KIOXIA

Why Replace Enterprise SATA SSDs with Data Center NVMe™ SSDs?

A new NVMe™ SSD category has evolved and designed primarily for read-intensive applications in cloud and in hyperscale data centers where they deliver much higher performance and larger capacities versus enterprise SATA SSDs, but within a comparable price tag. These new data center NVMe SSDs feature capacities up to 7.68 terabytes (TB[1]) and deliver complete end-to-end data protection with emphasis on Quality of Service (QoS). There are a number of reasons to replace enterprise SATA SSDs with data center NVMe SSDs in servers, but the top five include:

1. **Big Boost in Performance**
   (Data center NVMe SSDs consistently achieved significantly better performance results and better performance per dollar than enterprise SATA SSDs[2])
2. **Mainstream Availability**
   (Data center NVMe SSDs are available from leading server vendors that include Dell™ and HPE®)
3. **NVMe Monopolizes SSD Deployments**
   (NVMe SSDs, both enterprise and data center, will represent the majority of SSDs used in the enterprise over the next 5 years[3])
4. **Enterprise SATA in Rapid Decline**
   (Enterprise SATA SSDs use will decline from 42% in 2019 to 13% in 2021 to 4% by the end of 2023[3])
5. **Future Cloud Data Center Architectures based on NVMe-oF™**
   (NVMe-oF helps to disaggregate the NVMe SSDs from compute nodes and make them available as shared network-attached resources)

Big Boost in Performance

NVMe SSDs are replacing enterprise SATA SSDs in servers. They connect directly to server CPUs through the PCIe® interface bus, bypassing the host bus adapters (HBAs) that were used in legacy hard drive deployments. Up to sixteen PCIe Gen3 lanes are available for data to flow, where each lane can move data independently, and at maximum speed to keep today's fast CPUs continually fed with data.

The NVMe specification provides efficient host processing of each command and supports a queue depth of 64K commands in 64K queues (versus 32 commands in a single command queue supported by SATA). This enables NVMe SSDs to perform multicore processing of large I/O operations over multiple PCIe bus lanes, thus reducing bottlenecks while keeping data flowing into and out of server CPUs. While SATA theoretically provides 6 gigabits per second (Gb/s) performance, NVMe SSDs deliver a bigger boost in bandwidth and IOPS (input/output operations per second) performance with the ability to transfer data at 1 gigabyte per second (GB/s) per lane, or upwards of 4GB/s in a typical x4-lane configuration. This big boost in performance enables NVMe SSDs to handle more I/O requests quickly and efficiently.

	Enterprise SATA SSDs 6Gb/s		Data Center NVMe SSDs Gen3x4		Data Center NVMe SSD Advantage	
Read / Write Operation	**Interface Limit**	**Practical Specification**	**Interface Limit**	**Practical Specification**	**Interface Limit**	**Practical Specification**
Sequential Read	600MB/s	550MB/s	4,000MB/s	3,000MB/s	6.67x	5.45x
Sequential Write	600MB/s	550MB/s	4,000MB/s	1,800MB/s	6.67x	3.27x
Random Read	146,000 IOPS	85,000 IOPS	976,000 IOPS	500,000 IOPS	6.68x	5.88x
Random Write	146,000 IOPS	35,000 IOPS	976,000 IOPS	50,000 IOPS	6.68x	1.43x

Chart 1: Performance comparisons between enterprise SATA and data center NVMe SSDs

Chart 1 compares the interface limit[4] and practical specification[5] of enterprise SATA SSDs versus data center NVMe SSDs and includes benchmark testing in a lab environment[6]. From the internal test results, in every case, data center NVMe SSD performance far exceeded enterprise SATA's capabilities.

In addition to internal lab testing, external third-party testing was conducted by Principled Technologies® that compared transactional database performance and database analytics performance between data center NVMe SSDs and enterprise SATA SSDs[7]. The tests included transactions per second (TPS), operations per second (OPS), and associated per dollar costs.

TPS database performance tests were run on a Dell EMC™ PowerEdge™ R740xd server. From the test results, data center NVMe SSDs delivered an increase in transactional database performance by more than 2x when tested against enterprise SATA SSDs (Figure 1). The test methodology and results are publicly available[8].

