SAMSUNG

White Paper:

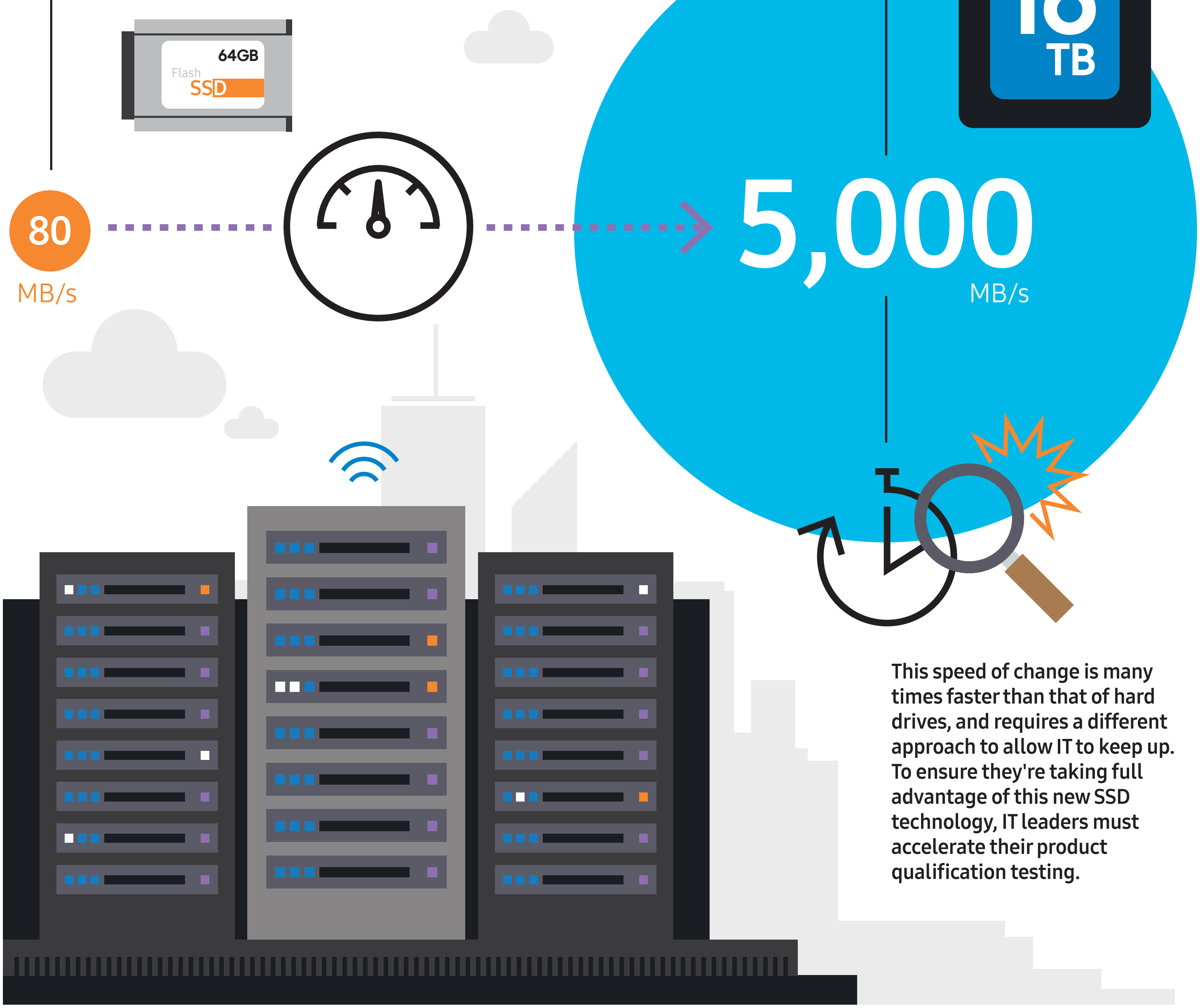
Managing the Enterprise Storage Life Cycle: How to Keep up With Fast-Evolving SSD Technology

Introduction

The world of storage looks utterly different than it did just a few years ago, and technology professionals need a finger on the pulse of these advancements.

2007 — Solid state drive (SSD) technology is advancing at a rapid rate. In 2007, SSDs were available with 64 GB capacities and 80 MB/s throughput.

Ten years later, we have SSDs with capacities as large as 16 TB and others with throughputs up to 5,000MB/s. — 2017



This speed of change is many times faster than that of hard drives, and requires a different approach to allow IT to keep up. To ensure they're taking full advantage of this new SSD technology, IT leaders must accelerate their product qualification testing.

Increased Performance of Newer SSDs

With SSDs, the pace of improvements is faster than ever — now doubling more than every 18 months.

