KIOXIA

Application Brief

Using KIOXIA EDSFF E1.S SSDs and RocksDB to Improve Database Engine Performance and Latency

Featuring KIOXIA XD7P Series Data Center E1.S SSDs Deployed in an E1.S Server

Business applications utilize databases to store various types of data generated from on-premises server clusters or the cloud. They require rapid responses to queries and very high storage performance from the database workloads at the backend. A database management system (DBMS) manages a database utilizing a database (storage) engine that enables interaction through create, read, update and delete (CRUD) operations. Common systems used by IT professionals today include MySQL®, PostgreSQL®, MongoDB® and Apache® Cassandra®.

Some database engines may not be optimized for the underlying storage hardware (such as SSDs) which can lead to performance degradation of those database queries and workloads. RocksDB is a database engine optimized for high read/write throughput performance and low latency storage. It features a log structure that accommodates a variety of production environments that require SSDs.

The underlying storage hardware plays an important role for meeting the performance requirements of data heavy applications. KIOXIA XD7P Series Data Center NVMe™ E1.S SSDs provide high read/write performance, low latency access to large data volumes and a rich feature set. They also support new Enterprise and Datacenter Standard Form Factor[1] (EDSFF) E1.S form factors that can address the data center objective for high storage performance and increased storage density.

This application brief presents a comparison of RocksDB database engine performance and latency tested in a lab environment supporting two SSD configurations. One configuration included an E1.S server deployed with four 7.68 terabyte[2] (TB) KIOXIA XD7P Series E1.S SSDs. The other Vendor A configuration used four similar capacity E1.S SSDs deployed in the same server.

The db_bench[3] test tool was used to benchmark RocksDB performance and latency in both SSD configurations running multi-random read tests. The database tests included throughput and read latency at the 99th, 99.9th and 99.99th percentiles. Additional drive metrics included aggregate read throughput, IOPS (input/output operations per second) and average read latency.

Test results show that the KIOXIA XD7P Series E1.S SSDs delivered higher throughput and lower latency of the RocksDB database engine for the database and drive tests when compared with Vendor A's E1.S SSDs in an E1.S server environment. Of particular interest, over 35 million database operations per second were performed by the RocksDB database engine and KIOXIA XD7P Series SSD configuration.

The test results include a brief description of each workload test, a graphical depiction of the test results and a test analysis. Appendix A covers the hardware and software configuration – Appendix B covers the configuration set-up and test procedures.

Test Results Snapshot

The KIOXIA XD7P Series E1.S SSDs improved RocksDB database engine performance when compared with Vendor A's E1.S SSDs per these test results:

Average Drive Read Latency

33% Lower

Database Read Latency:

99.99th Percentile

45% Lower

99.9th Percentile

17% Lower

99th Percentile

8% Lower

Database Throughput:

8% Higher

Aggregate Drive Read Throughput:

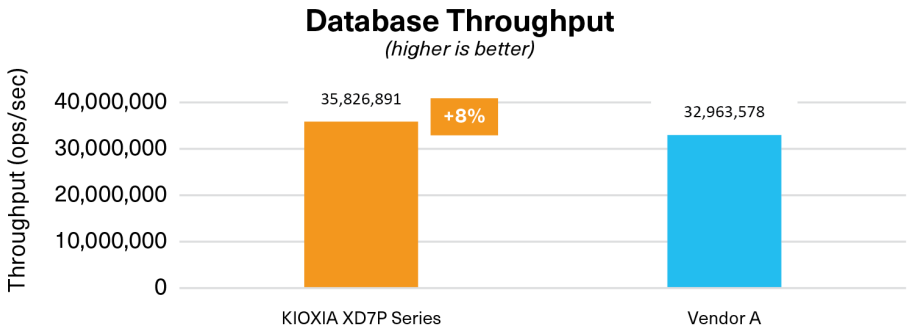
7% Higher

Aggregate Drive Read IOPS:

7% Higher

Test Results[4]

Database Throughput *(in operations per second - ops/sec)*
This test represents the number of operations per second that the RocksDB database engine completed on average. If DBMSs and database engines are able to complete more operations in a given timeframe, high database throughput can be achieved. By completing database operations in bulk, both concurrently and quickly, a DBMS is able to service multiple users at the same time, which in turn contributes to high application performance and a good user experience. The higher result is better.



