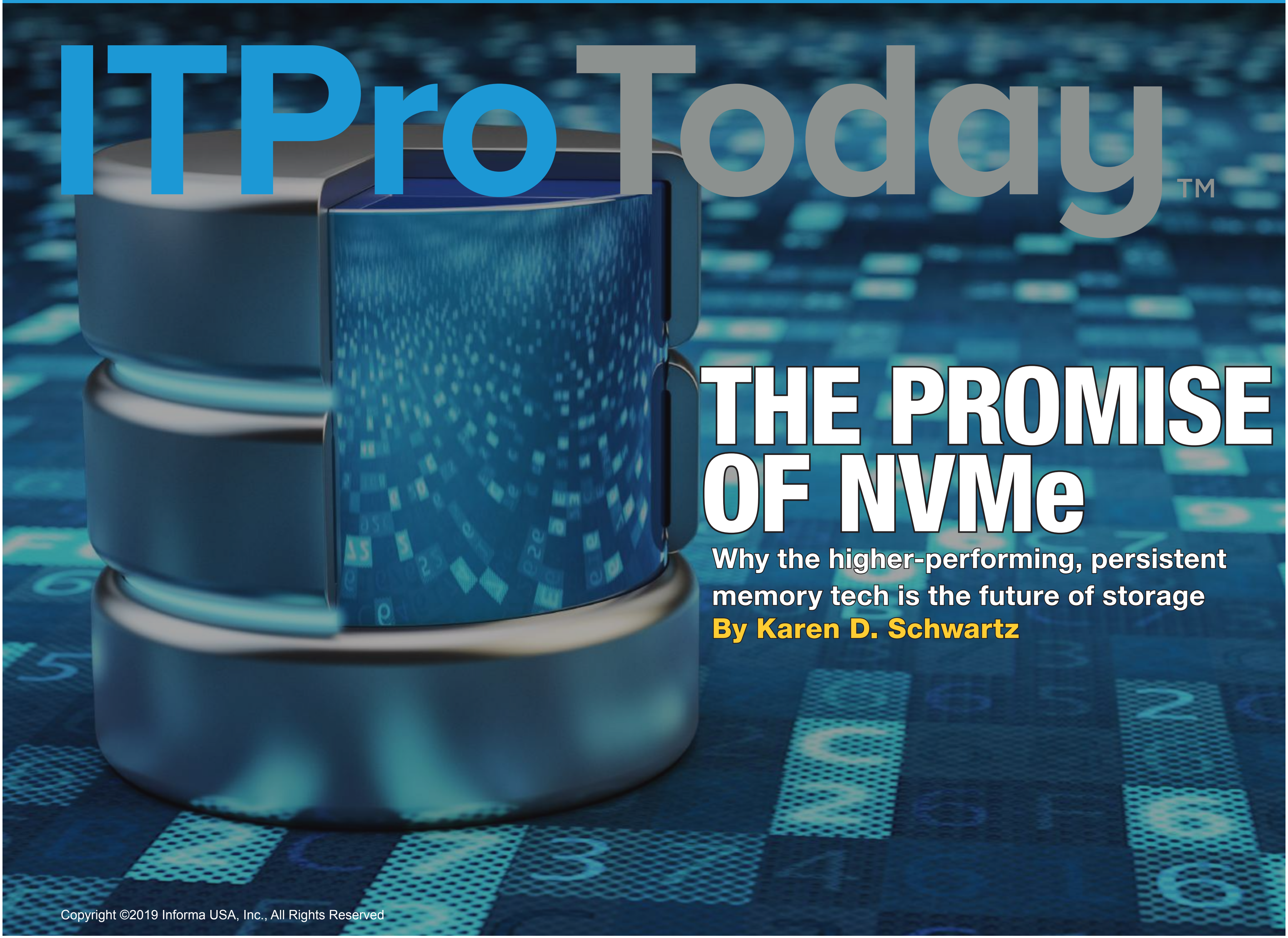
ITProtoday.com presents

Vol. 2, Issue 5 May 2019

ITPro Today™



THE PROMISE OF NVMe

Why the higher-performing, persistent memory tech is the future of storage

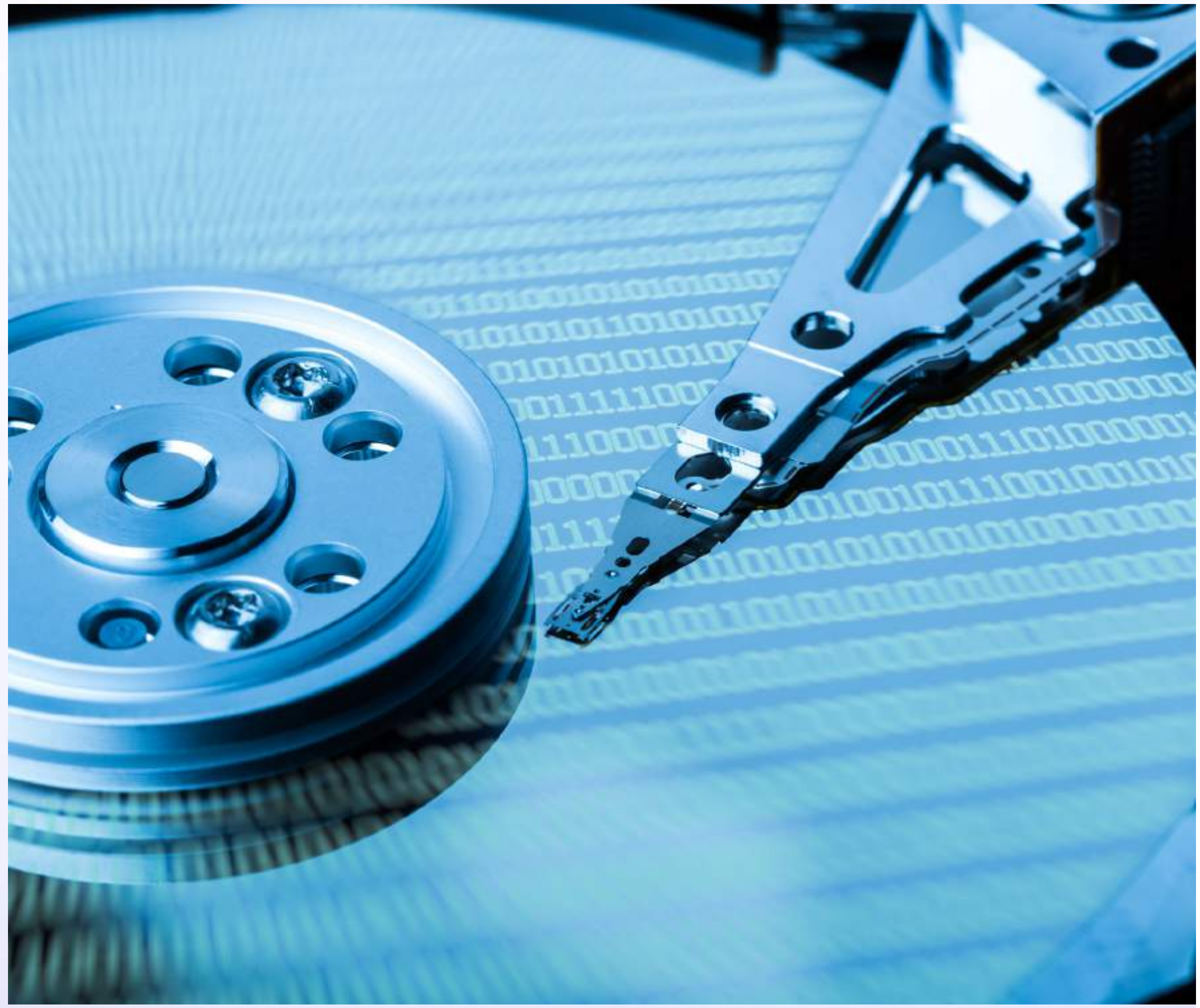
By Karen D. Schwartz

TABLE OF CONTENTS

The Promise of NVMe

WHY THE HIGHER-PERFORMING, PERSISTENT MEMORY TECH IS THE FUTURE OF STORAGE

In the 1980s, data storage took a big leap forward with innovations like the CD and 3.5-inch floppy drive. In the '90s, the big innovations were optical disk, flash technology and rewritable CDs. Those advances were critical in their time to address the increased use of compute resources and the resulting explosion in data requiring storage.



