Busting the NVMe Flash Myths

By George Crump

Storage Switzerland, LLC

Tintri

NVMe™ is the next wave of all-flash systems coming to market. These systems are fast, dense and scalable. They promise to fundamentally change the way IT designs storage architectures in support of existing applications, and they open up new possibilities for artificial intelligence, machine learning, and deep learning. As with any new technology NVMe flash storage systems come with their own set of myths, often amplified by individual vendors, which in actuality are attempting to steer customers toward their products. The goal of this eBook is to bust those myths.

MYTH 1 — Regular Data Centers Don't Need NVMe Flash

One of the most common myths propagated by vendors is that "regular" data centers don't need NVMe flash. The implication is that these environments don't have the workloads required to push NVMe flash to its limits. Proponents of this myth insist that the data center must have workloads like artificial intelligence, machine learning, deep learning or high-velocity analytics to benefit from NVMe. In actuality, there is a certain amount of truth to this first myth. Many data centers can't push NVMe flash to its limits, but there is a difference between pushing technology to its limits and getting some benefit from its use.

A case in point is laptops. Most end-user laptops today are equipped with flash drives, and most end-users don't come close to pushing those drives to their limits, but those end-users won't give up their flash drive for any reason. The reason for the loyalty to flash is not how many IOPS can the workloads create, but more about how much can the flash drive reduce latency.

NVMe drives, especially those placed inside a storage system, significantly reduce latency. The drive bus and the drives themselves are the funnel point for all IO. It makes sense then that these drives become saturated when serving IO to the diverse data center. NVMe can deliver genuine benefits to many data centers, and the technology enables them to reduce latency.

The Use Cases for NVMe in the Regular Data Center

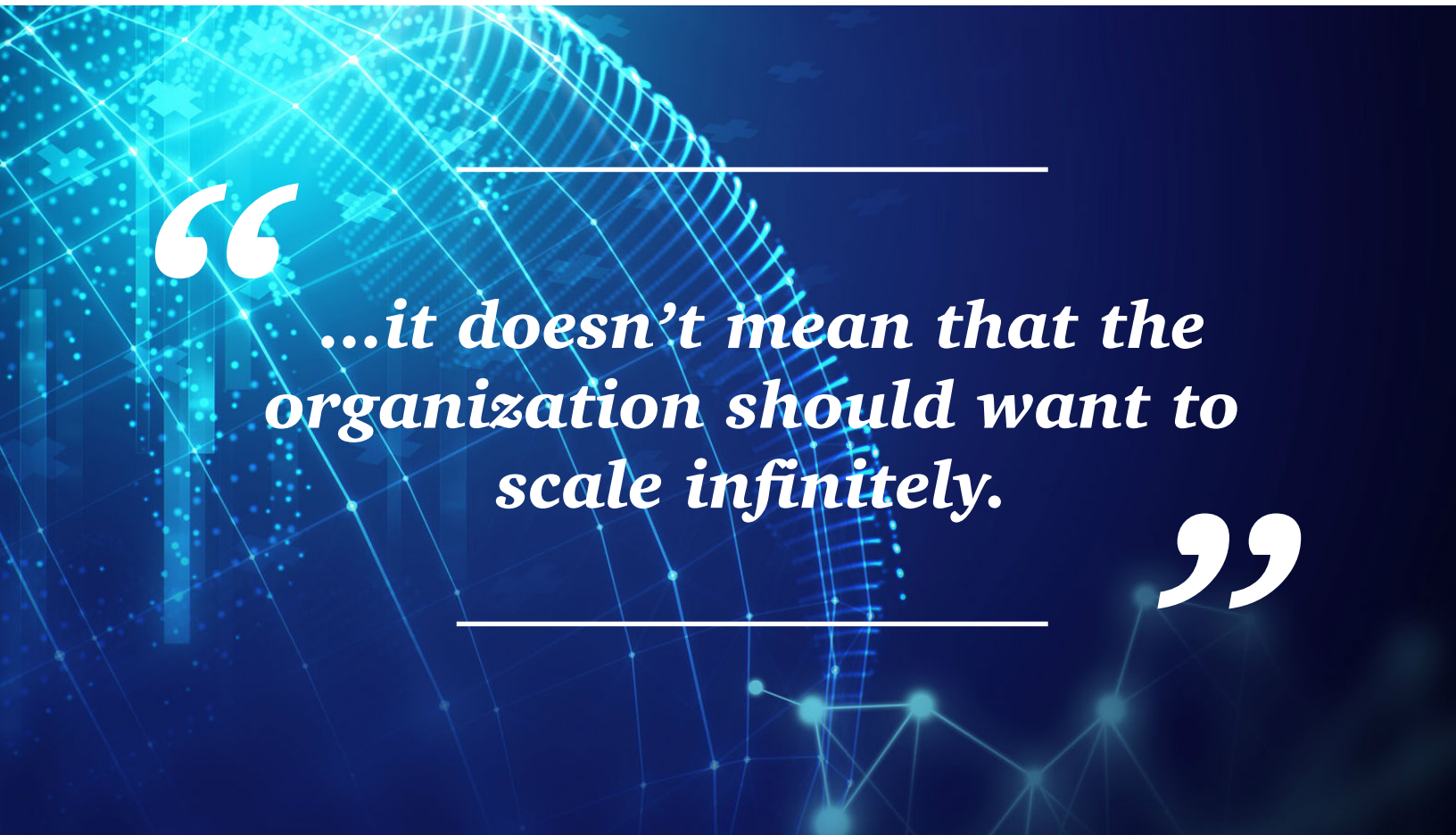
The first use case for NVMe Flash Storage Systems in the "regular" data center is scalable databases like Oracle®, MS-SQL, and others. These environments often use multiple application instances and even multiple storage systems. NVMe flash storage systems enable the organization to make the servers running these instances work harder and utilize more of the available CPU resources. If the storage system can push the physical server hardware, the outcome is a reduction in the number of physical servers and database application licenses to purchase.

