The Business Case for NVMe PCIe SSDs

Business Environment

Data-intensive businesses that thrive in today's environment move quickly, and data centers must move with them. To that end, technologies—such as 40-gigabit (and faster) Ethernet and multicore servers—are chasing latency out of the data center. One of the easiest targets for latency removal is the spinning platters of a server's direct-attached storage, its hard disk drives (HDDs).

Fast solid state drives (SSDs) are quickly replacing HDDs to garner more performance (measured in I/O operations per second, IOPS) in servers. While using SSDs instead of HDDs in any storage platform can significantly improve application performance, the potential gains are even more dramatic with SSDs connected directly to the server's PCIe® bus via NVMe™[1].

NVMe is a standard storage protocol[2] aimed at bringing storage I/O closer to the system processor for both faster IOPS performance and lower average latency. With the broadening use of NVMe SSDs (including those supporting hot-swap and hot-add for real time replacement and real time expansion), IT teams have access to low latency, high IOPS, expandable capacity and familiar supportability to better enable applications, users and management demand.

If high IOPS performance with low average latency is a main driver for your storage upgrade path, NVMe SSDs may be a great option.

SSD Landscape

SSDs are currently available in three host-to-drive interfaces: SATA, SAS and PCIe[3]. SATA and SAS SSDs require a host bus adapter (HBA), either integrated into the chipset on the main server board or as an expansion card in a PCIe slot. The host operating system controls these SATA or SAS HBAs and facilitates communications with the rest of the computing architecture.

All of these drive types offer remarkably better IOPS performance than legacy HDDs, but their performance may be limited by external factors such as the operating system (OS), driver stack and a poor efficiency caused by protocol conversions (note that these protocol conversions are necessary for SATA and SAS to get data from the host system, to the storage devices, and back).

1. This paper uses the abbreviated NVMe SSD to indicate NVMe SSDs that use the PCIe bus. This abbreviation is common in the industry.
2. For additional information about NVMe, visit http://www.nvme.org/
3. PCIe is a physical interface like SATA or SAS. NVMe is a storage protocol running over PCIe.

Micron®

NVMe SSDs enable much faster IOPS because they are not subject to the same extensive protocol conversions and OS driver stacks that SATA and SAS SSDs and their HBAs require. When we couple this more direct connection with the protocol advantages NVMe brings, we get a tremendous boost in IOPS performance.

Figure 1 shows an example of 4KiB random read IOPS among three high performance Micron® SSDs (9100 PRO NVMe, S610DC SAS and 5100 ECO SATA) along with a performance SAS HDD (15,000 RPM).
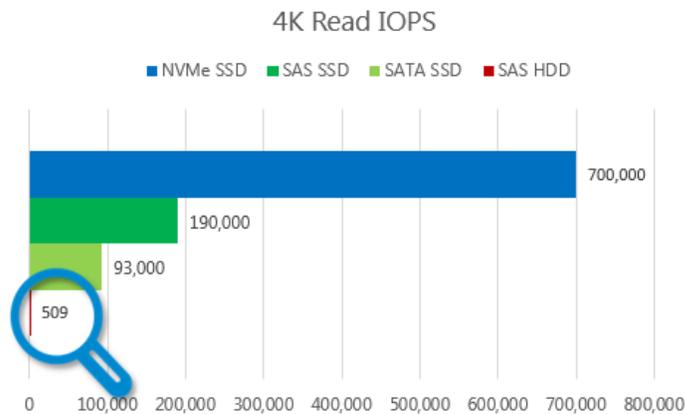

Figure 1: 4K Read IOPS by Drive Type

The data in Figure 1 for the PCIe NVMe, SAS and SATA SSDs is from Micron's SSD datasheets[4,5,6] (note: different capacities and versions have different performance characteristics). The SAS HDD performance was measured in our lab following the SNIA Performance Test Specification Enterprise version 1.1 IOPS test (HDD datasheets typically do not publish a 4K random read IOPS value – more information about the SNIA Performance Test Specification is available from www.snia.org).