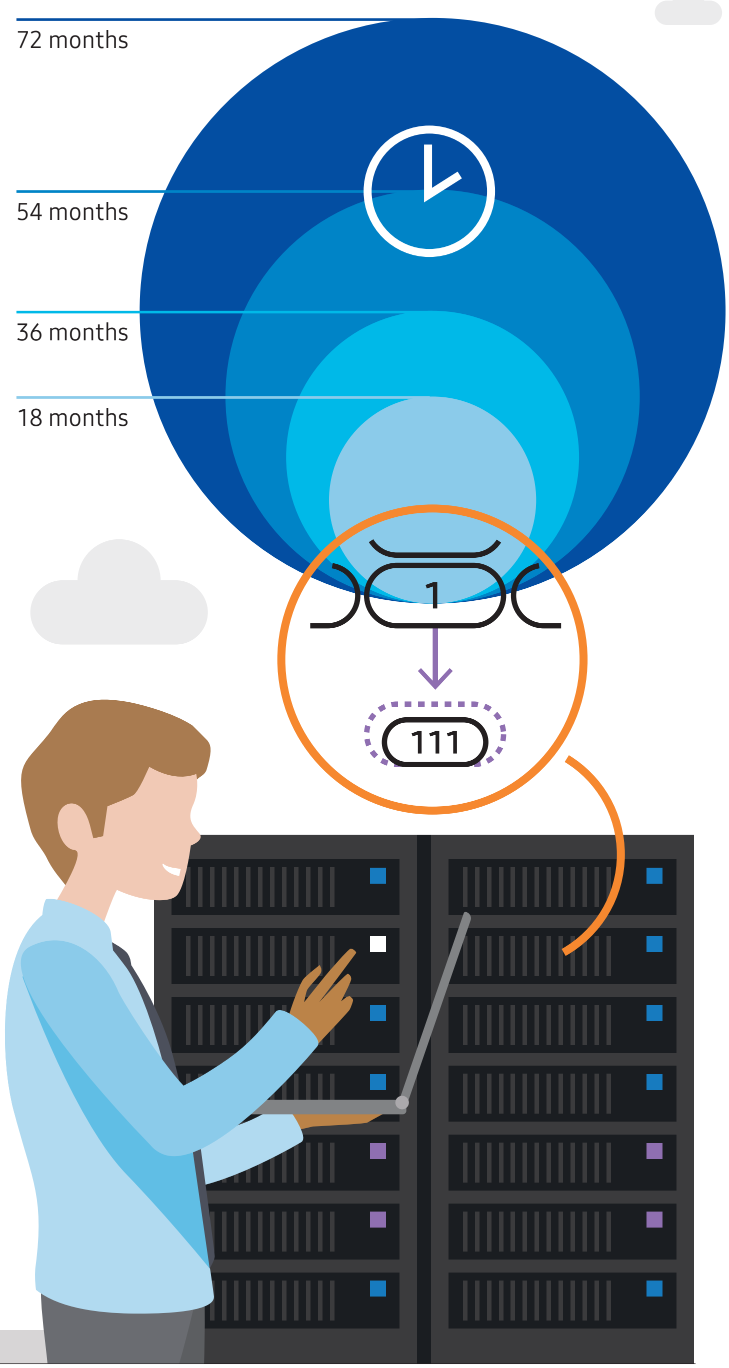
SSD performance now easily exceeds not only traditional hard disk drives (HDDs), but the Serial ATA (SATA) bus interface itself. This has led to the development of new interfaces, such as the Non-Volatile Memory express (NVMe). Current SSD capacities of 16 TB, 32 TB and greater are larger than the highest capacity mechanical HDDs, and Read/Write (R/W) operations and input/output operations per second (IOPS) continue to progress at a rate comparable to — or faster than — CPU advances, as opposed to the relatively lethargic rate of traditional HDD progress.

Compared to the Moore's Law operating in CPUs and memory, which sees capacity doubling every 18 months, upgrades to old, mechanical HDDs were fairly slow, with improvements being made on a three- to five-year cycle. But with all-electronic SSDs, the pace of improvements is faster than ever — now doubling roughly every 18 months.

Initially, increasing storage capacity in SSDs involved shrinking the physical storage cells, similar to how a city planner would put more houses closer together in order to fit more people in the same space. But a breakthrough came in the ability of SSDs to store two or three bits of data per cell, rather than just one. In our example, this means adding one or two more roommates to your house. Each one of these changes (shrinking the size and adding bits) led to performance challenges that had to be corrected via new firmware. Consequently, each new product release required new qualifications and testing by IT engineers who were using those drives.

Eventually, the overall geometry of the cells became to small, and cell to cell interference began to occur, which required additional error correction by the controller. Manufacturers then took a step back, increasing the sizes of cells and pathways to the point where they were reliable again, and then stacked multiple layers of cells on top of each other. For example, a 48-layer chip has 48 separate layers of cells stacked on top of each other. Adding new stacks does not impact the cell size, reducing the need for constantly evolving performance-correcting firmware and constant qualification. The controller is already designed to allow for multiple layers, so a layer transition doesn't mean that a new controller has to be developed. But a controller can be updated to take advantage of new technologies.

Consistent with Moore's Law, flash-based SSDs have continued to double in capacity roughly every 18 months, due to ever-increasing transistor density.

