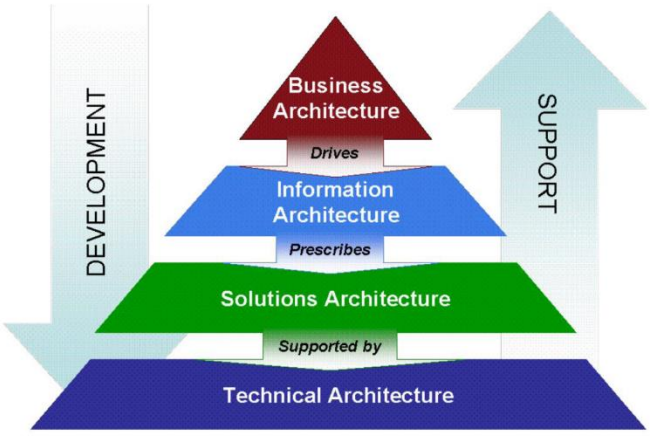
Virginia Information Technologies Agency

Enterprise Architecture Technical Brief

Non-Volatile Memory Express (NVMe)

Robert Kowalke

April 2017



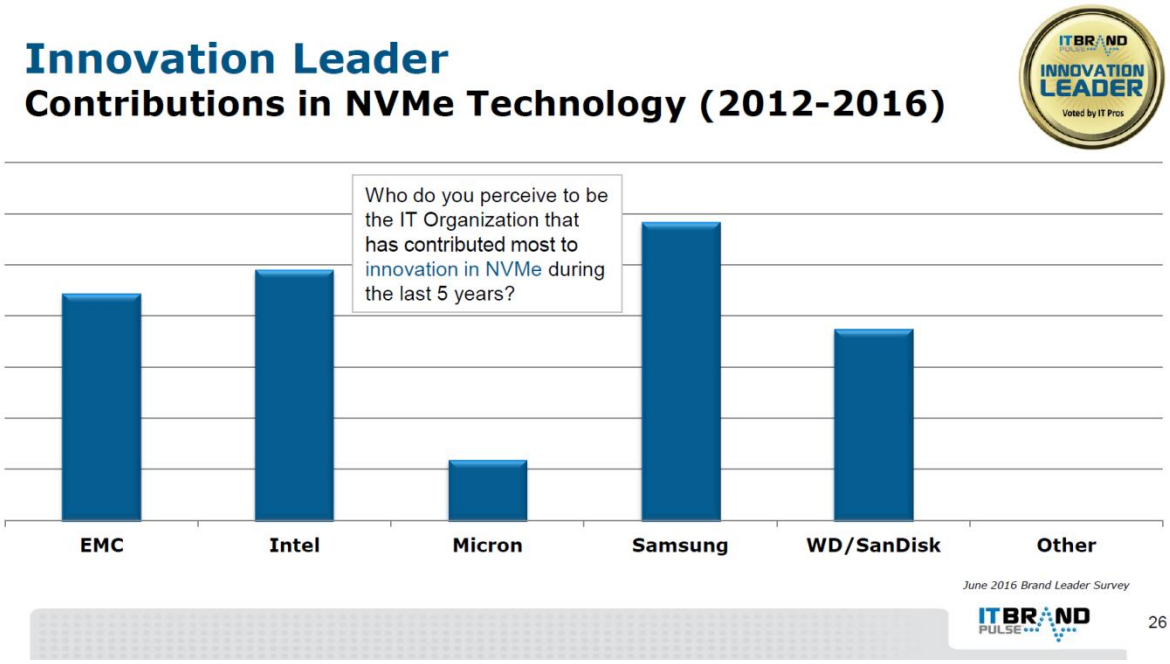
Contents

Photo courtesy of Samsung NVMe SSD 960 PRO/EVO – http://www.samsung.com/semiconductor/minisite/ssd/product/consumer/ssd960.html