The data in Figure 1 is color-coded: the NVMe SSD at the top in blue, the SAS SSD just below in green, the SATA SSD in light green and at the very bottom the SAS HDD in red. In Figure 1, a longer horizontal bar is better.

- The PCIe NVMe SSD is the 4K random read IOPS performance leader with 700,000 IOPS.
- The SAS and SATA SSDs are second and third with 190,0005 and 93,000 respectively.
- The SAS HDD measures so low that it is hard to see – measuring just 509 IOPS.

This raw speed advantage can translate to noticeable performance gains in I/O-intensive workloads that require fast responses such as those running on cloud and data center servers, video-on-demand, and virtualization applications, among others.

While PCIe NVMe SSDs show much faster read IOPS than other types of SSDs (and their read IOPS are orders of magnitude faster than HDDs), their superior speed alone may not be enough of a reason to switch: eye-popping IOPS will not help if serviceability suffers - IT simply can't afford to take servers offline to replace drives or increase capacity.

4. 9100 PRO U.2 or HHHL
5. S610DC SAS SSD
6. 5100 ECO SATA SSD

In these cases, reliability, availability, and serviceability (also called 'RAS') should also figure prominently. That means choosing SSDs that offer superior performance, are hot-swappable, and easily accessible. IT can then avoid having to shut down a server and open the chassis to add capacity or replace a drive.

NVMe SSDs support these capabilities. As the 2.5-inch PCIe SSD form factor – also called "U.2" – gains popularity, IT enjoys this same flexibility that SAS and SATA SSDs have enjoyed for generations, but with all the advantages of NVMe. Features like hot swap and hot add (for replacement and capacity expansion) without having to take the server offline are now part of the PCIe NVMe SSD landscape.

Phenomenal Performance Has Real Benefits

NVMe SSDs can clock in at 700,000 random 4K read IOPS, as reflected in Figure 1. With the growing number of servers capable of keeping these NVMe SSDs busy with demanding workloads (high core counts and greater memory density), application performance and throughput can quickly and easily be increased by replacing HDDs with much faster NVMe SSDs.

For example, a virtual desktop image store that serves a large number of virtual desktop users has to support a phenomenal number of small, random IOPS (VDI is typically a small, random IO profile that's very heavily read focused).

According to VMware's "Server and Storage Sizing Guide for Windows 7 Desktops in a Virtual Desktop Infrastructure"[7], each power user needs 17-25 IOPS[7]. Atlantis Computing's "Windows 7 IOPS for VDI: Deep Drive"[8] puts the number higher at 25-50 IOPS[8].

Using 25 IOPS per user to estimate the number of each type of drive needed[9], we see that we have to spread the load across several HDDs. For example, if we wanted to estimate how many of each drive we'd need if we wanted to service 500 power users, we would need at least 25 SAS HDDs, while we would need just one SSD (the minimum number we could deploy). Figure 2 shows this comparison.
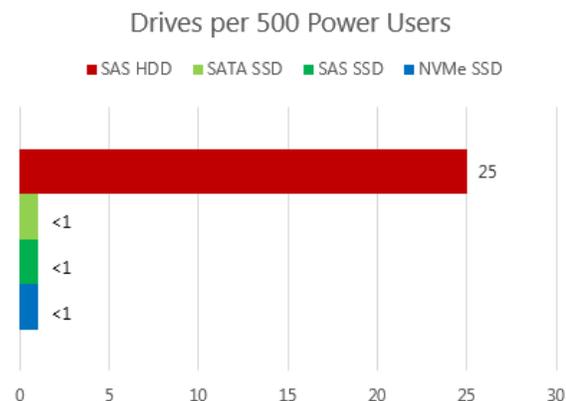

Figure 2: Drives to Support 500 VDI Power Users

7. "Server and Storage Sizing Guide for Windows 7 Desktops in a Virtual Desktop Infrastructure"
8. "Windows 7 IOPS for VDI: Deep Drive"
9. Estimated values are based only on storage device 4K random read IOPS. Actual number of users will vary based on multiple factors and may differ.