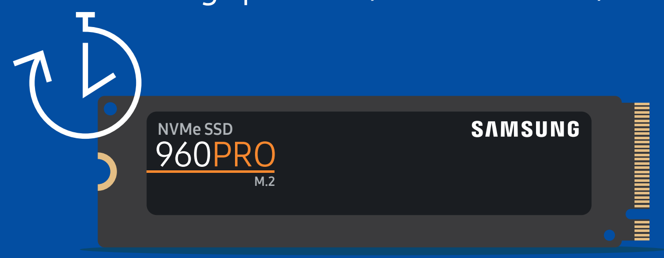


Greater Capacities and Endurance

There are now enterprise data center-grade SSDs that can be written to constantly for years at a time without wearing out cells.

The Non-Volatile Memory express (NVMe) standard improves upon SATA and is specifically designed for solid state drives.

The most recent NVMe drives for client PC applications have throughput of R/W 3.5-2.1 GB/s.



It took hard drives more than 50 years to go from 10 MB capacities to 10 TB, while flash-based SSDs have gone from 18 GB in 1999, to 15 TB in 18 years. At this same time, prices have dropped precipitously.[1] Part of the rapid rate of increase to date has been due to the adoption of "layering" chips, referred to as Vertical NAND (V-NAND) or 3D-NAND. V-NAND has allowed SSDs to go from one-layer planar technology to the current 64-layer V-NAND technology, with 72-layer and 96-layer chips on the horizon. This cannot continue indefinitely, but for now, the progress is rapid and, for the first time, has resulted in SSDs with higher capacities than hard drives.

NVMe and Transfer Speeds

Changes from first-generation SSDs to later versions increase performance by 5x or more, and the move to NVMe increases it at least another 5x. Initial runs of SSDs from about 10 years ago had performance rates in the 100-150 MB/s range (which was 2-3x the hard drive performance); the latest SSDs are in the range of 400-500 MB/s; and the most recent NVMe drives have performance rates of 2.5-5 GB/s or 2500 MB/s.

SSDs don't have the mechanical delays faced by traditional magnetic hard disk drives, which must move a physical head from track to track, and then wait for the rotation of the disk to bring the necessary bit under the head. This means that the response time for every bit on an SSD is the same as every other. In addition, improvements in the relative speeds of the cells have pushed the overall performance of SSDs to the point that a single drive can deliver higher performance than the SATA bus interface can support. SATA was designed to improve data transfer speeds between the serial bus and traditional HDDs from legacy Integrated Drive Electronics (IDE). Because of this, SATA is not designed for SSDs and has limitations on bandwidth speeds between the SATA bus and SSDs. As a result, the NVMe specification was developed, and NVMe SSDs can currently sustain performance numbers 5x that of the best SATA-based SSDs. NVMe or Non-Volatile Memory Host Controller Interface Specification (NVMHCI) is a logical device interface specification for accessing non-volatile storage media attached via PCIe bus. The combination of NVMe with the next generation of PCI Express (PCIe) should support even faster systems. In the meantime, some manufacturers are even putting flash memory on the DIMMs intended for RAM in order to boost performance.

The endurance of SSDs has improved as well, as early SSDs had issues with the total number of times data could be written to each cell. While this number is not infinite, it has improved dramatically, yielding drives that can sustain 24x7 operation over a period of years without cells failing, and without performance degradation during heavy use. In addition, the reliability of SSDs has proven to be 5-10x that of hard drives.[2] There are now enterprise data center-grade SSDs that can be written to constantly for years at a time without wearing out cells.

Need for Regular Evaluations

Companies need to evaluate new SSD technologies quickly, or they run the risk of buying already outdated technology.

Hard drives have gotten to the point where improvements in performance and capacity are incremental in two years; capacity might go from 6 TB to 8 TB, and access times from 18 ms to 16 ms, but no real difference in storage architecture is needed to cope with these changes. On the other hand, SSDs in the same two years might see a 5x jump in performance and a 2-4x increase in capacity. These kinds of changes have pushed manufacturers to develop the new PCIe 4.0 specification and the NVMe interface, among others. Companies that are used to making buying decisions based on the slow pace of hard drive development face the potential for real problems if they assume that SSDs can be evaluated in the same fashion. A change in hard drive models would simply require a new part number on the bill of materials.

A change in SSD models to the next version, however, may require a completely new hardware specification, interface and storage architecture, as well as different protocols for error correction, disaster recovery and more.

How Can Companies Improve the Evaluation Process?

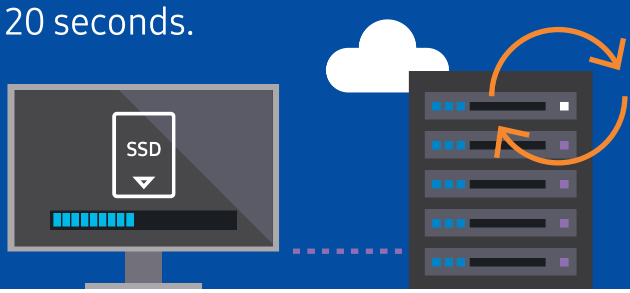
The latest versions of SSDs now coming to market offer a 5x improvement over the last generation.