As history tends to do, it continues to repeat itself. One often-cited statistic from IDC predicts that the amount of data worldwide will reach 175 zettabytes by 2025. To put that into perspective, one zettabyte equals one billion terabytes. IDC expects about half of that data to restore on-premise, with an equal amount stored in the cloud.

That's a lot of data—data that increasingly is being used in real-time workloads, for big data analytics, and in applications that involve artificial intelligence and machine learning. All of these applications require higher throughput, lower latency and greater scalability than ever before.

So what about SATA? It was designed for mechanical hard drives, and could only transfer data up to about 6GB/s. This type of throughput began causing bottlenecks for many organizations several years ago, and it's just getting worse. As organizations create more data and require faster, more scalable storage, they will probably move on. That's why most SSD vendors decided several years ago to stop investing in future generations of SATA.

Enter NVMe. The protocol, which stands for Non-Volatile Memory Express, is a series of standards that vendors are using to create higher bandwidth, lower latency storage, typically in the form of solid-state drives, all-flash arrays, fabrics and servers. According to NVME Express, the consortium responsible for developing the NVMe specification, is designed to provide efficient access to storage devices built with non-volatile memory, from today's NAND flash technology to future, higher performing, persistent memory technologies.

> "NVMe solid state drive has a read speed six to seven times faster than a comparable SATA device."

Why NVMe?

According to research from IDC, most storage professionals believe NVMe will increase overall throughput, increase storage security, reduce overall latency and increase IOPs. Clearly, NVMe holds a lot of promise. Here is how it is helping to achieve storage professionals' goals.

Better scalability to meet performance requirements over time: According to G2M Research, nearly 80 percent of respondents believe that current processing and storage architectures won't be able to handle the amount of data in their industry in the next five years. Think about it this way: SAS and SATA each have one queue, which allows for one thread. In SATA, that allows up to 32 commands, while SAS enables up to 256 commands. NVMe, on the other hand, allow for up to 4,000 queues, with 4,0000 commands per queue. That increases the scalability of CPUs, fabrics and other devices exponentially.

Higher throughput: While SAS- and SATA-based storage controllers running flash storage can handle only one command at a time, NVMe storage controllers, with 64,000 queues available, can handle many times that, reducing bottlenecks. In addition, today's mainstream NVMe solid state drive has a read speed six to seven times faster than a comparable SATA device, along with a write speed more than five times greater. According to research from Enterprise Strategy Group, NVMe SSDs can reach throughput rates of 32 GBps, with higher-end drives reaching up to 10 million IOPs.

And that's running PCI Express (PCIe) Gen 4. When PCI-E Gen 5 arrives sometime later this year, the NVMe SSD will scale with it. Jonmichael Hand, NVMe Express Marketing Work Group Chair and a strategic planner for Intel's Non-Volatile Memory Solutions Group, expects PCI-E Gen 5 to double the current speed. And it's only going to increase from there.

"They have already shown up to 4.6 million IOPS on a single Xeon core," Hand says. "Just two years ago, you could do two million IOPS with SATA and it would take 100% CPU on a dual socket, 28-core Xeon server. That means we're looking at NVMe being 50 or 60 times more efficient as far as IOPS for CPU utilization."

Lower latency: Next-generation generation memory technologies may have read access latency of under one microsecond, requiring a streamlined protocol that enables an end-to-end latency of under 10 microseconds, including the software stack. Today, the fastest SATA SSDs have an average latency of about 100 microseconds, while NVMe-based SSDs have an average latency of less than 10 seconds. That's because all of the information for an NVMe command is contained in one 64-byte command and there are no unchangeable register reads. In contrast, SATA includes a storage controller along with register updates—both of which case latency. The upshot is that the core protocol is more efficiently designed.

