, set-up and procedures), and the test results (visual representation and analysis), are presented in this application brief.

PCIe 4.0 Performance Overview

PCIe 4.0 devices, such as SSDs, utilize multiple PCIe lanes simultaneously to achieve increased performance. In comparison to PCIe 3.0, PCIe 4.0 delivers an increase in data transfer speed from 8 gigatransfers per second (GT/s) to 16 GT/s, and can move data at approximately 2 gigabytes per second (GB/s) per lane (versus about 1GB/s per lane). The end result is that PCIe 4.0 can double performance, delivering a 4-lane (x4) bandwidth of approximately 8GB/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

Table 1: PCIe 4.0 can double GT/s and by-lane performance versus PCIe 3.0 (Source: PCI-SIG)

In enterprise applications, PCIe 4.0 is well suited for workloads that need the utmost data storage performance and require robust reliability features for 24x7 operation. Enterprise SSDs typically include high-end features such as dual-port capabilities, larger capacities, high endurances and data protection schemes. For data center requirements, PCIe 4.0 is well suited for scale-out and hyperscale environments where SSDs are expected to deliver high read-intensive performance in support of cloud compute workloads. Data center SSDs are less expensive than fully-featured enterprise SSDs.

For the purposes of this application brief, benchmark tests were conducted by KIOXIA in a lab environment. The test results demonstrate the TPM performance and average combined latency that can be achieved when performing queries against the MySQL database with an AMD EPYC based PCIe 4.0 server, compared to the same server with PCIe 3.0 SSDs.

Description of Benchmark

Benchmark tests were conducted that compared TPM performance and transactional read/write latency of an AMD QCT S5K server platform. The storage configuration featured KIOXIA's CD6 Series PCIe 4.0 NVMe SSDs and the latest and currently shipping PCIe 3.0 NVMe SSD model from a leading vendor (referred to as Vendor A).

KIOXIA was the first storage vendor to deliver U.3 conformant enterprise and data center SSDs based on the PCIe 4.0 interface and NVMe specification, with their CM6 Series and CD6 Series SSDs, respectively.

The tests utilized an OLTP MySQL database workload based on comparable TPC-C™ benchmarks created by HammerDB[2] software.

Test Criteria

The hardware and software equipment used for these benchmark tests included:

- **AMD Based QCT S5K Server:**
  One (1) dual socket server with two (2) AMD EPYC™ 7702P processors, featuring a total of 128 processing cores, 2.0 GHz frequency, and 128 gigabytes[3] (GB) of DDR4
- **Operating System:** CentOS™ v8.2
- **Application:** MySQL v8.0
- **Benchmark Software:** Comparable TPC-C benchmark tests generated through HammerDB v3.3 test software
- **Storage Devices (Table 2):**
  One (1) KIOXIA CD6 Series PCIe 4.0 SSD with 7.68 terabyte[3] (TB) capacity
  One (1) Vendor A PCIe 3.0 SSD with 7.68 TB capacity

Specifications	CD6 Series	Vendor A
Interface	PCIe 4.0	PCIe 3.0
Capacity	7.68 TB	7.68 TB
Form Factor	2.5-inch (15mm)	2.5-inch (7mm)
NAND Flash Type	BiCS FLASH™ 3D flash memory (96-layer)	3D TLC
Drive Writes per Day[4] (DWPD)	1.3 (3 yrs)	1.3 (3 yrs)
Data Warehouses[5]	5,000	5,000
Virtual User Count	64	64
InnoDB Buffer Pool Size	64 GB	64 GB

Table 2: SSD specifications and set-up parameters

Set-up & Test Procedures

The test system was configured using the hardware and software equipment outlined above. The MySQL application was configured to store both the log and database on the SSD that was tested. Additionally, HammerDB software was run locally on the server as there was excess CPU available to not cause bottlenecks in performance. It was configured with a test schema based on the TPC-C benchmark (to emulate a MySQL OLTP database environment). The MySQL application was then loaded with 5,000 data warehouses that comprised about 442 GB of the server's storage capacity. The MySQL InnoDB storage engine buffer pool size was set to 64 GB, which corresponds to the amount of memory that is allocated to the MySQL application.

Additionally, the test tool was configured for 64 virtual users to simultaneously send query threads in order to obtain responses. The query response time was also set to one millisecond (ms), demonstrating the ability to achieve very fast responses.

TPM and latency tests were run once for each tested SSD series and the latency results were gathered while the TPM test was ongoing. The objective of the tests was to showcase how AMD based PCIe 4.0 servers provide significantly higher performance and lower latencies for common database applications when compared to PCIe 3.0.