With NVMe SSDs offering performance advantage of several orders of magnitude (single drive comparison as shown in Figure 1), the same user base could access their images stored on far fewer SSDs than HDDs.

In addition to breakthrough performance, PCIe NVMe SSDs provide more useful work per dollar spent. Because of this, PCIe NVMe SSDs save money despite the comparatively high initial acquisition cost.

When performance is the primary driver, the metric to consider should be cost per IOPS. This metric shows the value that SSDs deliver when they are doing what they do best: quickly storing and delivering data to applications. That is why the initial investment can be misleading—because cost per gigabyte does not factor in performance advantages.

We can evaluate each device in terms of relevant work (using our 4K random IOPS example from Figure 1) per dollar using publicly available selling price data from searching http://www.froogle.com/ and averaging the first three results.

Figure 3 shows a relative IOPS per dollar comparison for the same four example drives (three SSDs and one HDD) and a 4KiB 100% read IOPS (as might be seen in a virtual desktop image store).
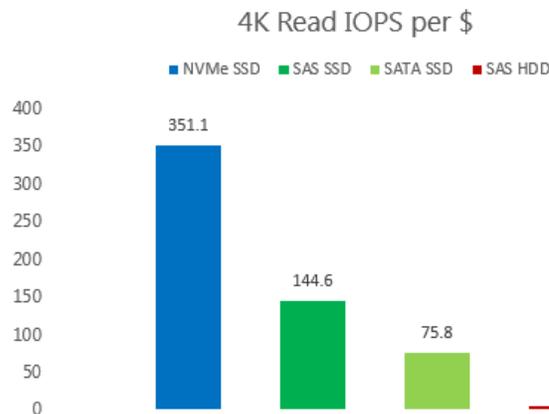

Figure 3: 4K Random Read IOPS per Dollar

The values in Figure 3 were calculated taking the number of 4KiB read IOPS for each drive type (from Figure 1) and dividing by the average of three single drive selling prices from http://www.froogle.com/– giving IOPS/$.

Note that this is an example calculation – if other transfer sizes or read/write ratios are used, the results would differ. As pricing changes, the results would also differ.

In Figure 3, the calculated IOPS/$ is shown at the top of each vertical bar. As in Figure 1, the data is color-coded. The NVMe SSD at the left in blue, the SAS SSD in green, the SATA SSD in light green and the SAS HDD in red. In Figure 3, taller is better.

More Energy Efficient

Energy consumption is another important consideration when choosing a drive type. NVMe SSDs draw a maximum energy load of 25 watts compared to about 7.2 watts for a 15,000 RPM HDD[10]. However, one NVMe SSD can potentially replace many HDDs (again referring Figure 1 as one example).

If we again consider a virtual desktop image server equipped with 25 HDDs, those drives would draw about 180 watts on average. If the array were larger, say 50 HDDs, it would draw approximately 360 watts. If larger still – 100 drives, the power requirement is even greater, about 720 watts. Thus, the NVMe SSD has the potential to result in a significant reduction in energy use for that storage platform.

PCIe NVMe SSDs: More than Speed

It makes good business sense to upgrade to PCIe NVMe SSDs.

Because of the SSD's incredible performance, multiple HDDs can be replaced with one or a few NVMe SSDs, which can provide a vast improvement in the per IOPS dollar and much better power efficiency. And with integrated management, NVMe SSDs help IT ensure smooth operation. These benefits add up to big savings that can more than offset the acquisition cost of NVMe SSDs.

Contact your Micron sales representative or visit micron.com to explore the opportunities awaiting you with Micron® PCIe NVMe SSDs.

10. Example datasheet, active power value used in this example

If high IOPS performance with low average latency is a main driver for your storage upgrade path, NVMe SSDs may be a great option.

micron.com

Micron®