robert.kowalke@vita.virginia.gov

NVMe Recommendations

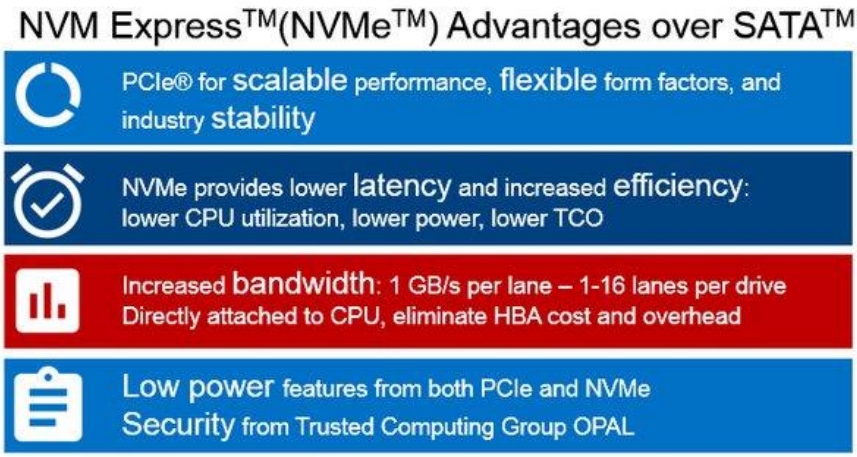
- NVMe adoption is significant among those buying storage servers and should be considered as part of your strategy around such storage efforts as:
  - database applications;
  - big data analytics;
  - as cache in front of a storage array;
  - in a software defined storage solution, or
  - Hyper-V servers
- Generally, NVMe is a good candidate anywhere:
  - traditional flash storage solutions are used, or
  - any application or usage scenario that demands high availability from the drive
- NVMe does not replace Serial Attached Small Computer Serial Interface (SCSI) (SAS) or SATA.
  - They all will co-exist for years to come, enabling different tiers of server storage I/O performance tailored to different requirements inside the same platform.
  - One will simply align applicable technology (SATA, SAS, and NVMe) to meet different performance and cost parameters.



For any comments, questions, and/or concerns with this technical brief, please contact VITA EA: ea@vita.virginia.gov

NVMe Benefits

- NVMe, unlike SCSI, was built from the ground up specifically to support solid state storage devices.
  - Streamlines the SW I/O stack by reducing unnecessary overhead introduced by SCSI.
  - Performance benefit since it does not have to support legacy protocols, meaning its vendors can focus on taking full advantage of memory based storage.
  - Compatibility advantage because there is now one software-interface standard for PCIe SSD vendors to adhere to, and they don't have to write their own.
  - PCIe is the underlying data transport layer for graphics and other add-in cards.
- NVMe is nice to have for SSDs, but it's critical for 3DXP (pronounced 3D cross point) SSDs. https://en.wikipedia.org/wiki/3D_XPoint
  - Samsung notes SSDs are steadily displacing Hard Disk Drives (HDDs) in more apps, but NVMe is shaping up to be the dark horse (https://en.wikipedia.org/wiki/Dark_horse) that may put the venerable HDD to rest.
  - NVMe has been in development since 2007, with the first commercially available NVMe chipsets released in August 2012.
  - NVMe will never work over SATA, it is PCIe only.
    - It is the optimized, high performance, scalable host controller interface designed for NVM-based storage.
  - SATA was originally designed for spinning disks, and with the advent of SSDs, the SATA protocol has become a limiting factor
- A good pictures/graphs overview of NVMe is available here → http://www.tomshardware.com/reviews/intel-750-series-ssd,4096-2.html