Strong security: NVMe works well with industry-defined security specifications and security standards, including Trusted Computing Group OPAL (for self-encrypting drives) and eDrive. The protocol includes trusted send and receive commands to tunnel security commands. Both devices and software must manage the keys and encryption schema, but NVMe has everything in place to make it all work.

The Flavors of NVMe

There are three core parts of NVMe: the NVMe Express base specification, NVMe over Fabrics, and NVMe-oF/TCP. A fourth piece, the NVMe Management Interface, helps pull everything together.

> "NVMe works well with industry-defined security specifications and security standards, including Trusted Computing Group OPAL (for self-encrypting drives) and eDrive."

The NVM Express base specification is the industry standard for PCI-E solid state drives. It contains the queueing mechanism and transport commands necessary for PCI-E SSDs to communicate with an operating system. It works with a variety of form factors including M.2 (solid-state drives typically used in slim laptops and some desktop motherboards), U.2 2.5-inch drives that operate over PCIe, and SSD add-in cards that operate over PCIe. Drive capacity ranges from 128Gb to 4TB.

NVMe supports all major operating systems and works with everything from low-power mobile devices to full-power enterprise devices. Most major storage vendors today have NVMe-based solid state drives, including Toshiba, Huawei, Seagate, Micron, Intel and Samsung.

NVMe over Fabrics (NVMe-oF) is a common architecture supporting a wide range of networking fabrics. A fabric is basically a way for network nodes to connect to network switches. At a basic level, NVME-oF extends NVMe outside of a computer or server, providing distance connectivity to both individual NVMe devices and storage systems with no more than 10 microseconds of additional latency over a native NVMe device inside a server, according to NVM Express. NVMe-oF can transport data within fabrics over RDMA, RoCEv2, iWARP, InfiniBand and, FCoE (Fibre Channel over Ethernet). Benefits include faster, more efficient connections to servers and performance improvements, Many vendors offer NVMe-oF compliant products including Toshiba, Cisco, Pure Storage, Mellanox, Broadcom, QLogic, Emulex and Brocade.

NVMe-oF/TCP is a transport binding for the Transmission Control Protocol (TCP), a standard for how to establish and maintain network communications as application data traverses a network. Recently ratified, NVMe-oF/TCP is designed to connect NVMe-oF hosts and NVMe-oF controller devices. Typically, NVMe-oF commands are sent over standard TCP/IP sockets, with each NVMe queue pair mapping to a TCP connection. It can be layered over existing software-based TCP transport implementations and typically has much lower latency levels than existing storage networking protocols. In fact, experts tend to think of NVMe-oF/TCP as a replacement for iSCSI. When mounting remote storage, for example, using NVMe-oF is two to three times more efficient than iSCSI. It's also seen as a good way to move away from the complex Remote Direct Memory Access (RDMA) protocol. Vendors providing NVMe-oF/TCP-based products today include Toshiba, Solarflare and Newisys.

The NVMe Management Interface (NVMe-MI) is an infrastructure to help manage NVMe-based devices in areas like with configuration management, problem diagnosis, firmware updates, namespace management, inventory, and security management. For example, it is commonly used to foster communication between SSDs and in-band host software. It also supports out-of-band management. It can also discover NVMe devices and their capabilities, store data about the host environment and make it available to the management controller, monitor the health and temperature of NVMe devices and preserve the security of at-rest data. The interface is system-agnostic.

Practical Uses for NVMe Today

There are many ways that organizations have begun exploiting the power of NVMe today. It's ideal for high-performance workloads that require very low latency, consolidating workloads, managing data portability, and much more. Here are some of the most important.

Modern application processing: Applications today must provide real-time interactions and experiences to end users. This means that applications need faster storage that can read and write a tsunami of data very quickly. Legacy storage technology simply cannot keep up with the performance demands imposed by new application environments.