Test Results

TPM and latency benchmarks were conducted. For TPM, the higher the value, the better the result. For read and write latency, the lower the value, the better the result.

Transactions Per Minute

In an OLTP database environment, TPM is a measure of how many transactions in the TPC-C transaction profile that are being executed per minute. The HammerDB software, executing the TPC-C transaction profile, randomly performs new order transactions and randomly executes additional transaction types such as payment, order status, delivery and stock levels. This benchmark simulates an OLTP environment where there are a large number of users that conduct simple, yet short transactions that require sub-second response times and return relatively few records.

The TPM test results are presented (Table 3) and visually represented (Figure 1):

Benchmark Test: MySQL TPC-C Workload	CD6 Series (PCIe 4.0)	Vendor A (PCIe 3.0)
Transactions per Minute	1,064,222	560,216
CD6 Series PCIe 4.0 Advantage	+89.9%	-

Table 3: TPM comparison of data center PCIe 4.0 vs PCIe 3.0 NVMe SSDs (higher is better)

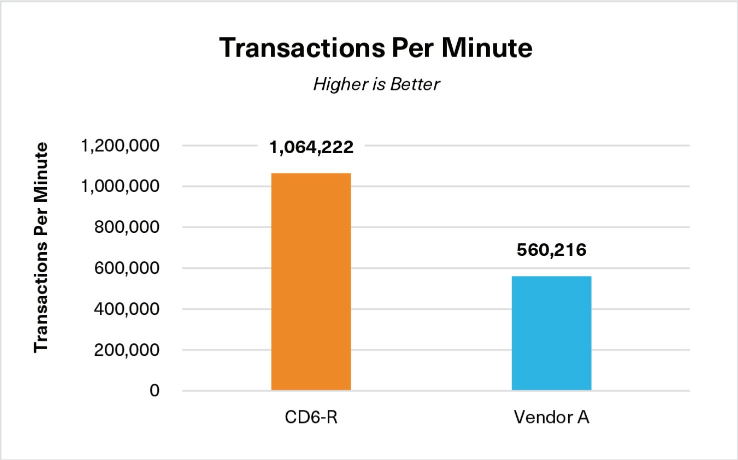

Figure 1: TPM comparison of data center PCIe 4.0 and PCIe 3.0 NVMe SSDs (higher is better)

Average Read and Write Latency

Latency is the delay in time before a storage device completes a data transaction following an instruction from the host for that request, which can greatly affect application performance and the user experience. Application response time often has built-in latency between the user and the server, so maintaining low-latency within the server will usually translate into an overall better user experience.

The read and write latency test results are presented (Table 4) and visually represented (Figure 2):

Benchmark Test: MySQL TPC-C Workload	CD6 Series (PCIe 4.0)	Vendor A (PCIe 3.0)
Average Read Latency (ms)	0.16	0.48
CD6 Series PCIe 4.0 Advantage	-66%	-
Average Write Latency (ms)	0.73	7.94
CD6 Series PCIe 4.0 Advantage	-90%	-

Table 4: Latency comparison of data center PCIe 4.0 vs PCIe 3.0 NVMe SSDs (lower is better)