Analysis:

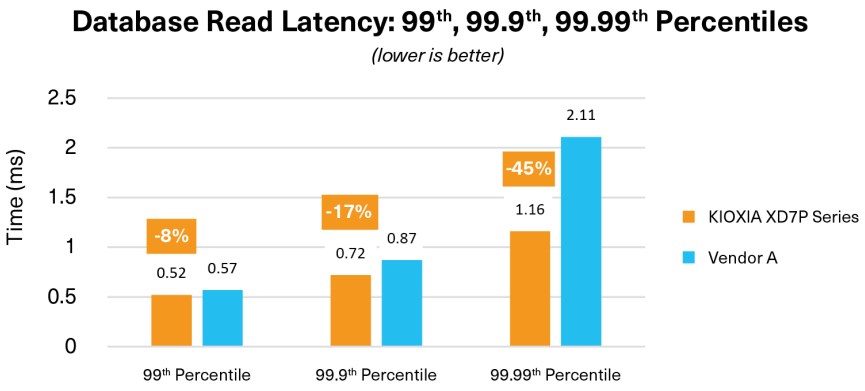
The KIOXIA XD7P Series E1.S SSDs demonstrated 8% higher throughput when compared with Vendor A's E1.S SSDs.

Database Read Latency *(in milliseconds - ms)*

99th Percentile:
This test represents the time it took to perform a read database operation at the 99th percentile, or when 99% of the total number of operations within the test were able to be completed at or less than the reported value. The lower result is better.

99.9th Percentile:
This test represents the time it took to perform a read database operation at the 99.9th percentile, or when 99.9% of the total number of operations within the test were able to be completed at or less than the reported value. The lower result is better.

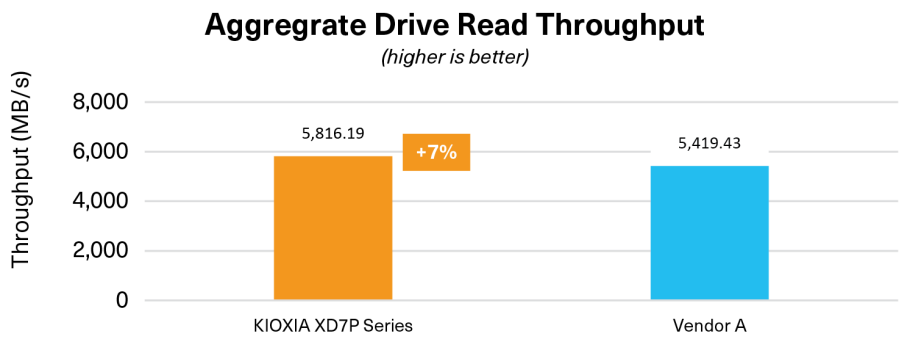
99.99th Percentile:
This test represents the time it took to perform a read database operation at the 99.99th percentile, or when 99.99% of the total number of operations within the test were able to be completed at or less than the reported value. The lower result is better.



Analysis:

The KIOXIA XD7P Series E1.S SSDs demonstrated lower database read latency at the 99th, 99.9th and 99.99th percentiles when compared with Vendor A's E1.S SSDs.

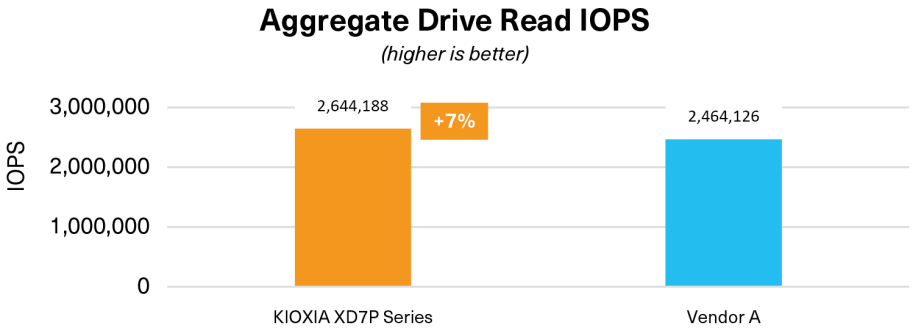
Aggregate Drive Read Throughput *(in megabytes per second - MB/s)*
This test represents the total throughput achieved from all of the underlying SSDs in both configurations while the workload was running. The higher result is better.