Companies searching for ways to remain competitive are always trying to keep pace with new technologies. Not only are they enabling new hardware, but they're using new technologies like big data, Online Transaction Processing (OLTP) and the Internet of Things, as well as old standbys like database servers and e-commerce. All these technologies can stress any storage system to the utmost. This means that there will always be a drive to improve storage performance, which will require evaluating each improvement in components and systems. Since the pace of SSD development is extreme, systems architects and storage administrators should get used to the idea of at least annual testing. New versions are available every year, and between capacity, transfer speeds, latency and input/output operations per second, vast improvements are occurring in very short time frames.

Administrators should also be constantly evaluating the new SSDs that come along. The latest versions of SSDs now coming to market offer a 5x improvement over the last generation. In addition to their use in storage arrays, small numbers of drives added to existing servers can improve their performance enough to extend their useful lives by years. NVMe drives might at first glance seem incompatible with older server motherboards that don't have NVMe support, however, inexpensive PCIe cards known as 'sleds' give any motherboard with an open 4x to 16x PCIe slot the ability to support one to four NVMe SSDs.

Testing may be done easily using Iometer and other low-level drive testing applications, but the most interesting results can be achieved with apps that simulate the loads created by those running in the data center. For example, Microsoft's LoadSim generates loads on a Microsoft Exchange Email server. An open source projects application originally created by Microsoft, also called LoadSim, allows companies to create simulated loads on web servers. There are many different tools for generating loads on databases, from the TPC-C OLTP benchmark to apps specific to Oracle, Microsoft or open-source versions of SQL servers.

New SSDs can reduce server reboot times from 1-2 minutes to less than 20 seconds.



In addition to the obvious improvements in speed and latency, the new SSDs can reduce reboot times in servers, often from a minute or two to less than 20 seconds. This is of substantial interest to administrators trying to keep their offline time for mission-critical servers as close to zero as possible.

In addition to regular testing, administrators should take the following actions to improve their evaluation process:

Adding Efficiencies

After testing, SSDs can be used to upgrade existing servers or workstations (perhaps for the executives who approve the test budget). Once the evaluation of the new drives is completed, the test units can be added to existing servers. Administrators interested in extending the usable life of servers will find few other add-ons that have as much potential to keep a server useful than speeding up storage performance by a factor of five.

Information Sources

In addition to in-house testing, there are a variety of sources to turn to for information on improvements in storage devices. There is no substitute for real-world testing, especially if you have apps beyond the usual five or six that most analysts or reviewers are likely to test with. That also applies if you have more complex apps, such as sales force automation, data lake tools or Security Information and Event Management (SIEM), which may show drastic improvements in performance when adding SSDs to the servers, but which most publications and analysts will not test due to the complexity of the test setup. On the other hand, there are many websites that will test SSDs to verify manufacturers' claims on throughput, IOPS and latency, which will give you a good place to start when selecting candidates for testing.

For example, Calypso Systems created a white paper, "Datacenter Server Workload and Performance Analysis," in which they tested several data center SSDs and SAS HDDs on a 24-hour SQL Server workload from a 2,000-outlet retail web portal. Unlike typical synthetic lab workloads, Calypso is using new IO capture and analysis tools that allow them to capture, analyze and test data center Real Word Storage Workloads (RWSWs) for a 2,000-outlet retail chain web portal. Calypso conducted the testing, captured the Block IO level data and published the data results of the various data center SSDs and SAS HDDs in the white paper. The white paper and tools are free and available to download at www.TestMyWorkload.com.

Magazines and Websites

Magazines and websites such as Storage Magazine, TechTarget, Network World and Tom's Hardware perform tests on everything from individual SSDs to large storage systems from vendors like EMC. There are many websites — both current and former print publications, analyst firms and manufacturer's sites with data, blogs and how-to papers — that can give you good data on the basics of SSD performance, what the new standards might mean for storage systems that incorporate the new SSDs, and technical details on what criteria might be easily overlooked during evaluations, as well as potential problems to look for. It's also worthwhile to look at the credentials of the various sites, as some are 20-30 years old, with reputable writers, while others may be only a year or two old.

Analysts

Analyst groups perform the same kinds of testing that your company might want. The costs vary from one group to another, but between the time saved on testing and the costs of buying the parts to test, the fees may be worth it, particularly if you don't have in-house expertise.
Some analyst groups do extremely thorough testing designed to show the effects of real-world applications, while others, even some very large and well-known firms, take money from both sides, charging vendors to be part of the test and consumers to access the data. Therefore, it's important to perform adequate research on the background of analyst groups before deciding which one to work with.

User Groups

User groups such as the Storage Networking Users Group can provide information on real-world usage and results of SSDs.

Vendor user groups can also be good sources of information — particularly when it comes to specialized information such as how to get a certain storage system configured to best support a particular database, or what might be causing problems when connecting an iSCSI system from a particular vendor to a Windows 2016 server.