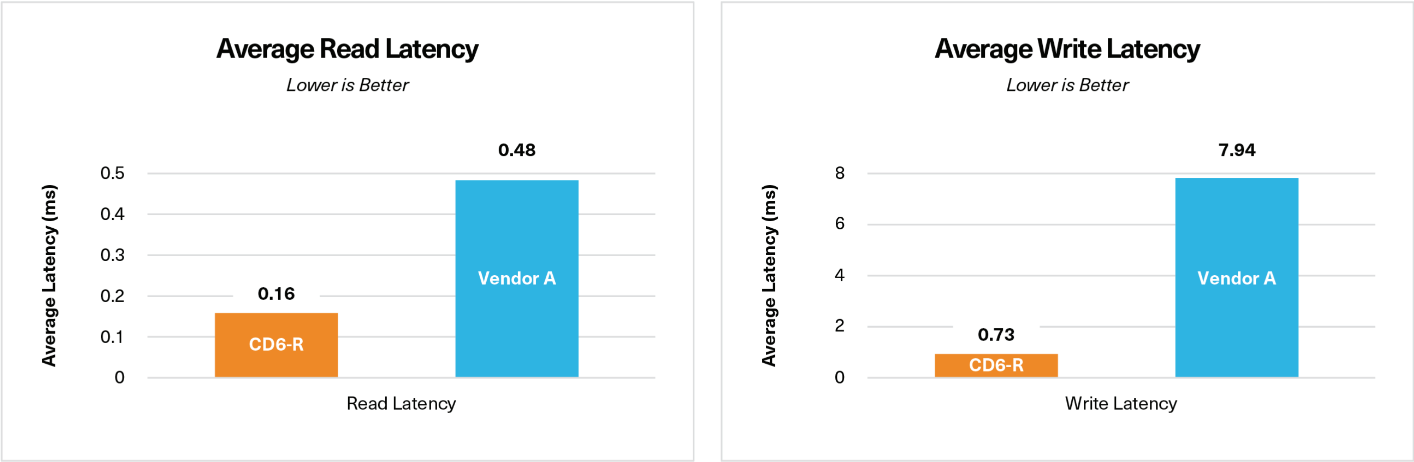

Figure 2: Latency comparison of data center PCIe 4.0 vs PCIe 3.0 NVMe SSDs (lower is better)

Test Analysis

From the benchmark test results, a data center PCIe 4.0-based SSD enabled the server to deliver almost 90% higher transactions per minute when compared to the tested PCIe 3.0-based SSD from a leading vendor. The PCIe 4.0-based CD6 Series TPM performance enabled the AMD EPYC 7702P based S5K server platform to support significantly more transactions, all while servicing transactions with over 66% better read latency.

Better storage performance also has a domino effect on overall system performance, to the point of reducing DRAM requirements or other factors. To demonstrate the advantages of PCIe 4.0 technology regarding memory allocation, additional tests were conducted, but limited the MySQL application to 32 GB of available

DRAM (versus 64 GB as set in the original performance test). For this DRAM-specific test, the PCIe 4.0-based server with a CD6 Series SSD delivered 662,975 transactions per minute, while the same server with a PCIe 3.0-based drive from Vendor A delivered a TPM of 337,419. This resulted in a 96% performance advantage for the server configured with PCIe 4.0 SSD storage.

The PCIe 4.0-based CD6 Series SSD result of 662,975 TPM, when under a DRAM constraint of 32 GB, outperformed the PCIe 3.0-based Vendor A system configured with 64 GB of DRAM, delivering 560,216 TPM. Even with half of the DRAM allocated, the CD6 Series with 32 GB of DRAM delivered over 18% better TPM when compared to Vendor A's PCIe 3.0 performance with 64 GB of DRAM allocated (Figure 3). This result demonstrated that increased SSD performance impacts system performance regardless of a scale-up or a scale-down.

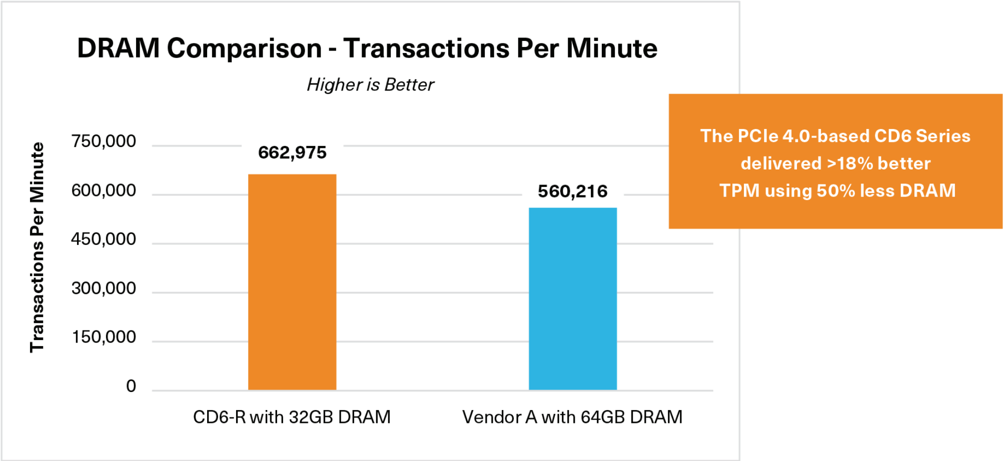

Figure 3: TPM comparison of data center PCIe 4.0 vs PCIe 3.0 NVMe SSDs using 50% less DRAM (higher is better)

It should also be noted that the AMD EPYC based QCT S5K server processor settings were changed from a default Nodes per Socket5 (NPS) that equaled '4' to an NPS setting that equaled '1.' The NPS settings impact how memory is allocated to the CPU where an 'NPS=1' setting enables interleaving of all eight (8) channels for each socket, and creates a single NUMA node per socket. This change from 'NPS=4' to 'NPS=1' increases per-core memory bandwidth (at the expense of memory latency), which in turn, increases the CPU's ability to move data from memory to the SSD across the PCIe bus. Since all memory channels are available to all cores, the CPU is able to process transactions to disk faster, resulting in overall higher performance from both the system and the KIOXIA CD6-R SSD.