The second use case is high-density VMware®. While VMware and the hypervisors scale by adding physical nodes to the cluster, it doesn't mean that the organization should want to scale infinitely. Each additional node in the cluster costs money. The fewer nodes the organization must purchase the better. Once again, the challenge is making sure the storage system can keep up with the IO demands of the virtual machines, which enables IT to increase the density of VMs per nodes. The fewer nodes required, the lower the hardware acquisition cost, power and cooling costs, and the software licensing cost.

The third use case is high-performance file systems. Most file IO today is still based on either SMB or NFS, especially for performance sensitive workloads. NVMe enables the SMB or NFS based file system to respond almost instantly to a wide variety of IO requests, ranging from creative design, test and development to analytics processing.

The fourth use case is to build a pathway to tomorrow's use case. Most organizations are set to embark on an Artificial Intelligence, Machine Learning, and Deep Learning journey. Many have their first projects already up and running. Instead of buying an entirely new storage architecture just for these use cases, an NVMe storage system can serve as the building block for these new projects as they develop and mature.

NVMe Flash Systems are available now, and for most data centers it makes sense to invest in something they can use now to increase workload density because of its raw performance capabilities. The modern flash array investment should last at least five years. Flash vendors have warranty programs that accommodate this new reality. The primary limiter to that longevity is no longer capacity but performance. NVMe flash media almost means that the organization shouldn't hit the performance limit anytime soon.



MYTH 2— NVMe Flash Has to Be End-to-End

One of the advantages of the NVMe Flash standard is that it is networkable. NVMe over Fabrics (NVMf) brings direct attach storage like latency to shared storage systems. NVMf eliminates the need for modern applications like Apache Hadoop® and others to attempt to reduce latency by using direct-attached storage with replication. NVMf however, enables organizations to realize the latency reduction of NVMe without losing the benefits and efficiencies of a shared storage system. The promise of NVMf is so great that some vendors suggest that an end-to-end NVMe infrastructure is a requirement for moving to NVMe.

Requirements for NVMf

For an organization to create an end-to-end NVMe infrastructure it needs an NVMe host bus adapter (HBA) in each server, an NVMe ready network switch and a storage system with not only NVMe flash drives, but also NVMe host bus adapters. There are three problems with these requirements today. First, they are expensive. Most NVMe ready products carry a premium price. Second, while the switch vendors have done an excellent job of providing simultaneous access to legacy protocols and NVMf, it is a different environment with different HBAs, so there is a new learning curve. Third, there is a limited degree of compatibility between vendors. While working configurations are possible, the customer is generally very restricted as to which HBAs they can use with each operating system.

Is NVMf Worth it Today?