Real-time workloads: Real-time workloads that are sensitive to latency, such as transaction processing, are an ideal use case. These workloads typically require high bandwidth and low latency. Traditionally, organizations have used SAS and SATA in these environments, but they can bog down performance. The higher bandwidth and lower latency of NVMe provides much more responsive performance. This can apply to virtually any sector—e-commerce, finance, life sciences, energy and more.

Big data analytics: Analyzing vast and constantly growing data sets requires fast, efficient access to data. Without the right capabilities, frustrating I/O bottlenecks can slow processes to unusable levels. By incorporating NVMe over Fabric, big data analytics can be processed much more quickly, because potential I/O bottlenecks in indexing and search operations are reduced dramatically.

Relational databases: Relational databases like OLTP, Oracle, MySQL and Microsoft SQL Server often have performance and scalability issues, especially as they grow. Because NVMe flash storage improves scalability, reduces latency and improves performance, fewer physical devices are required. PCIe flash SSD cards, in particular, can vastly improve the write-through cache of relational databases, because hot files, indices and metadata can all be placed into the SSDs as cache. Data also can be pinned there. This can accelerate databases by as much as an order of magnitude when it comes to queries, sorts, and calculations, Hand said.

Artificial intelligence and machine learning: Storage bottlenecks are a very real problem for applications that involve AI (the analysis of data in real-time) and machine learning (the analysis of data at rest), partly because they involve very large data sets and partly because they require quick response to multiple queries. That's why NVMe is such an important and timely advance in this area. Because NVMe supports both high bandwidth and low latency, it's ideal for these applications. Most importantly, its support of multiple I/O queues allows it to process multiple requests simultaneously.

> "NVMe supports all major operating systems and works with everything from low-power mobile devices to full-power enterprise devices."

More efficient workload consolidation: Because NVMe is much denser in terms of performance than SCSI, it's possible to replace a SCSI-based array that takes up 30U of rack space with an NVMe-based array that takes up only 12U of rack space. The result is more efficient performance in much less space. The other benefit is reduced latency. Because NVMe supports a much higher number of queues, along with more commands per queue, it can handle larger and more varied workloads simultaneously.

High-performance computing: Whether it's automating stock trading, editing video, predicting and tracking storms or identifying fraud, high-performance computing requires extremely low latency, high performance, fast compute and storage. The characteristics of NVMe can accelerate workloads, allow more jobs to run simultaneously, and reduce the number of processing cores necessary to complete the tasks.

Managing data mobility: According to G2M, more than half of respondents expect the movement of large data stores between storage systems, storage devices, and servers to be a significant problem for their organization either today or within the next 12 months. Many also believe data movement will impact server, networking, or storage costs, as well as application performance. The higher bandwidth and higher capacity NVMe SSDs mean data is read faster, can be moved in and out of storage more easily

Building the next-generation data center. While much of data storage is going to the cloud, data centers aren't going away any time soon. Including NVMe-based storage in the data center opens up a lot of possibilities, such as using storage-class memory as a cache. With NVMe in the data center, data-intensive applications will get the performance they need. For example, all-flash arrays that combine NVMe solid state drives and NVMe over Fabrics can vastly improve response times. According to NetApp, it can result in running 60 percent more workloads on a single system, with data access of up to 50 percent faster than Fibre Channel protocol.

Devops: An effective develops environment requires small, fast releases and continuous monitoring of changes. A storage technology that allows developers to test scenarios in less time has a lot of value.

Internet of Things (IoT): Internet-attached sensors are everywhere today—in smart buildings, drones, robots, machines, equipment, surveillance cameras. All of these sensors collect data that is used to identify and analyze actions and trends. And it's growing fast; according to one recent report, there are seven billion connected devices today, and that number is expected to rise to 21.5 billion by 2025. More importantly, each of those connected devices can generate terabytes of data annually. NVMe is an effective way to aggregate data sources from IoT devices into databases and analyze them without unacceptable levels of latency.

So Is NVMe Ready for Prime Time?

The short answer is "yes". With NVMe version 1.4 well underway, the specification continues to improve, including hyperscale innovations for performance, improve data latency, congestion management, security enhancements, multi-pathing access and in-band authentication.