Total Transactions per Second (TPS) – higher is better:

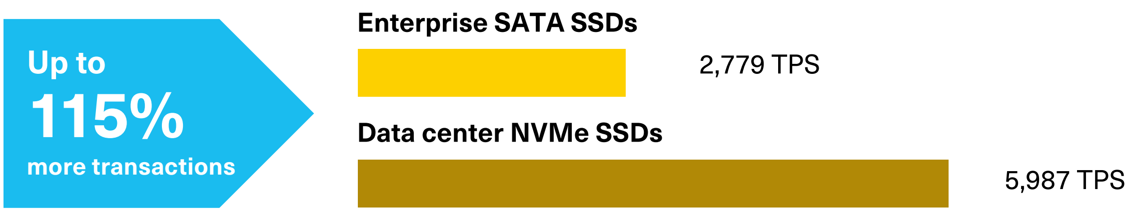

Figure 1 compares transactions per second between enterprise SATA and data center NVMe SSDs

In addition to the TPS test results, a cost assessment as to how many transactions per dollar that each configured SSD was able to achieve was also compared. As a result, NVMe SSDs demonstrated similar results for transactions per second per dollar than enterprise SATA SSDs (Figure 2). The test methodology and results of the test are also publicly available[8].

Total Transactions per Second (TPS) per Dollar – higher is better:

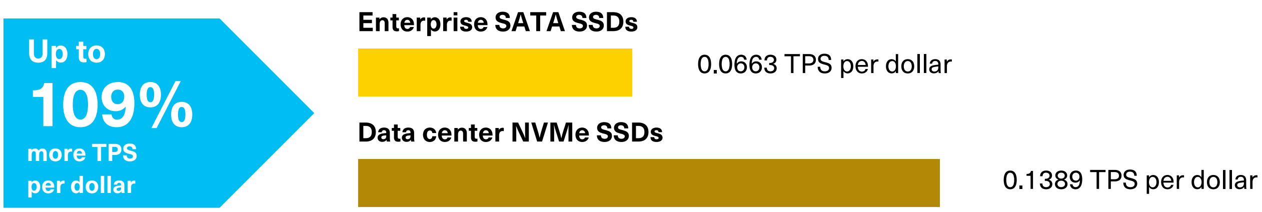

Figure 2 compares transactions per second per dollar between enterprise SATA and data center NVMe SSDs

The operations per second (OPS) database analytics performance tests indicate how many operations per second that the SSD can process when configured on a Dell EMC PowerEdge R840 server with enterprise SATA and data center NVMe SSDs. The results indicated that data center NVMe SSDs improved OPS performance by almost 2.5x when tested against enterprise SATA SSDs (Figure 3). The test methodology and results of the test are publicly available[9].

Total Operations per Second (OPS) – higher is better:

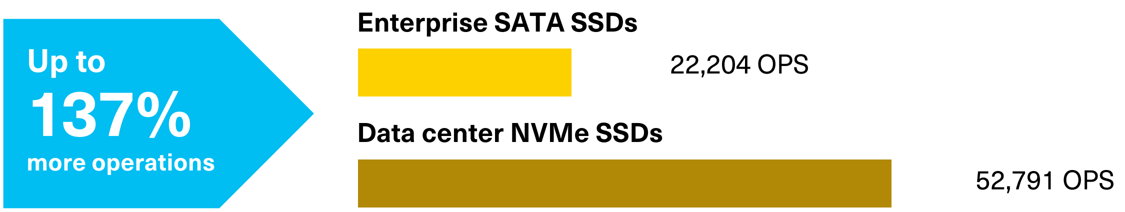

Figure 3 compares operations per second between enterprise SATA and data center NVMe SSDs

In addition to the OPS test results, a cost assessment as to how many operations per dollar that each configured SSD was able to achieve was also compared. As a result, NVMe SSDs demonstrated similar results for operations per second per dollar than enterprise SATA SSDs (Figure 4). The test methodology and results of the test are also publicly available[9].