NVMf is the protocol of choice for the future, but for today, IT planners should proceed with caution, especially given the above concerns, and given the reality that most organizations only use a small potential of their current network's capabilities. There is also room to grow the current network's capabilities as higher bandwidth Ethernet and Fibre Channel network switches and HBAs are now available. While bandwidth doesn't always help latency, it still improves performance as more servers, and virtual machines can send more I/O across the same network. The challenge with the ever-improving network and the increasing server computing power is that I/O eventually bottlenecks at the storage system.

NVMe Flash Arrays

What's needed first for most organizations is a high-performance all-flash array that is native NVMe throughout. This type of NVMe system solves the bottleneck of many servers and virtual machines all sending I/O to a single storage system. It also enables the organization to use existing network infrastructure and networking protocols. Additionally, it enables a gradual transition to NVMe that allows application and workloads to run without disruption.

Vendors providing these systems need to make sure their storage software takes full advantage of NVMe, and they likely need to tweak that software to better integrate with NVMe. They also need to provide more powerful processors in their storage systems and make sure their storage software can take advantage of those processors so that I/O can move quickly through the storage system. Finally, they should continue to provide the enterprise features that users have come to count on like SAN/NAS support, snapshots, replication, deduplication, and compression.

—

There is little doubt that over time most data centers will move to NVMf but that transition needs to be gradual and can't disrupt current applications and workloads. During that transition, IT does need to address performance bottlenecks and more than likely an NVMe Flash Array solves most of those issues.

MYTH 3—
NVMe Flash Has to be Featureless?

The primary goal of the NVMe standard is to reduce latency. It accomplishes this by leveraging the PCIe interface and increasing both command counts and queue depth. NVMe Flash Systems however, are more than just the flash media. Vendors create these storage systems out of computing resources, networking resources, and storage software to control how data flows to and from the NVMe flash storage. The NVMe flash is a component within an ecosystem that creates a storage array. Each of the components can impact latency, potentially negating the latency reductions that NVMe offers.

The Storage Software Challenge

A primary point of concern is the storage software itself. The first challenge is the efficiency of the storage software. The software needs to be optimized for today's modern multi-threaded processors. The features it provides must also be efficient so that the capabilities offered by NVMe flash to reduce latency, which enterprise IT has come to count on, can be fully realized.

Vendors have taken several approaches to work around the latency introduced by the storage software. One of the more common is to offer featureless storage software that only provides the most basic storage functions like volume creation and management. These vendors' stance is that modern applications such as Hadoop and NoSQL databases have the most common storage features built in, so IT doesn't need those features in the storage software. The problem with this approach is that most data centers have not cut over 100% to modern applications and most never will be 100% modern applications.

The second challenge is forcing the application to perform storage management functions and deliver storage features. This means the featureless array diverts some CPU power, intended for the processing of the application, to storage management functions. The organization ends up over-buying computing power for the application, and it is much more difficult to deliver consistent performance.

The third challenge is that while it is reasonable to assume that feature-rich storage software adds somewhat to latency, IT planners can't ignore the reality that most organizations can sacrifice some latency in return for a richer feature set and never notice a performance impact.

Striking a balance— Features and Low Latency

For most organizations, the better choice is to sacrifice some latency in return for full-featured performance. The storage vendor needs to optimize its software for today's multi-threaded processor designs and make sure their features such as deduplication, compression, snapshots, and replication impact performance as little as possible.

A full-featured NVMe flash array might introduce some latency, but the organization can quickly realize a gain by its introduction. First, the computing power dedicated for the applications can remain dedicated. Second, most of the features in a new NVMe all-flash array are there to either protect data from media failure or to reduce operational costs. If the organization is using a featureless NVMe flash array and counting on the application to provide the missing features, they often see an increase in application processor consumption and in flash capacity consumption since most applications use three-way replication instead of a RAID-based data protection method.

Finally, a featureless NVMe flash array might promote additional storage sprawl since it more than likely is a block-only device. This type of NVMe flash array also needs a full-featured file system to support artificial intelligence, machine learning, and high-velocity analytics. A full-featured system may also be multi-protocol, which means it can provide an NFS file system for AI and ML workloads while at the same time being the storage solution for more traditional workloads such as Oracle and VMware through either NFS, iSCSI or Fiber Channel (FC) protocols.

MYTH 4—
NVMe Flash is the End of SAS All-Flash

Before NVMe, most flash systems were either serial attached SCSI (SAS) or Serial Advanced Technology Attachment (SATA) based. In all-flash arrays, SAS was the preferred protocol because of its higher bandwidth and ability to support multi-port connections. NVMe is the logical upgrade to these technologies because of its higher command count, greater queue depth, and native PCIe connectivity; however, NVMe does not replace SAS. Today SAS flash drives are still less expensive than their NVMe alternatives, and most of the high capacity and ultra-high capacity flash drives are SAS based.