At the same time, experts are quick to note areas for improvement. ESG, for example, has cited the need for better heat throttling, faster transfer rates for data transfer from SSD to SSD and more advanced storage functionality.

And then there is cost. Although SSD prices are dropping, the technology is still more expensive than it should be. ESG notes, however, that most data centers can get by with mid-level NVMe drives. It advises to avoid overspending on NVMe capacity you won't use. Another factor to consider is how NVMe lowers Total Cost of Ownership by increasing virtual machine density in virtualized environments.

So should you wait? While it's reasonable to wait until end-of-life for SAS or serial-attached SCSI if they are legacy workhorses for enterprise storage in your organization, that's not true of SATA. SATA's time has come and gone. Tom Coughlin, president of storage consultancy Coughlin Associates, puts it this way: "Within just a couple of years, NVMe will displace SATA completely, at least in the enterprise space."

In fact, NVMe is coming to your organization—later, if not sooner. It's already the dominant storage interface for data center SSD, with flash drives capable of up to 32TB coming soon.

And that's only the tip of the iceberg. According to NVM Express, this year will bring even more: not only increased adoption of NVMe-based technologies in general, but new types of non-volatile memory like Quad Level Cell Flash, which Hand expects to eventually replace hard drives, and SCM (Storage Class Memory).

By the middle of 2019, NVM Express expects to add many features to the basic specification, including I/O determinism to eliminate read latency outliers without giving up high performance, multipathing, rebuild assist, persistent memory region, a secure erase extension, enhanced command retry, and more. Next up for NVMe-oF is enhanced discovery and TCP Transport Binding, while NVMe-MI will get an in-band mechanism that allows applications to tunnel NVMe-MI commands through an NVMe driver. There will also be new NVMe administrative commands, and better management for NVMe-oF within NVMe-MI.

The NVMe Timeline

March, 2011: Version 1.0 of the NVMe specification released

October, 2012: Version 1.1 released

July, 2013: Samsung release the first NVMe drive

November, 2014: Version 1.2 released

November, 2014: NVMe drives become commercially available

June, 2016: Release of NVME-oF specification

2016: Apple releases the first mobile deployment of NVMe over PCIe in smartphones

May, 2017: Version 1.3 of NVMe specification released

November, 2018: Release of NVMe-oF/TCP specification

2019: Expected release of NVMe Version 1.4

The Future of Mobile Storage may be NVMe

You may not realize it, but the mobile device you can't put down may already incorporate NVMe in some way. Apple led the way in 2015 by replacing its traditional SDIO storage with PCIe and NVMe, and followed that up by including NVMe in some of its other products. Google also has made progress, shipping most of its Chrome-based mobile devices with the storage technology. As of now, however, Android phones aren't incorporating NVMe.

The main reason mobile device manufacturers are coming around to NVMe is because of the higher performance and battery life it can bring. The protocol also can reduce latency, delivering a better quality of service (QoS) and application responsiveness. In 2016, the NVM Express organization issued version 1.3 of the specification, which focuses on the low power consumption and other features required by mobile devices.

All of this is becoming more and more important as users depend more heavily on their mobile devices for more critical applications and processes. For example, according to Statistica, there were 197 billion application downloads in 2017, and that number is expected to rise to nearly 353 billion by 2021. Many of those apps require more capacity and greater performance than previous generations. In addition, users are storing more photos (1.2 trillion photos were taken on smartphones in 2017). And according to one survey, the amount of time spent watching videos on a smartphone is growing 88 percent year over year.

In addition to smartphones and tablets, NVMe has become part of other portable media. In the past few years, vendors have introduced everything from high-performance portable SSD units to microSD Express memory cards that integrate PCIe and NVMe. The newest specifications for microSD Express memory cards can provide up to a 985 MB/sec transfer rate. These cards can be used in a variety of mobile devices to reduce latency and increase transfer speeds across applications. It's also being used to increase performance and reduce latency in edge computing, such as Internet of Things (IoT) and Industrial IoT devices.