MayaData

intel

Case Study

Lightning-fast storage solutions with
**OpenEBS Mayastor and
Intel® Optane™**

OpenEBS Mayastor is low latency, low overhead open source storage for Kubernetes. The Intel® Optane™ SSD P5800X sets the standard for high-performance Data Center SSDs.

Results

OpenEBS Mayastor, realizes the full performance potential of Intel® Optane™ SSD P5800X drives, for example, delivering 1.424 million IOPS for 100% random reads workloads

Lab and simulation testing showed the underlying OpenEBS Mayastor imparted less than a 10% overhead vs. the raw performance of the PX5800X

OpenEBS Mayastor does not require NVMe drives or cloud volumes though it can fully leverage their capabilities

These benchmarks were performed with the assistance of Kubera Propel from MayaData, open source management software that includes an intuitive interface and APIs and offers built-in benchmarking

MayaData offers 24x7 support for OpenEBS and the use of Kubera Propel

The Data Agility Company

Evan Powell

CEO, MayaData

"

We are extremely appreciative of the work Intel is doing to deliver radically faster, small footprint NVMe connected Optane SSDs to the market. Our benchmarking work with Intel shows users can now achieve the massively high performance potential of Optane SSDs via the very low overhead Kubera Propel and underlying OpenEBS Mayastor."

Intel® Optane™ SSD P5800X

The revolutionary Intel® Optane™ SSD P5800X with next generation Intel Optane storage media and advanced controller delivers "no-compromises" I/O performance - read or write - and high endurance, providing unprecedented value over legacy storage in the accelerating world of intelligent data.

The Intel® Optane™ SSD P5800X is the World's Fastest Data Center SSD and is PCIe gen 4.0 capable, delivering unprecedented performance to a wide range of applications and storage work loads.

Sequential Read	Up to 7.2 GB/s
Sequential Write	Up to 6.2 GB/s
Random 512B Read	Up to 4.6MIOPS
Random 4KB Read	Up to 1.5MIOPS
Random 4KB Write	Up to 1.5MIOPS
Random 4KB Read/Write (70/30)	Up to 1.8MIOPS
QoS (4KRR, QD=1, 99%)	<6us[1]
QoS (4KRR, Read Write Mixed, QD=1, 99.999%)	<25us
Endurance	100 DWPD

Table 1 Key performance and endurance attributes of P5800X

The P5800X, when compared with Intel® Optane™ SSD P4800X (first generation of Intel® Optane™ SSD) delivers significantly higher performance[2]. P5800X delivers up to 3x more IOPS in 4KB mixed random read/write scenario. Up to 40% better QoS for 99% percentile in 4KB random read, queue depth 1 scenario. Up to 3x higher seq bandwidth and 67% higher endurance as measured in Drive Writes Per Day (DWPD). P5800X is available in U.2 form factor and different capacities: 400GB, 800GB, 1.6TB, 3.2TB.

[1] Latency measured using Intel® Storage Performance Developer Kit (Intel SPDK) and FIO 3.3 (with fio plugin) based on random 4KB transfer size with total queue depth of 1 (QD=1 workers/jobs=1) workload.
[2] See Appendix for workloads and configurations. Results may vary.

MayaData Kubera Propel

OpenEBS Mayastor is designed to enable extremely low latency performance. MayaData, the originator of OpenEBS, has also developed Kubera Propel for easy management. OpenEBS Mayastor includes a cloud-native declarative data plane written in Rust and utilizes technologies such as NVMe and SPDK.

Kubera Propel is a simple to use, easy to provision, and high-performance Container Attached Storage solution enabling per workload storage.

Company

MayaData delivers Data Agility. MayaData is one of the top contributors[1] to CNCF projects and leads the development of the fastest open source storage for Kubernetes, OpenEBS. MayaData founded two CNCF projects, OpenEBS – the #1 open source container attached storage solution – and Litmus – the #1 Kubernetes native chaos engineering project, which is now led by the MayaData spin-off ChaosNative. Well-known users of MayaData software include the CNCF, Bloomberg, Comcast, Arista, Verizon, Flipkart,

[1] CNCF DevStats - PR Authors Companies: https://all.devstats.cncf.io/d/21/prs-authors-companies-table?orgId=1
[2] OpenEBS Adopters: https://github.com/openebs/openebs/blob/master/ADOPTERS.md

intel.

Fast databases

Mayastor is based on new technology, SPDK, a new generation of storage software designed for high speed, and low latency NVMe devices like the Intel® Optane™.

OpenEBS Mayastor, when run on Intel® Optane™ SSD DC P4800X Series and the second generation of Intel® Xeon® Scalable Processors delivered maximal sustain throughput of 20kqps (thousands of queries per second) with a peak of 30kpqs, nearly utilizing the full performance potential of P4800X device.