Total Operations per Second (OPS) per Dollar – higher is better:

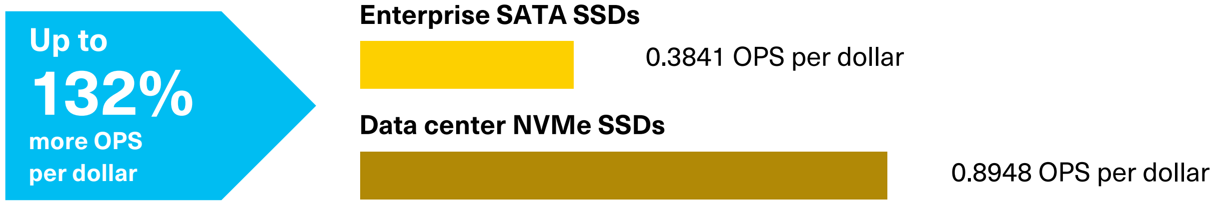

Figure 4 compares operations per second per dollar between enterprise SATA and data center NVMe SSDs

Mainstream Availability

Data center NVMe is a new category of PCIe/NVMe SSD media designed and optimized for read-intensive and latency-sensitive applications. KIOXIA's [CD5 Series] of data center NVMe SSDs are developed using 64-layer BiCS FLASH™ 3D memory technology and available in 960GB to 7,680GB capacities. The CD5 Series delivers significantly more bandwidth and IOPS for performance-oriented applications when compared to enterprise SATA SSDs.

CD5 Series NVMe SSDs are now available on selective leading servers from Dell EMC PowerEdge and HPE ProLiant® servers (Chart 2).

CD5 Series SSD Capacity	Dell™ PowerEdge™ SKUs	HPE® ProLiant® SKUs
960GB	KCD5XLUG960G	P10208-B21
1,920GB	KCD5XLUG1T92	P10210-B21
3,840GB	KCD5XLUG3T84	P10212-B21
7,680GB	KCD5XLUG7T68	Contact HPE

Chart 2 outlines CD5 SKU availability on leading servers

NVMe Monopolizes SSD Deployments

NVMe SSDs will represent a majority of the SSDs used in the enterprise, about 45% (in units) of the total SSD consumption by the end of 2019[3]. This consumption is expected to grow to 78% by the end of 2021 and 91% by the end of 2023[3], completely dominating SSD server deployments. So why have NVMe SSDs become so dominant?

The NVMe interface increases flash storage performance between PCIe-based servers and SSDs by eliminating the SCSI command stack (and associated protocol translation overhead). When SATA SSDs are deployed within a server, I/O commands must traverse through multiple layers of the command stack (Figure 5), which is acceptable for hard drives, where the media is the bottleneck, but not for SSDs. As SATA is half-duplex, only one lane / one direction can be used to transfer data at one time. With CPUs getting faster and DRAM bandwidth increasing, the single SATA lane can become a bottleneck where servers are left waiting for data transactions to complete leaving some of the compute, storage and memory resources underutilized.

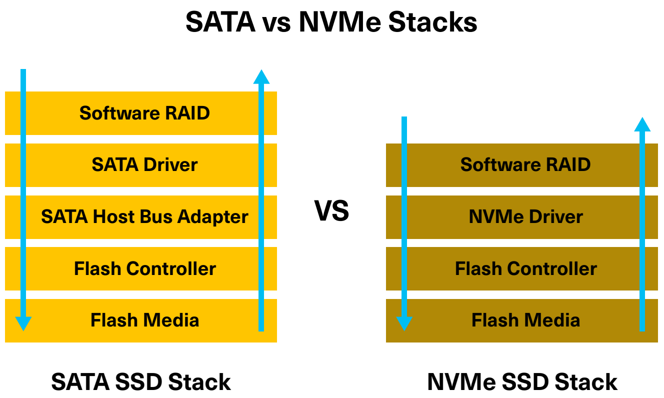

Figure 5 shows SATA and NVMe SSD stacks

Alternatively, NVMe SSDs connect to servers through the PCIe interface bus where each PCIe Gen3 lane can move data at speeds up to 1GB/s, or up to 4GB/s in an x4 configuration. This enables multicore processing over multiple PCIe lanes that reduce bottlenecks and keep data flowing into and out of server CPUs so that more demanding storage workloads can be handled quickly and efficiently.