Standards Bodies

Organizations such as the Storage Networking Industry Association (SNIA) not only ensure that new standards such as NVMe are interoperable between vendors, they also test systems to provide benchmarks that show the real improvements from generation to generation.

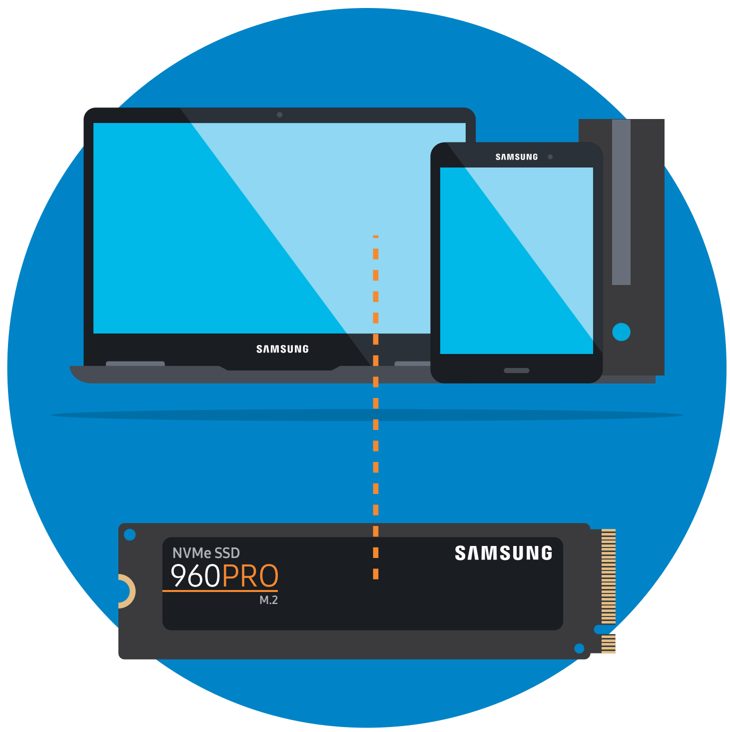
SNIA is an industry group intended to ensure that storage projects from different vendors can interoperate properly so that switches from Brocade will work with storage systems from EMC and so forth. In addition to testing interoperability, the group publishes benchmarking standards, and the website holds many member-authored papers that can be excellent resources for technical troubleshooting and making architectural decisions on new storage systems. The website also provides information on upcoming standards and new products. Members can be good resources for specific questions, as most major vendors have representatives who can be contacted through the group.

What Are the New Technologies?

In addition to 3D V-NAND, which is the vertical stacking of cells that results in higher-capacity SSDs, NVMe is the other new technology that's reshaping SSDs, replacing SATA or Fibre Channel and allowing the new, faster SSD devices to achieve their full performance levels. There are two parts to NVMe: the physical interface and the protocol. The protocol is essentially the same as the PCIe bus used for expansion cards in PCs, using a 4-lane (4x) channel for each drive. Sleds with one to four slots for NVMe drives can be plugged directly into a PCIe slot, with a 1x card supporting one drive and a 4x card supporting up to four drives. There are two physical specifications for NVMe interfaces: the U.2 connector and the M.2 connector. SSDs with the U.2 connector look similar to a 2.5-in. SATA drive, but have a different connector for the PCIe bus. M.2 NVMe SSDs are long and narrow similar to a "gums stick" and connect through a M.2 connector. Additionally, the M.2 NVMe SSD will contain flash memory chips on a single side (single-sided) or on both sides (double-sided). They're being adopted in laptops and other mobile devices, but have a future in high-density storage systems as well.

The most important part of the NVMe specification is the increased performance. While the SATA bus is currently limited to 6 Gbps with SATA Express 3, which works out to about 500 MB/s of throughput, the NVMe specification supports speeds more than 5x that, with lower latency as well. Since each NVMe SSD has its own 4x PCIe lane, multiple drives can each reach the same high levels of performance, resulting in aggregate performance that's very high indeed.

New technologies are being adopted in laptops and other mobile devices, but have a future in high-density storage systems as well.

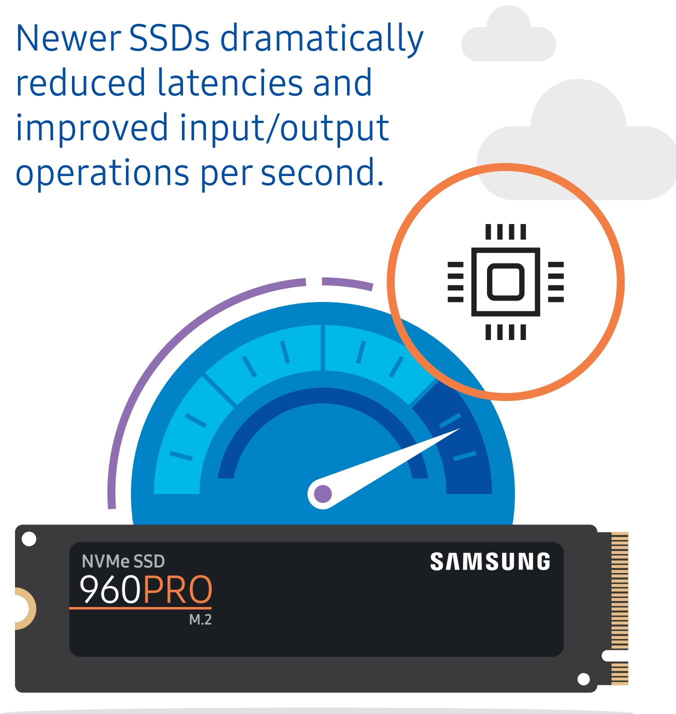


What Do These Technologies Mean for Server Performance?