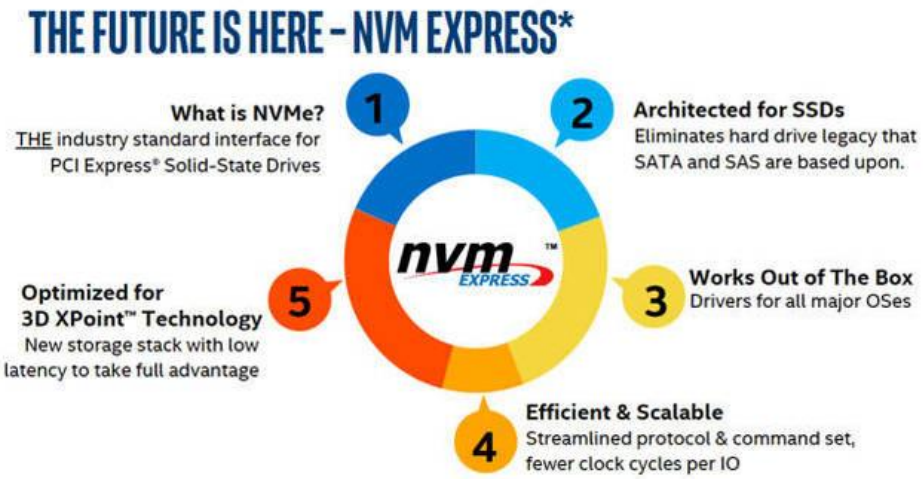


- NVMe capitalizes on the low latency and parallelism of PCIe SSDs.
  - PCIe is raw throughput requiring a standard software interface, command set, and feature set allowing datacenters to unlock the true potential of PCIe SSDs.
    - NVMe is that standard software interface for PCIe.
  - NVMe improves bandwidth to each SSD by connecting the CPU to SSDs directly over PCIe, which means no need for an intervening Host Bus Adapter (HBA), providing a greater number of PCIe lanes to be employed.
  - Considered the SSD interface for the foreseeable future.

robert.kowalke@vita.virginia.gov

NVMe Background

- NVM Express™ (NVMe) is an optimized, high performance, scalable host controller interface / protocol (firmware) designed for accessing Solid State Drives (SSDs) attached through the Peripheral Component Interconnect Express (PCIe) bus.
    - NVMe is considered the SSD interface for the foreseeable future.
    - It's a completely new architecture for storage, from the software stack to hardware devices and systems.
    - Capitalizes on the low latency and parallelism of PCIe SSDs.
- NVMe, Advanced Host Controller Interface (AHCI), and Integrated Drive Electronics (IDE) are transfer protocols (languages) running on top of transfer interfaces such as PCIe or Serial Advanced Technology Attachment (SATA).



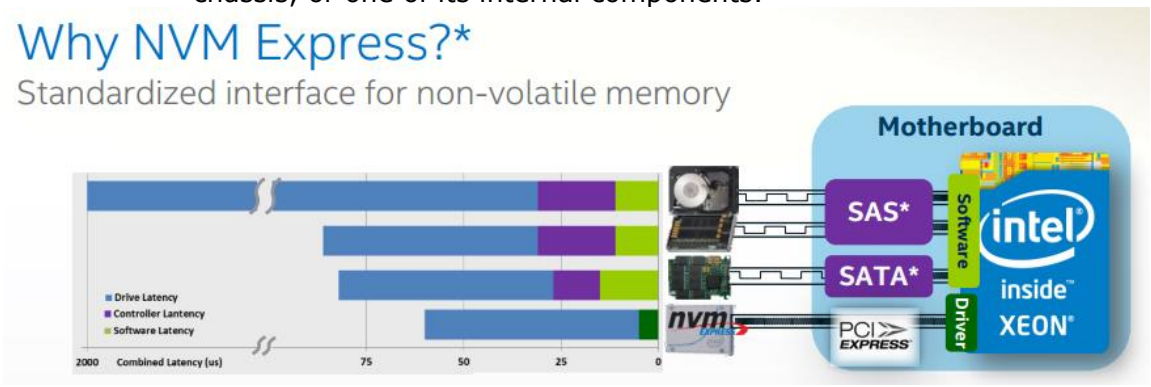
- NVMe, AHCI, and IDE are transfer protocols (languages) that run on top of transfer interfaces such as PCIe or SATA.
    - NVMe is the latest high performance and optimized protocol which supersedes AHCI and compliments PCIe technology.
        - Offers an optimized command and completion path for use with NVMe based storage.
        - Developed by a consortium of manufacturers specifically for SSDs to overcome the speed bottleneck imposed by older SATA connections.
        - NVMe is akin to having a more efficient language between storage device and PC:
            - NVMe can handle 65,000 queues of data each with 65,000 commands, instead of one queue of 32 commands, and it only has seven major commands (read, write, flush etc.).
        - Delivers better performance, reduced latency, and is scalable.
        - It's the protocol of choice for next generation storage technologies such as 3D XPoint.
    - SATA is the dominant interface for connecting an SSD to the PC.
        - It employs the command protocol AHCI, which also supports IDE.
        - Built with slower spinning disks in mind rather than flash memory.
        - SATA transfer rates begin at 150 MB/s and max out at 600 MB/s.
        - For most consumer uses of SSDs the SATA interface is absolutely adequate.