With this change, CD6-R Series performance improved from 993,593 transactions per minute ('NPS=4') to 1,064,222 transactions per minute ('NPS=1'), and an over 7% improvement in TPM performance (Figure 4).

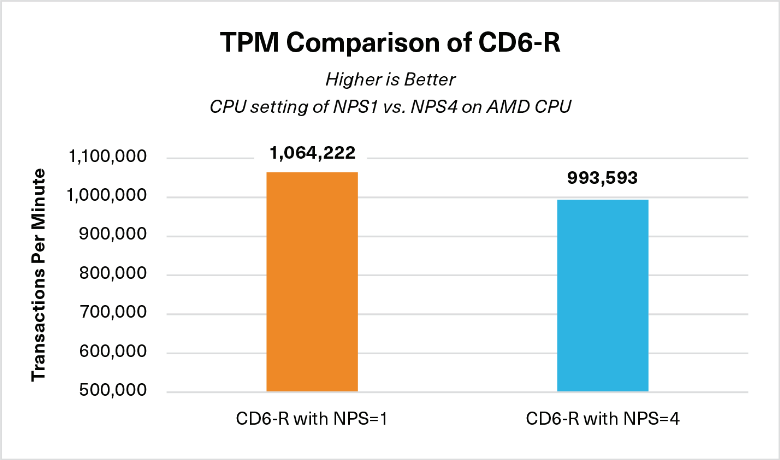

Figure 4: TPM comparison of CD6 Series PCIe 4.0 NVMe SSDs with the CPU setting at '1' and '4' (higher is better)

Summary

The testing demonstrated that a mainstream AMD EPYC based PCIe 4.0 server with less DRAM can outperform a PCIe 3.0 server with larger DRAM, and can have a dramatic impact on server sizing. If extra performance is needed, then PCIe 4.0-based SSDs can be utilized to get greater storage throughput out of the server. However, if cost is the primary factor, then purchasing PCIe 4.0 servers with reduced DRAM (and PCIe 4.0 SSDs deployed) can still outperform a more expensive server configurations that have additional DRAM and are configured with PCIe 3.0 SSDs.

Driven by the PCIe 4.0 interface and NVMe specification, the combination of the AMD based S5K server with EPYC 7702p processors and the KIOXIA CD6 Series SSD deliver more storage bandwidth and input/output operations per second (IOPS) than the same server with a PCIe 3.0 SSD. As it relates to MySQL applications, PCIe 4.0 enables more database transactions to be completed and more users to be serviced simultaneously.

The benefits of significantly improved application performance and cost-efficiencies make it an easy decision to adopt PCIe 4.0 technology. Additional CD6 Series PCIe 4.0 NVMe SSD information is available here.

NOTES:

[1] Source: https://db-engines.com/en/ranking, January 2021.

[2] HammerDB is benchmarking and load testing software that is used to test popular databases. It simulates the stored workloads of multiple virtual users against specific databases to identify transactional scenarios and derive meaningful information about the data environment, such as performance comparisons. TPC Benchmark C is a supported OLTP benchmark that includes a mix of five concurrent transactions of different types, and nine types of tables with a wide range of record and population sizes and where results are measured in transactions per minute.

[3] 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 1Gbit = 2^{30} bits = 1,073,741,824 bits, 1GB = 2^{30} bytes = 1,073,741,824 bytes and 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, and/or pre-installed software applications, or media content. Actual formatted capacity may vary.

[4] 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, for the specified lifetime. Actual results may vary due to system configuration, usage, and other factors.

[5] The 5,000 data warehouses, in combination with the 64G buffer pool size, represents a real-world database configuration that creates a normal database size. In the TPC-C benchmarks created by HammerDB, 5,000 data warehouses equates to a database size of 442 GB, with the ability to cache 14.48% of the database with 64 GB of RAM.

[6] At a Nodes per Socket (NPS) setting of 4, and 2 CPUs, the system regards this set-up as 8 logical CPUs, with multiple cores, and assigns 1/8 of memory to each logical node. The memory split means that less processing cores can access each segment of memory, however, memory access within that segment is faster. When the Nodes per Socket (NPS) setting was changed from default ‘4’ to a ‘1’ setting, TPM performance increased.

KIOXIA