Storage performance is one of the major bottlenecks of server performance. Unlike PCs, which generally have only one user at a time, servers have dozens, hundreds or even thousands of users simultaneously accessing data on the system. This leads to contention for access to the storage system, and even the fastest hard disk winds up with queues of applications waiting for data to come from storage. SSDs, with both faster throughput and lower latency, reduce the time apps must wait for data to be delivered by anywhere from 5x-100x, depending on the application and type of data.

Overall Performance

Newer SSDs dramatically reduced latencies and improved input/output operations per second.



Transfer rates tend to get the most attention, because it's a single number that's easily understood, and going from 500 MB/s to 2500 MB/s is a great advance. However, newer SSDs also have dramatically reduced latencies (generally measured in microseconds) and improved numbers of IOPS. These numbers can vary widely even within a single device, depending on the size of blocks, whether the operation is a read, a write or a combination of both and whether the blocks are being read sequentially or from random places on the device.

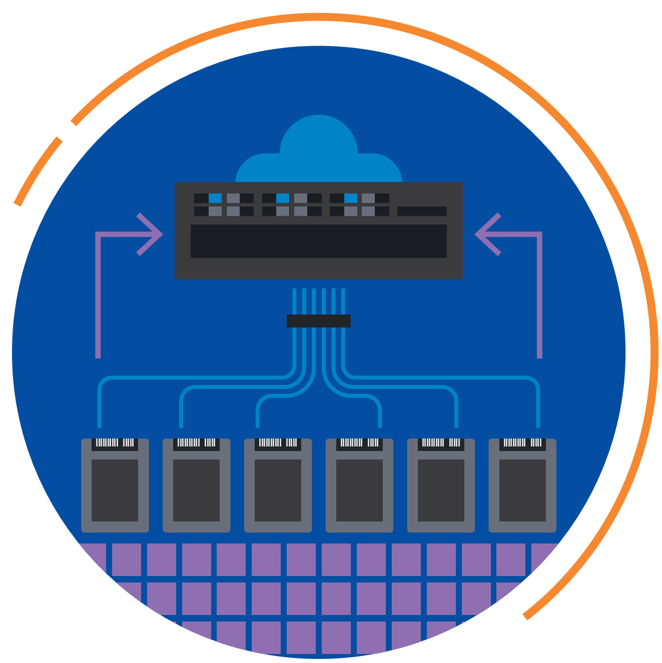
IOPS and latency can have major effects on some applications, and little on others. If a user opens a file, they generally won't know whether data started moving in a millionth of a second or three ten-thousandths of a second. However, an application that is collecting and processing data from hundreds or thousands of locations simultaneously will be greatly affected if the response time drops by a factor of 100x. Virtual desktop interface (VDI) applications, which emulate dozens or hundreds of PCs, typically can use all the IOPS they can get, since they're servicing requests from dozens to hundreds of users at once. Server Virtualization can have the same issues since it tends to consist of many different server applications all running at once.

Higher Capacities

Higher capacities can also improve performance: Keeping data local on the server rather than sharing it across a large cluster greatly reduces the time it takes to access it. This is because the internal connection to data is hundreds or even thousands of times faster than going across the network.

Large databases are typically clustered — not so much to improve reliability or to add processing horsepower, but because they typically need more RAM than a single server can hold in order to process terabytes or even petabytes of data at once. In particular, indices of the metadata are usually kept in RAM since they can heavily impact overall system performance. However, even large servers are limited by two factors: the total number of DIMM slots and the costs of RAM. Higher-capacity RAM modules tend to cost much more than smaller ones — even though a server could theoretically hold 16 256 GB DRAM modules, this would cost four or six times what 16 128 GB modules would cost, and would still only be 4 TB of RAM.

By keeping these files in NVMe-based SSDs, enough capacity can be configured to keep all of the indices and metadata on a single server. Even though the performance is less than that of DRAM, the gains from not sending data over the network to other servers in the database cluster may be sufficient to improve overall system performance. Additionally, it's at a greatly reduced cost compared to keeping all of the data in DRAM.