robert.kowalke@vita.virginia.gov

Note: While the fastest hard drives can offer read speeds of around 70mb/s on average, an SSD is easily able to offer read speeds of up to 2500mb/s.

- PCIe supersedes SATA as the latest high bandwidth interface.
  - Entry level PCIe SSD speeds are two to three times faster than older generation SATA 3.0 SSDs, mainly due to the number of channels contained by each to transfer data (roughly 10 for SATA and 25 for PCIe).
    - Depending on usage, real world benchmarks may not reflect this massive gain due to bottlenecks elsewhere in the PC.
  - Think of PCIe as the physical interface while NVMe is the protocol for managing NVM devices using that interface.

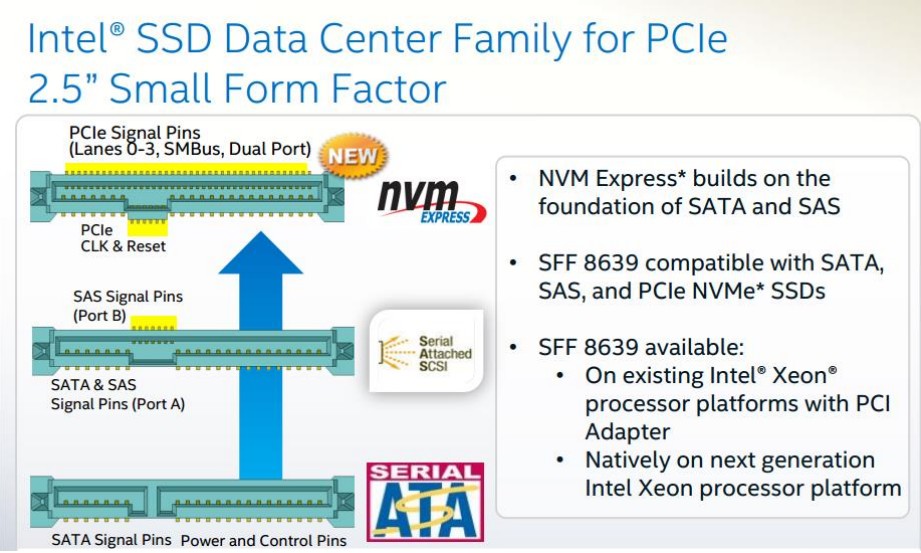
- NVMe may alternately be referred to as M.2 drives, and while technically correct, it is a bit of a misnomer.
  - From a Dell perspective, all NVMe drives utilize the M.2 form factor, but not all M.2 SSDs utilize NVMe technology – M.2 is simply the form factor.
    - Form factor in computers is the size, configuration, or physical arrangement of a computing device; the term is commonly used in describing the size and/or arrangement of a device, a computer case or chassis, or one of its internal components.