The Role of High Capacity Flash

Several vendors are committing to deliver 50TB+ high-density flash drives in 2019. A 24-drive flash array, using 50TB drives, delivers 1.2PB of raw capacity and potentially over 1PB after protection. Using high-capacity flash drives, organizations can achieve Petabytes of capacity in just a few rack units. These systems are ideal for active archive use cases where extreme performance isn't required, but rapid access is.

The challenge for IT is how to leverage these high capacity flash drives without increasing the IT burden of having to manage separate systems. IT needs to not only look for vendors that can provide a choice of NVMe or SAS based flash arrays but also the ability to use the same storage software to manage those systems. Many vendors offer the option of SAS or NVMe flash systems, but few provide the option under a single management GUI.

Hybrid Flash

When flash first came to the data center, the most common method was in a hybrid array, which mixed SAS based flash with hard disk drives to help offset the cost of the premium-priced flash technology. The continual decline in flash pricing combined with data efficiency technologies like deduplication and compression is creating price parity between flash and hard disk systems. Price parity between the two technologies reduces the value of hybrid arrays for many vendors. Vendors with the intelligence in their software to automatically tier between two different types of storage though, can leverage that technology to mix SAS and NVMe drives thus creating a new type of hybrid flash array. The flash system can leverage the NVMe tier for high-performance workloads and the high-capacity SAS-based flash tier for other workloads or dormant data.

The tiering between SAS and NVMe flash tiers requires further development though since the software can no longer base data placement decisions solely on access patterns. The storage software now needs to understand “data intensity” to make sure accessed data can take advantage of the higher performing NVMe tier.

Object Flash

Another option is for the organization to leverage the high-capacity flash drives in object storage systems to create an active archive. Most object storage systems already support flash drives, but most can’t move data automatically between a primary storage system and an object storage system. Third party solutions exist to perform the analysis and movement of data between the two storage systems, but vendors that provide both technologies should have an advantage in providing a deeper more seamless integration between the tiers.

—

In this eBook we’ve busted several of the myths surrounding NVMe. While NVMe offers unprecedented performance to organizations of all sizes, each organization needs to evaluate its needs carefully. In most cases, leveraging a flash system that leverages NVMe flash while still taking advantage of traditional connectivity instead of replacing or upgrading the existing network is the best fit for the organization. At the same time, they should make sure that the features, like snapshots and replication, are still available to them when they move to an NVMe array. Finally, organizations should not rule out SAS or even Hybrid based storage systems to balance performance requirements against costs.

The Firm

Storage Switzerland is the leading storage analyst firm focused on the emerging storage categories of memory-based storage (Flash), Big Data, virtualization, and cloud computing. The firm is widely recognized for its blogs, white papers and videos on current approaches such as all-flash arrays, deduplication, SSD's, software-defined storage, backup appliances and storage networking. The name "Storage Switzerland" indicates a pledge to provide neutral analysis of the storage marketplace, rather than focusing on a single vendor approach.

About Our Partner

Based in Silicon Valley, Tintri by DDN is a wholly owned subsidiary of DataDirect Networks (DDN), the data-at-scale powerhouse and world's largest privately held storage company. Tintri solutions serve the needs of enterprise data centers with intelligent infrastructure. Our product portfolio includes advanced technologies from Tintri (acquired September 2018), Nexenta (acquired May 2019), and IntelliFlash (acquired from Western Digital November 2019). This portfolio offers a fresh choice and a new experience to elevate your user experiences through automation, analytic insight, and tight integration with application workflows for your enterprise applications, including mission-critical databases and business applications, virtualization, and flexible software defined solutions.

Experience different! For more information on how Tintri's portfolio of solutions can turbo-charge your business success through a simple, intelligent infrastructure, visit https://tintri.com

The Analyst

George Crump is the founder of Storage Switzerland, the leading storage analyst firm focused on the subjects of big data, solid state storage, virtualization, cloud computing and data protection. He is widely recognized for his articles, white papers, and videos on such current approaches as all-flash arrays, deduplication, SSDs, software-defined storage, backup appliances, and storage networking. He has over 25 years of experience designing storage solutions for data centers across the U.S.