Total Cost of Ownership

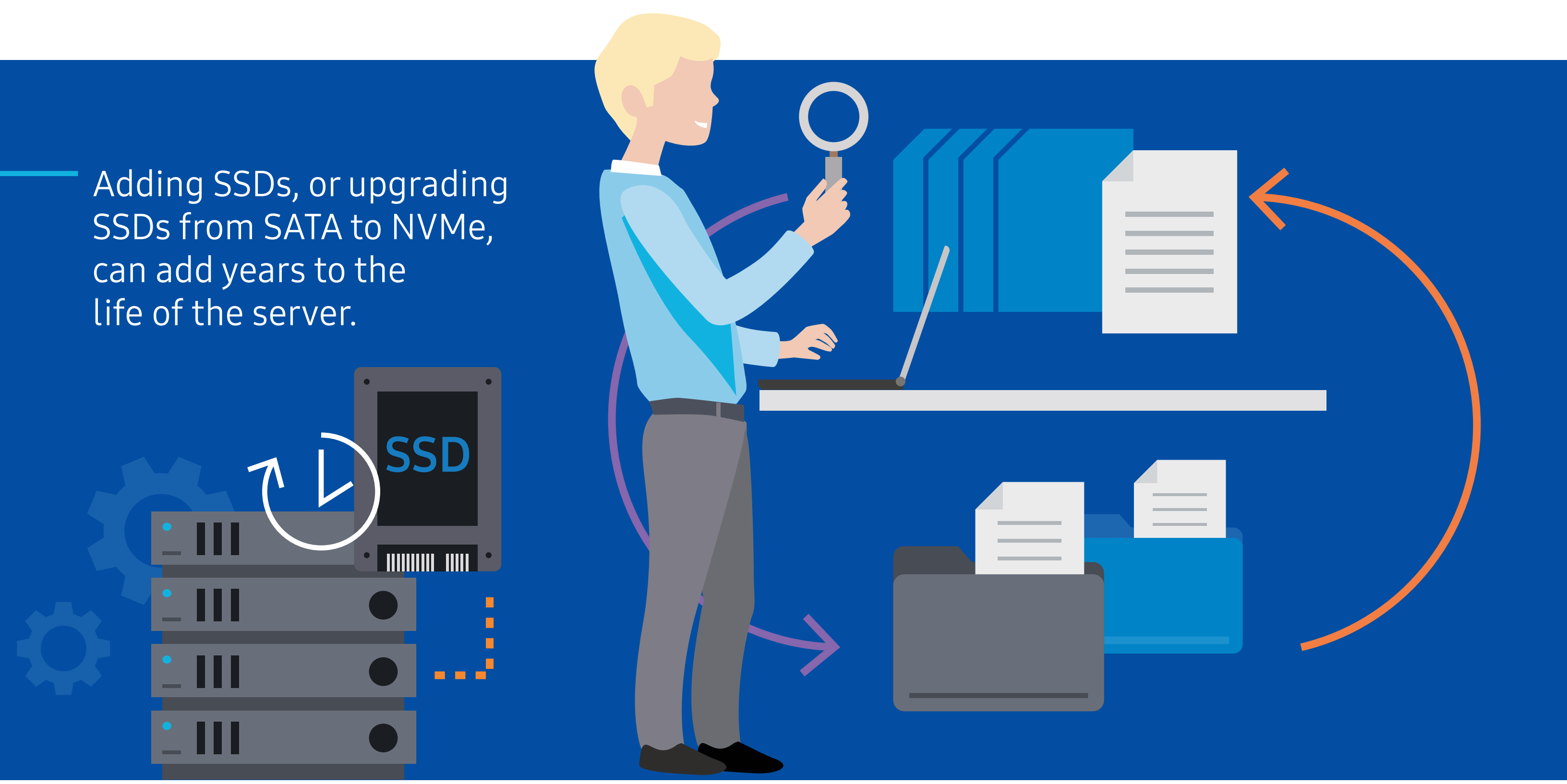
SSDs are fast enough that the data stored on them can be duplicated, compressed or both, resulting in near parity with HD costs. Compression shrinks the size of files being stored, while deduplication removes extra copies of files or parts of files.

The increases in performance that can be gained from newer, faster SSDs may allow organizations to put off server refreshes. Since the move to 64-bit CPUs came a decade ago, and multicore CPUs made the processing speed of a single CPU less of an issue. Now, servers are seldom upgraded because they've reached the capacity of their CPUs. Many aging servers are still only reaching 15-20 percent CPU utilization. In addition, many applications are not CPU-bound, waiting for the next task to complete processing through the CPU, but are I/O bound, waiting for the next piece of data to come from the storage subsystem. Therefore, it's more typical that storage speed or capacity is nearing its limits in aging servers. Adding SSDs, or upgrading SSDs from SATA to NVMe, can increase performance sufficiently enough to add years to the life of the server, at a cost much less than having to purchase a new server. Additionally, other new technologies, such as inline deduplication, can offset the cost of new SSDs, allowing effective compression on active storage without impacting performance.

While the cost per gigabyte of raw capacity is greater for SSDs than for hard drives by 4-10x, SSDs are fast enough that the data stored on them can be duplicated, compressed or both, resulting in near parity with HD costs. Compression shrinks the size of files being stored, while deduplication removes extra copies of files or parts of files. For example, if a number of documents all contained the same graphic, each file would be modified so that a small pointer was left, indicating that the removed data could be found at another location.