Enterprise SATA in Rapid Decline

Enterprise SATA SSDs are widely deployed in servers based on their cost-effectiveness and are used mostly for application workloads that don't require exceptionally high performance or large capacities. The interface itself has reached a performance plateau with no speed improvements planned in its future. For many applications, the limited SATA SSD performance may become a bottleneck to server CPUs, preventing them from reaching their operational or transactional potential. The end result may be an underutilization of the server's compute capabilities which can also limit the server's ability to service many clients simultaneously.

Enterprise SATA SSDs will represent about 42% of the total enterprise SSD consumption (in units) by the end of 2019[3]. The total SATA SSD consumption in the enterprise will rapidly decline to 13% by the end of 2021, and expected to decrease further to 4% by the end of 2023[3].

Future Cloud Data Center Architectures based on NVMe-oF

Today's data infrastructures (on-premises and cloud) are typically based on general-purpose architectures that utilize a pre-configured allocation of compute and storage resources to meet a variety of application workload demands. As these data-intensive applications proliferate on a larger scale, the one-size-fits-all approach comes up short relating to performance, capacity and scalability, and are challenged to meet varied application requirements. The need to disaggregate these resources has become an IT priority so that organizations can independently scale storage and compute, stretch their IT budgets, and maximize hardware resource utilization.

Specialized software that works in conjunction with a widely deployed network protocol is required to enable flash-based NVMe SSDs to disaggregate from the compute nodes and make them available as network-attached shared resources. NVMe over Fabrics (NVMe-oF) technology has quickly become the protocol of choice. It can pool flash resources, and provision the right amount of storage and compute for each server and application in the data center. In an orchestrated cloud environment, NVMe-oF technology enables the server to borrow compute resources from lower priority applications and provide them to key applications during peak demand times. While this disaggregated approach is similar to other block storage technologies, such as iSCSI, storage is accessed at much lower latencies using NVMe-oF pooled storage for sharing by different hosts.

The storage strategy is to move toward a shared infrastructure and cloud orchestration so that the right amount of storage and compute can be allocated for each application workload while delivering almost the same high-performance, low-latency benefits as if the NVMe SSDs were locally-attached.

KIOXIA's KumoScale™ software delivers 'NVMe as a service' and enables IT users to create a disaggregated rack-scale NVMe-oF storage solution for the data center.

KUMOSCALE™

Learn more about KIOXIA's CD5 Series and Life After SATA:
https://business.kioxia.com/en-us/ssd/life-after-sata.html

Notes:

[1] Definition of capacity: KIOXIA Corporation defines 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 1GB = 2^{30} bytes = 1,073,741,824 bytes, 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, such as Microsoft Operating System and/or pre-installed software applications, or media content. Actual formatted capacity may vary.

[2] Based on transactional and operational database performance, tested internally or externally, or based on published vendor specifications.

[3] Source: IDC. - "Worldwide Solid State Drive Forecast, 2018-2023, Market Forecast-Table 12, Jeff Janukowicz, May 2019, IDC #43828819.

[4] The interface limit represents the highest read or write performance that 'could' theoretically be obtained from an SSD, and takes into account the physical limitations of the electrical connections.

[5] The practical specification accounts for typical performance at the product level based on benchmark testing in a lab environment.

[6] 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 existing KIOXIA SSD products, tested at 1 DWPD (Drive Write per Day) and include the HK4 Series of SATA SSDs and CD5 Series of NVMe SSDs.

[7] For these tests, Principled Technologies used data center NVMe SSDs based on the KIOXIA CD5 Series (960GB capacity, 2.5" SSDs) and enterprise SATA SSDs based on the Intel® D3-S4510 Series (960GB capacity, 2.5" SSDs).

[8] The Principled Technologies transactional database performance test results are included in the report entitled, "Get significantly better transactional database performance for less from a Dell EMC PowerEdge R740xd server with value SAS and data center NVMe SSDs from KIOXIA," at the following link: http://facts.pt/q7mwkro.

[9] The Principled Technologies database analytics performance test results are included in the report entitled, "Get better database analytics performance at a lower cost with Dell EMC PowerEdge servers equipped with value SAS and data center NVMe SSDs from KIOXIA," at the following link: http://facts.pt/r43qggb.