- NVMe is better for almost every application and never worse.
  - NVMe price premium over SATA has dropped significantly; as low as 10% now.
  - Performance boost is significant enough to switch if currently using SATA SSD's.
    Note: In terms of read and write speeds, SATA hard disks support several hundred megabytes (MBs) per second, while NVME supports several gigabytes (GBs) per second – one order of magnitude greater for unparalleled performance gains.
  - Lower latency due to improve drivers and increased queues (and queue sizes).
  - Lower CPU used to handle larger number of I/Os (more CPU available for useful work).
  - Higher I/O activity rates (IOPs) to boost productivity unlock value of fast flash and NVM.
  - Bandwidth improvements leveraging various fast PCIe interface and available lanes.
  - Dual-pathing of devices like what is available with dual-path SAS devices.
  - Unlock the value of more cores per processor socket and software threads (productivity).
  - Various packaging options, deployment scenarios and configuration options.
  - Appears as a standard storage device on most operating systems.

robert.kowalke@vita.virginia.gov

    - Plug-play with in-box drivers on many popular operating systems and hypervisors.
- Streamlines the software I/O stack by reducing unnecessary overhead introduced by the SCSI stack.
- NVMe brings a performance benefit since it does not have to support legacy protocols, meaning its vendors can focus on taking full advantage of memory based storage.
    - An NVMe drive can select up to 64,000 blocks at a time, whereas a SCSI based drive must select one block of data at a time.
    - PCIe is 1GB/s per lane, and with seamless multi-lane support, a vendor can easily build a single device that is 4 or 8 GB/s – substantially outrunning SAS or SATA potential.



- NVMe brings a compatibility advantage because there is now one software-interface standard for PCIe SSD vendors to adhere to, and they don't have to write their own.
    - Interface is designed specifically for memory based storage.
    - Implementation IT professionals no longer need to vet vendors based on their compatibility with a particular operating system, rather they can look at the specific capabilities and cost of the card to determine which is best for their environment.
    - NVMe form factor is a standard, supporting both PCIe cards and U.2 form factor SSDs.
        - Support for U.2 SSDs means that a storage system can be built using 2.5" SSD drives for ease of service while not having to give up performance.

robert.kowalke@vita.virginia.gov

NVMe Conclusions

- Key workloads work better with NVMe SSDs:
  - Database – NVMe shines in traditional relational databases with consistent, low latency, high bandwidth performance.
  - Big Data – Analytics and NoSQL databases fully utilize NVMe performance to provide near real time results.
  - Private Cloud – Software Defined Infrastructure is made affordable with high performance NVMe SSDs.
  - High Performance Computing (HPC) – NVMe SSDs support high bandwidth demands of HPC designed to speed up workflow times.
  - Virtualization – NVMe SSDs lower enterprise total cost of ownership (TCO) by enabling increased virtual machine (VM) scalability while optimizing platform utilization.
- Samsung at its Global SSD Summit shared its future trends noting that SSDs are steadily displacing HDDs in more applications, but NVMe is shaping up to be the dark horse that may put the venerable HDD to rest.
- NVMe is better for almost every application and never worse.
- Suits the highly virtualized data center perfectly.
- NVMe is nice to have for flash SSDs, but it's critical for 3DXP SSDs.