This can be especially powerful in the case of VDIs or server virtualization, where dozens to hundreds of virtual disks can be created, each containing the operating system files and applications needed. If there are 100 virtual disks containing all the Windows 10 operating system files, nearly all of those files will be identical on each virtual disk. Using deduplication can reduce the space required for 100 virtual disks to little more than the space required for one.

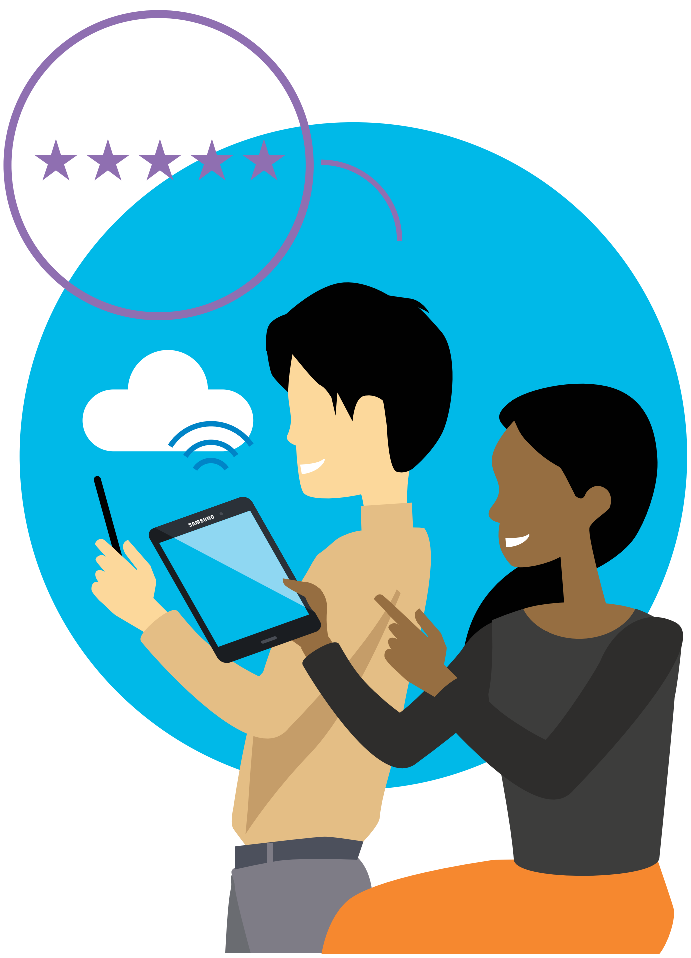
Adding SSDs, or upgrading SSDs from SATA to NVMe, can add years to the life of the server.



Heat and Power

Both the overall power and the heat generated by SSDs is less than HDDs by about half, and SSDs when idle, use about an eighth of the power needed by HDDs.

SSDs also use less power than hard drives, especially when they're in standby mode when data isn't being actively accessed. Both the overall power and the heat generated by the SSDs is less than HDDs by about half, and, when idle, SSDs use about an eighth of the power needed by HDDs. Over time in large data centers, this can add up to a considerable amount. This is especially the case where loads in the data center go down outside normal business hours, as many HDDs never enter idle mode since earlier versions of idle proved problematic.



Increased User Satisfaction

Improving the storage system by 5-10x can add life to a server and improve user satisfaction without the expense of all-new server hardware.

Generally, users look for applications that don't make them wait, and SSDs deliver sufficient performance to reduce wait times to negligible levels and keep them there for years to come. Since many applications involve waiting for storage rather than waiting for CPU processing, improving the storage system by 5-10x can add life to a server and improve user satisfaction without the expense of all-new server hardware. It can also reduce the time and expense required to upgrade the infrastructure and migrate applications to the new server.

Conclusion

The rate of progress in SSD technology has ramped up from the desultory pace of improvements in HDDs to a rate as fast or faster than the progressions in CPUs. This means that going even a year without evaluating the new technologies will cause organizations to miss out on substantial improvements in storage. This is not only in storage performance for its own sake, but in increased efficiencies that can lower data center operating costs by reducing power and cooling expenses and extending the life cycles of server hardware. Increasing user satisfaction and improving measurable application performance at the same time will make the administrators look good, and any new hardware bought for testing can be used afterward to improve existing servers. The rapid rate of progress in SSD technology is something that companies can't afford to ignore; testing and evaluation once a year should be a minimum.

Learn more: samsung.com/enterprisessd

Footnotes

1. SSD Prices Plummet Again, Close in on HDDs, Lucas Mearian. PCWorld. March 3, 2016.

http://www.pcworld.com/article/3040591/storage/ssd-prices-plummet-again-close-in-on-hdds.html

2. Grueling Endurance Test Blows Away SSD Durability Fears, Ian Paul. PCWorld. December 5, 2014.

http://www.pcworld.com/article/2856052/grueling-endurance-test-blows-away-ssd-durability-fears.html