When we run the same benchmark (sysbench) on the next generation of Intel® Optane™ SSD DC P5800X Series and third generation of Intel® Xeon® Scalable Processors we observed a significant performance increase resulting in approximately 3 to 4 times higher throughput than on the P4800X. Sysbench peaked at about 80kqps while delivering sustained throughput at about 65kqps and relatively low resource utilization suggests that multiple such databases could be run concurrently on the host.

MayaData and Intel® together

By partnering with Intel® in the benchmarking and use of Intel® Optane™ and Intel® SPDK, MayaData has been able to validate the approach to deliver extremely high performance per container and workload performance using the open source storage engine OpenEBS Mayastor managed by Kubera Propel. More information about Intel® Optane™ is **available from Intel® here.**

Footnotes - MayaData

Benchmark results obtained against MySQL using the read/write oltp test distributed with the sysbench tool. Sysbench was configured to use 200 threads, with a data set of 8 tables and 10,000,000 rows per table.

Kubera Propel running on Intel® Optane™ SSD P4800X and Intel® Optane™ SSD P5800X Throughput Comparision
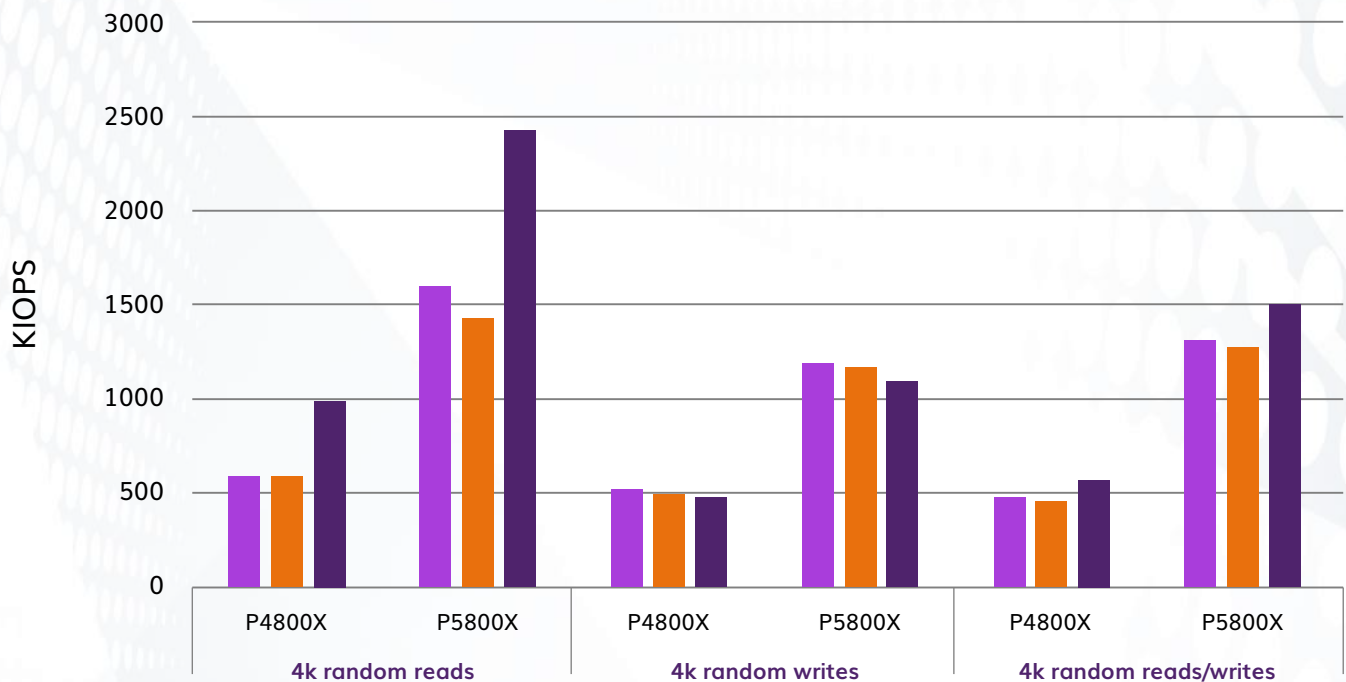
Workload Device & Configuration

● Baseline over PCIe   ● Single replica over TCP   ● Mirrored RAID 1 over TCP

Figure 1:

Chart above includes throughput comparison of OpenEBS Mayastor running Intel® Optane™ SSD P4800X on 2nd generation of Intel® Xeon® Scalable Processors server and Intel® Optane™ SSD P5800X on pre-production version of 3rd generation of Intel® Xeon® Scalable Processors server ("Ice Lake"). Intel® Optane™ SSD P5800X delivers 2-3 times throughput improvement over Intel® Optane™ SSD P4800X, depending on the configuration. Observed overheads introduced by SW layers were minimal, less than 10% when compared to the "Baseline" numbers.

In both cases, OpenEBS Mayastor was able to utilize the full Intel® Optane™ SSD potential.