Analysis:

The KIOXIA XD7P Series E1.S SSDs demonstrated 7% higher read throughput from all drives when compared with Vendor A's E1.S SSDs.

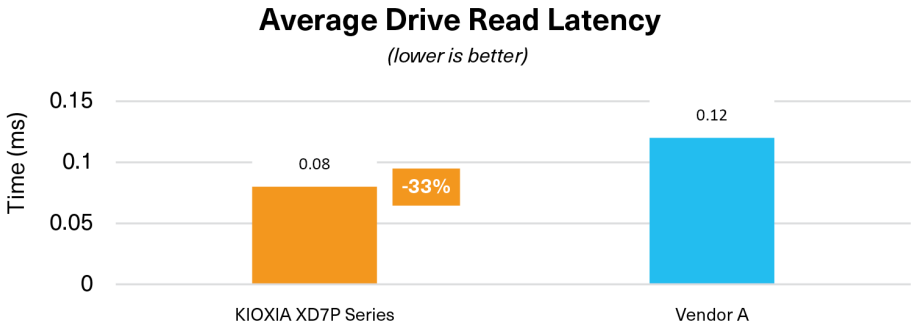
Aggregate Drive Read IOPS *(in input/output operations per second - IOPS)*
This test represents the total IOPS achieved from all of the underlying SSDs in both configurations while the workload was running. The higher result is better.



Analysis:

The KIOXIA XD7P Series E1.S SSDs demonstrated 7% higher read IOPS from all drives when compared with Vendor A's E1.S SSDs.

Average Drive Read Latency *(in milliseconds - ms)*
This test represents the time it took to perform a drive read operation. It includes the average time it took for the read operation to complete once the drive received the read operation request issued from the db_bench workload generator. The lower result is better.



Analysis:

The KIOXIA XD7P Series E1.S SSDs demonstrated 33% lower average read latency from all drives when compared with Vendor A's E1.S SSDs.

Summary

The KIOXIA XD7P Series E1.S SSDs demonstrated higher performance and lower latency with the RocksDB database engine for the drive tests performed by KIOXIA when compared with Vendor A's E1.S SSDs in an E1.S server environment. These drives are well suited for database engines and database applications as presented by the test results.

Database throughput is important with respect to storage density, primarily for determining how many users a single server can support. In other words, higher throughput enables more operations per server with the same storage density. The KIOXIA XD7P Series SSD and RocksDB configuration demonstrated over 35 million database operations per second.

Faster responses from database engines make database applications appear snappier, which is an important measurement in determining the end user experience. As such, database and drive latency performance can be a critical revenue measurement for online transactional and purchase-based websites. With high database throughput and low database read latency, coupled with high drive read throughput/IOPS and low drive read latency, the XD7P Series E1.S SSDs enable more users per server and a better overall end user experience when compared with the E1.S SSDs from Vendor A.

KIOXIA XD7P Series SSD Product Info

The KIOXIA XD Series product line aims to provide better total cost of ownership for customers. The latest generation KIOXIA XD7P Series PCIe® 4.0 data center NVMe SSDs are representative of a new category of Open Compute Project® (OCP) cloud-optimized SSDs and based on the EDSFF E1.S form factor. The series includes 9.5 mm and 15 mm E1.S form factor thickness and supports up to 7.68 TB capacities at 1 Drive Write Per Day[5] (DWPD). The XD7P Series performance specifications[6] include:

KIOXIA XD7P Series SSD7: E1.S 15 mm[7]

SPECIFICATION		Units	1,920 GB (up to)	3,840 GB (up to)	7,680 GB (up to)
Sequential Read	(128 KB; QD=32; 20 W)	MB/s	7,200	7,200	7,200
Sequential Write	(128 KB; QD=32; 20 W)	MB/s	3,100	4,800	4,800
Random Read	(4 KB; QD=256; 20 W)	IOPS	1,500,000	1,650,000	1,550,000
Random Write	(4 KB; QD=32; 20 W)	IOPS	95,000	180,000	200,000

Additional KIOXIA XD7P Series E1.S data center SSD specifications and information is available here.