Flash Storage/NVMe – June, 2016

Voted by IT Pros

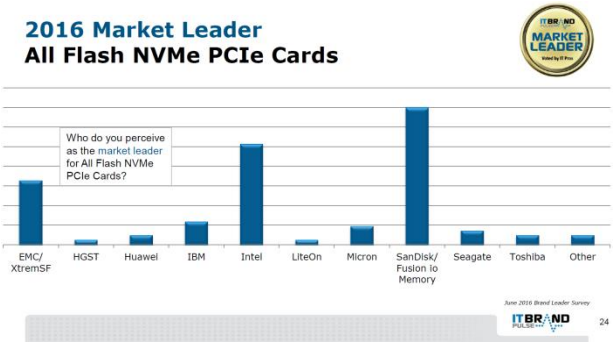
Press Release
Mini Report

	BRAND LEADER	MARKET LEADER	PRICE LEADER	PERFORMANCE LEADER	RELIABILITY LEADER	SERVICE AND SUPPORT LEADER	INNOVATION LEADER
All Flash Converged System	HPE	HPE	HPE	HPE	HPE	HPE	HPE
All Flash FC Array	EMC	Dell Storage	EMC	EMC	EMC	EMC	EMC
All Flash Hyperconverged System	Nutanix	Nutanix	Cisco	VCE	Cisco	Nutanix	
All Flash iSCSI Array	Dell Storage	Dell Storage	Dell Storage	Dell Storage	Dell Storage	Dell Storage	
All Flash NAS Array	NetApp	NetApp	NetApp	NetApp	NetApp & EMC (tie)	NetApp	
All Flash Unified SAN/NAS Array	NetApp	NetApp	NetApp	NetApp	NetApp	NetApp	
Flash Cache Software	SanDisk	SanDisk	SanDisk	SanDisk	SanDisk	SanDisk	
Hybrid HDD/SSD Array	EMC	Dell Storage	EMC	EMC	EMC	EMC	
NVDIMM	Micron	Micron	Micron	Micron	Micron	Micron	
NVMe SSD Modules	Samsung	SanDisk	Samsung	Samsung	SanDisk	Samsung	
All Flash NVMe Fabric Attached Array	EMC/DSSD	EMC/DSSD	EMC/DSSD	EMC/DSSD	EMC/DSSD	EMC/DSSD	
NVMe Network Adapters	Mellanox	Mellanox	Mellanox	Mellanox	Mellanox	Mellanox	
All Flash NVMe PCIe Cards	SanDisk	SanDisk	SanDisk	Intel	SanDisk	SanDisk	
Contributions to NVMe Technology (2012-2016)							Samsung
New Enterprise-Class NVM Product							Intel 3D XPoint NVDIMM

robert.kowalke@vita.virginia.gov

References

- http://nvmexpress.org/wp-content/uploads/2013/04/General_NVMe_FAQ.pdf
- https://storageswiss.com/2015/10/26/what-is-nvme/
- https://itpeernetwork.intel.com/why-you-should-care-about-nvm-express/
- http://info.grcooling.com/nvme-vs-sata-ssd
- http://storageioblog.com/nvme-place-volatile-memory-express/
- http://storageioblog.com/answer-nvme-questions/
- http://www.networkcomputing.com/storage/nvme-and-nvms-what-expect/345852419
- https://www.brighttalk.com/webcast/663/132761
- https://en.wikipedia.org/wiki/NVM_Express
- http://www.userbenchmark.com/Faq/What-s-the-difference-between-SATA-PCIe-and-NVMe/105
- http://blog.lenovo.com/en/blog/are-nvme-ssds-right-for-you/
- https://www.dell.com/support/article/us/en/4/SLN301149/frequently-asked-questions-regarding-nvme-solid-state-drives--ssds--and-their-interaction-during-the-installation-of-windows-on-systems-that-have-the-intel-skylake-chipset?lang=EN
- https://hardware.slashdot.org/story/16/10/01/2127209/with-hdds-on-the-ropes-samsung-predicts-ssd-price-collisions-as-nvme-takes-over
- http://www.nvmexpress.org/wp-content/uploads/NVMe_Overview.pdf
- http://www.thinkmate.com/inside/articles/nvme
- http://www.datacenterjournal.com/answer-nvme-questions/
- https://www.micron.com/solutions/why-nvme-should-be-in-your-data-center
- https://itbrandpulse.com/it-brand-pulse-presenting-2016-innovation-leader-awards-at-flash-memory-summit/
- https://www.slideshare.net/ITBrandPulse/flash-storagenvme-brand-leader-minireport
- https://www.flashmemorysummit.com/English/Conference/Proceedings_Chrono.html
- http://www.thessdreview.com/daily-news/latest-buzz/intel-ssd-dc-p3700-review-800gb-nvme-enterprise/



robert.kowalke@vita.virginia.gov