Spotlight on MayaData

MayaData focuses on enabling Data Agility for its customers and the broader community. MayaData's team of CKAs and CKDs provides 24x7 proactive support for customers and these customers also receive access to Kubera Propel for simplified management. Kubera capabilities include:

Simplified configuration, management, and monitoring of a fast cloud-native data plane for stateful workloads on Kubernetes, such as Kafka, PostgreSQL, Cassandra, and others

Automated lifecycle management of data layer components, including the OpenEBS Mayastor storage engine and underlying storage such as disks and cloud volumes.

The foundation of Kubera Propel, OpenEBS Mayastor, is a breakthrough in per container, open source performance, which approaches theoretical maximum performance measured by various parameters. The low overhead of OpenEBS Mayastor translates directly into less waste in the purchase of hardware and cloud volumes.

Learn more

If you wish to learn more about high-performance Kubernetes storage solutions, you may find these links helpful.
- Kubera product line
- MayaData home page
- Intel® Optane™ SSD P5800X
- Intel® Optane™ Technology for Data Centers

Benchmarking Results

The Container Attached Storage pattern has proliferated & is becoming a de facto standard for running stateful workloads on Kubernetes. The CNCF project OpenEBS has various underlying storage engines that can be deployed to provide the desired capabilities to individual workloads.

OpenEBS Mayastor is the first Container Attached Storage engine developed with the tremendous performance capabilities of Intel® Optane™ and the NVMe protocol itself in mind. Written in Rust and open source, Mayastor is fast becoming the preferred choice for workloads that need the ease of use and portability of OpenEBS plus the performance otherwise only delivered by the much harder to manage and to protect direct access to high performing disks and cloud volumes.

This benchmarking solution guide demonstrated that Kubera Propel and the underlying OpenEBS Mayastor imparted less than a 10% performance overhead. Even with two replicas, which doubles the work for writes, OpenEBS remained within the bounds of less than a 10% overhead. In the case of two replicas, the read performance increases significantly and fully doubles, again showing the ability of OpenEBS Mayastor and managed by MayaData's Kubera Propel to fully leverage the capabilities of the underlying hardware. Further optimization is possible, including configurations that would reduce the time spent on the CPU, an approach that is particularly relevant to hardware constrained edge environments.

In conclusion -- this examination has shown that the Container Attached Storage project OpenEBS Mayastor -- as provisioned and operated by the MayaData software Kubera Propel -- can deliver outstanding performance at the storage layer by leveraging the incredible capabilities of Intel® Optane™ SSD P5800X.

MayaData

intel

System configuration: k3s is used to provision a Kubernetes cluster. After deployment, we label two nodes such that only they run OpenEBS Mayastor, and the remaining node will be used for running the workload itself.

	Workload Type: fio, block size=4k, numjobs=8, qd=64	Baseline [kIOPS]	Single Replica [kIOPS]	Mirrored Raid 1 [kIOPS]	Single Replica Overhead (percent difference relative to Baseline performance)	Mirrored RAID 1 Overhead (percent difference relative to Baseline performance)
P4800X	Random reads (100%)	585	579	993	-1.03%	69.74%
	Random writes (100%)	516	490	473	-5.04%	-8.33%
	Random read/write (50/50)	476	454	568	-4.62%	19.33%
P5800X	Random reads (100%)	1595	1424	2425	-10.72%	52.04%
	Random writes (100%)	1186	1168	1094	-1.52%	-7.76%
	Random read/write (50/50)	1321	1272	1501	-3.71%	13.63%

Table above includes KIOPS numbers of OpenEBS Mayastor running on Intel® Optane™ SSD P4800X on 2nd generation of Intel(R) Xeon(R) Scalable Processors server and Intel(R) Optane™ SSD P5800X on pre-production version of 3rd generation of Intel(R) Xeon(R) Scalable Processor server ("Ice Lake"). "Single Replica Overhead" and "Mirrored RAID 1 Overhead" columns contain percentage of overhead. Negative percentage means performance is lower than "Baseline", positive percentage denotes performance improvement vs "Baseline". Observed overheads introduced by SW layers are minimal. Significant performance improvement for "Mirrored RAID 1" case is observed for configurations including read operations. This is because reads are distributed across different replicas.

Footnotes - MayaData

Source: Actual test results showed an average of 3.56% overhead on a single replica instance
Source: https://mayadata.io/assets/pdf/benchmarking-3-million-writes-with-cassandra-on-k8s-and-kubera-propel.pdf

MayaData

intel.

Disclaimers - Intel

Configuration 1: 3-node, 2x Intel® Xeon® Gold 6252 CPU on Intel® Server System R2224WFTZSR, Intel Optane SSD P4800X Production, Ubuntu 20.10

Test: FIO 3.20 Performance varies by use,

www.Intel.com/PerformanceIndex

Configuration 2: 3-node, 2x pre-production "Ice Lake" CPU on Intel reference platform (Coyote Pass), Intel Optane SSD P5800X Production, Ubuntu 20.10

Test: FIO 3.20 Performance varies by use, configuration and other factors. Learn more at

www.Intel.com/PerformanceIndex