Appendix A
Database Tests

Server Information	
No. of Servers	1
No. of CPU Sockets	2
CPU	Intel® Xeon® Gold 6444Y
No. of CPU Cores	16
CPU Frequency	3.6 GHz
Total Memory	128 GB DDR5 DRAM
Memory Frequency	DDR5-4800

Operating System Information	
Operating System	Ubuntu®
Version	22.04.3 LTS
Kernel	5.15.0-86-generic

SSD Information		
Model	KIOXIA XD7P Series	Vendor A
Interface	PCIe 4.0 x4	PCIe 4.0 x4
Interface Speed	64 gigatransfers per second (GT/s)	64 GT/s
No. of Devices	4	4
Form Factor	E1.S (15 mm)	E1.S (15 mm)
Capacity	7.68 TB	7.68 TB
DWPD	1 (5 years)	1 (5 years)
Active Power	up to 20 W	up to 12 W

RocksDB Information	
Version	8.8.0
Load Generator	db_bench
No. of Database Engines	4
Cache Size	20,000,000,000 bytes
Block Size	4,096 bytes
No. of CPU Threads	96

Appendix B
Configuration Set-up / Test Procedures

Set-up

An E1.S server was set-up with an Ubuntu® 22.04.3 LTS operating system.

Four KIOXIA XD7P Series E1.S SSDs were installed into the server.

Each drive was setup with an XFS® file system and mounted with the noatime[8], nodiratime[9], and nodiscard[10] flags.

The RocksDB database engine was downloaded into the server in production release mode.

Using db_bench provided with the RocksDB database engine, a database was created on each of the four KIOXIA XD7P Series drives that consisted of one billion keys and a block size of 4,096 bytes.

Procedures

A database test was conducted where a multi-random read workload was run against each of the RocksDB databases simultaneously. This 100% read workload represents a common workload that a database engine could receive, where caching is an example use case.

The results from the following conducted tests were recorded:

- Database Throughput
- Database Read Latency at the 99th Percentile
- Database Read Latency at the 99.9th Percentile
- Database Read Latency at the 99.99th Percentile
- Aggregate Drive Read Throughput
- Aggregate Drive Read IOPS
- Average Drive Read Latency

The same set-up outlined above was configured for the Vendor A E1.S SSDs.

The same tests outlined above were performed for the Vendor A E1.S SSD configuration.

The test results were compared with the KIOXIA XD7P Series E1.S SSD results.

NOTES:

[1] Developed by the Small Form Factor Technical Affiliate (SFF-TA) working group as part of the Storage Networking Industry Association (SNIA).

[2] Definition of capacity - KIOXIA Corporation defines a megabyte (MB) as 1,000,000 bytes, a gigabyte (GB) as 1,000,000,000 bytes, a terabyte (TB) as 1,000,000,000,000 bytes and a petabyte (PB) as 1,000,000,000,000,000 bytes. A computer operating system, however, reports storage capacity using powers of 2 for the definition of 1Gbit = 230 bits = 1,073,741,824 bits, 1GB = 230 bytes = 1,073,741,824 bytes, 1TB = 240 bytes = 1,099,511,627,776 bytes and 1PB = 240 bytes = 1,125,899,906,842,624 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] db_bench is the main tool used to benchmark RocksDB performance. RocksDB inherited db_bench from LevelDB, and enhanced it to support many additional options. It supports many benchmarks to generate different types of workloads.

[4] Read and write speed may vary depending on the host device, read and write conditions and file size.

[5] Drive Write(s) per Day (DWPD): 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.

[6] KIOXIA XD7P Series SSD product specifications provided by KIOXIA Corporation and is accurate as of this publication date.

[7] The KIOXIA XD7P Series SSD product image may represent a design model.

[8] The noatime option turns off access time recording so that the file system will ignore access time updates on files. If the file system is used for database workloads, specifying noatime can reduce writes to the file system.

[9] The nodiratime option disables the updating of access times when opening directories – this prevents modifications to the access time when enumerating directories.

[10] The nodiscard option disables the file system from informing the underlying block device to issue a TRIM command when blocks are longer used.

TRADEMARKS:

DISCLAIMERS: