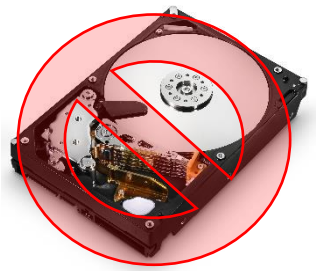
TRENTON Systems, Inc.
1725 MacLeod Drive • Lawrenceville, Georgia 30043
Sales (800) 875-6031 • Phone (770) 287-3100 • Fax (770) 287-3150

Internal Note: **The Future of Non-Volatile Storage**

Date: **September 12, 2016**

Non-volatile storage has long been one of the most restrictive bottlenecks in computing. Humanity generates 2.5 quintillion bytes of data daily. That data has to be stored somewhere, and in the modern computing ecosystem, consumers expect to be able to access it instantaneously. The speed increases provided by moving to flash storage from traditional HDDs in the last decade quickly saturated contemporary storage interfaces such as SATA and SAS. Thus, the industry has defined new standards such as Non Volatile Memory Express (NVMe) to outline how new storage devices should communicate over the PCIe bus. The PCI Express Bus' serial nature, high speed and low latency made it an ideal candidate to succeed SATA and SAS for storage interconnectivity. Standard PCIe connectors proved too large to reasonably leverage as storage interconnects, so a smaller or more traditional interface was required. Two emerging standards are competing to become the de facto physical standard in high-speed nonvolatile storage, M.2 and U.2. The Trenton TKL8255 is launching with provisions for M.2 drives, pictured right.



The enterprise storage market is moving away from not only rotational, magnetic media in large and mission critical storage deployments but also traditional SSDs. Available native PCIe lanes per processor have increased dramatically in recent years, resulting in the adoption of high-bandwidth, low-latency interfaces for storage devices becoming more palatable from a value perspective. More importantly, support for ever-larger Input/Output operations per Second (IOPS) at real-time latencies has become a major decision point as well, with the dawn of the plug and play consumer IoT and mobile-first thinking.

In some of the most IOPS-intensive enterprise applications, even these new standards are proving inadequate to meet performance goals of system designers. Development of Memory Channel Storage (MCS) to bring nonvolatile storage onto the system memory bus, on standard memory DIMMs, is underway in order to provide unparalleled latency, bandwidth and performance scaling. The ultimate goal being to make latency differences between system memory and persistent storage nominal.

Moving forward, Bob Brennan, SVP, Memory Solutions, Samsung Semiconductor Inc. believes that Flash will become tiered, as speed tradeoffs are realized and the need for additional long-term storage capacity at the cold or warm levels are increased. He believes the hierarchy will be NVMe -> Bulk "WORM" Write Once/Read Many Flash (SATA/SAS SSDs) -> Magnetic Disks or larger, slower flash SSDs as cold storage. Mark Peters of the Enterprise Strategy group says of Enterprise Flash, "There really is no longer much, if any, debate about the inherent value and desirability or even relative reliability and longevity of flash. Instead the questions are now about the type, speed and extent of adoption. Simply put, it is not *if* flash will be adopted, but where and how soon."

About NVMe

NVMe (Non-Volatile Memory Express) is a technology that allows solid-state flash storage to be addressed directly via the PCIe bus. This allows for much greater transfer speeds and lower latencies than legacy storage interfaces while reducing system size and power requirements.

The NVMe standard allows system designers to reduce system complexity and reliably increase performance in the area that has traditionally been the most restrictive bottleneck on computing systems, the non-volatile storage.

TRENTON Systems, Inc.
1725 MacLeod Drive • Lawrenceville, Georgia 30043
Sales (800) 875-6031 • Phone (770) 287-3100 • Fax (770) 287-3150

Transfer speeds are not the only improvement, NVMe supports 64k commands/64k queues whereas SATA supports 1 command/32 queues, further, the controller is cognizant of multiple processor cores and can prioritize requests. These advances should greatly improve multitasking performance.

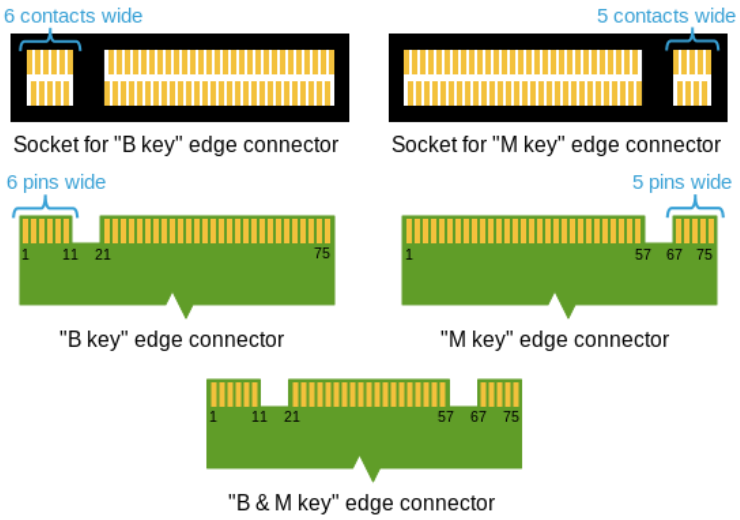
About M.2

NVMe on M.2 was initially popularized in the evolutionary 2013 MacBook Air and the Sony Vaio Pro. Initial drives were limited to PCIe 2.0, as the necessary PCIe lanes were pulled from the PCH, resulting in real world performance around 1GB/s read. Consumer adoption of Solid State Storage has been expanding since The Fourth Generation Intel Core "*Haswell"* introduced native support for booting from NVMe on the Z97 and X99 chipsets. The drive for lower power/lower weight solutions in mobile products has fueled expansion of NVMe storage products--the upper bound for TDP of M.2 SSDs is 10w. For our SWaP-conscious customers, this could be a very important purchase consideration. Having performance that exceeds that which previously took a whole RAID of SSDs onboard a SHB, no cabling required, under 10w, could provide an excellent platform for a customer to utilize our SHB in a creative or custom design.

M.2 Interface cards have several different physical form factors. The most important, or, most limiting, of these form factors are the Keys: A, B, E and M. These keys refer to the notches in the pins of the cards. **For our purposes, it is important to remember that the key on the TKL8255 is a "M" key which supports both M and B implementations.**

Most of our customers will be wanting to utilize this interface for storage as the SSDs utilize a technology known as NVMe, outlined below, that provides a substantial speed increase over the most common storage interconnects in use today, (SATA, SAS).

Keys A and E provide PCIe 3.0 x2, USB 2.0 and I^2C support and are generally used for WiFi/Bluetooth/LTE/3G/GPS transceivers, as these devices do not require much bandwidth. A and E-keyed devices will **NOT** work with the TKL8255.



Key B provides PCIe 3.0 x2, SATA, USB 2 & 3, analog and PCM audio, and I^2C. These devices will work in the M-keyed TKL8255. Performance of storage devices which utilize this key can be expected to be on par with current SATA drives.

Key M provides PCIe 3.0 x4 and SATA. These devices will provide the best storage performance, as it can fully saturate the provided links of the interface.

M.2 devices, no matter their key, come in a variety of card measurements, typically given in "XXYYY" where the two X digits indicate width in millimeters, the (up to) three Y digits indicate length. Assuming that a given device is keyed "B" or "M," the width measurement is not a concern, assuming the card conforms to that key's physical constraints. The TKL8255 accepts 42, 60 and 80mm long M.2 devices and has appropriate mounting locations for each. **The most popular form factor for storage-oriented M-keyed NVMe drives is 2280.**

TRENTON Systems, Inc.
1725 MacLeod Drive • Lawrenceville, Georgia 30043
Sales (800) 875-6031 • Phone (770) 287-3100 • Fax (770) 287-3150

The M.2 interface has largely taken over from the previous mSATA standard for small-footprint solid-state devices as it is generally more versatile and space-effective than mSATA. M.2 was explicitly designed to maximize utilization of card space while minimizing the effective footprint.
Samsung has been mass-producing x4 M.2 SSDs since Q2 2015 and most new designs of portable computers—laptops, tablets, convertibles, etc., utilize M.2 storage solutions as of Q3 2016, whether x2 or x4. Even the desktop/enthusiast/workstation market cannot ignore the performance benefits of having storage directly on the PCIe bus, and most have begun spinning motherboards that take advantage of the technology. The gaming workstation oriented Z170 chipset supports up to three x4 PCIe NVMe devices simultaneously, allowing for RAIDs of NVMe devices.

SNIA, a storage industry think-tank and standards promoter, predicts M.2 unit shipments of 175 million by 2018 (pp.8).

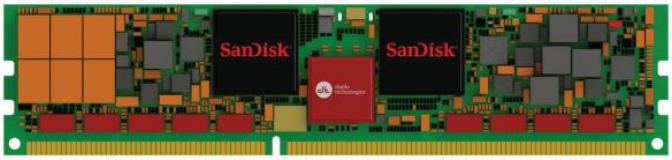
About U.2

If M.2 is a radical departure from storage device interconnect orthodoxy, U.2 is the reformation movement. Originally known as SFF-8639, U.2 is, essentially, "SATA Express on steroids with support for four PCIe lanes." SATA Express was an adaptation of the SATA interface, analogous to the early 2-lane M.2 slots. It was an improvement over the existing interface, providing more speed, however, the cabling scheme was complicated and it was never adopted in as wide a manner as early M.2 was by the industry or the public. The operative difference between U.2 and M.2 is that while M.2 devices are generally limited to board-mount solutions, which can be prohibitive for surface-mount components on a board, U.2 is a plug-and-cable type solution. This means that more traditional form factors, i.e., 2.5" drives and, thus, existing chassis, will continue to be useful in deploying U.2-based NVMe. Performance differences between the two interfaces can be expected to be nominal, assuming non-extended cabling lengths. Cabling can be expected to be more expensive and more delicate than legacy storage interconnects, so care must be taken in engineering U.2 solutions as to not have small-radius bends required of the cabling.

U.2 Host SFF8463 (left) and SFF8639 Target (Right)

About Memory Channel Storage

Memory Channel Storage is an architecture that enables flash memory to directly interface to a memory channel from the processor and be presented to the operating system as a traditional storage device in the traditional DIMM form factor. Other than speed and latency benefits (write latency as low as 3.3 µs) this paradigm offers packaging benefits similar to M.2, as there are no external connections required, power and data are supplied by the host board. Currently, most customers interested in this technology are high-end datacenters, however, economies of scale dictate that the benefits of MCS will begin to ripple out into the market in the coming years. Pricing is not currently readily available but can be expected to be prohibitive for all but the most cost-agnostic customers.



Diablo Technologies' MCS DIMM

MCS does have the benefit of being fully supported on existing SHB/Motherboard designs out of the box, with no hardware changes required, other than the caveat that for every MCS DIMM deployed, the system loses the ability to use that DIMM for system memory. Application software optimizations will be required to obtain peak performance out of MCS devices.

TRENTON Systems, Inc.
1725 MacLeod Drive • Lawrenceville, Georgia 30043
Sales (800) 875-6031 • Phone (770) 287-3100 • Fax (770) 287-3150

Comparisons

Assuming 4 lanes of PCIe 3.0, a theoretical maximum bi-directional data throughput rate of 8GB/s is possible allowing NVMe manufacturers much headroom in increasing the performance of their drives before an interface change or upgrade is required, current maximum transfer speeds are about ~2400MB/s. (See attached benchmarks, below.)

Market forces are moving towards the M.2 standard. Samsung, one of the world's foremost manufacturers of NAND, has not released "950" versions of their venerable 850 line of SATA SSDs, presumably because of the large bandwidth bottlenecks presented by the older interface. A retail, brick-and-mortar comparison of the two generations of drives shows a cost delta of ~$60, or about 43% for the same capacity. **The 950 Pro NVMe M.2 drive is about 240% faster than the 850 EVO.**

At press time, only Super Talent and Intel have released U.2-compatible drives. Only about 12 consumer motherboards currently support the U.2 spec, with many more M.2 to U.2 converter cards available. U.2 has more traction in the server market, where Supermicro utilizes it in their SuperStorage appliances. It is worth noting that in the Supermicro implementation, NVMe 1.2 support is provided in order to fully support hot-swapping.

M.2 has a price advantage over U.2 at the moment, likely due to M.2's more mature commercial availability window. The best price comparison at the moment is the U.2 400GB Intel SSDPE2MW400G4R5 at $449.99 (449.99/400GB=1.125 $/GB) versus the M.2 512GB Samsung MZ-V5P512BW at $315.62 (315.62/512GB=0.616 $/GB). This cost delta could either shrink or enlarge, depending on market receptivity to the newer form factor. It is likely that the enterprise will be more likely to adopt U.2 due to the aforementioned compatibilities with existing infrastructure.

MCS' major player with product in the field right now is Diablo Technologies. Intel is currently developing their own solution for MCS, which claims to be able to place 1TB of storage onto a standard DIMM sized "drive." It is also worth remembering that with multiple memory controllers onboard a processor, MCS technology scales exceptionally well.

Several industry players have created products from value-conscious entry level to cost-no-object best of the best appliances. Dell offers an all-flash array systems starting at $25,000 with little data provided on capacities or latencies at that price point. Violin Memory offers little in the way of pricing information on their systems which combine both hardware and data as a service components, but do tout 1 million IOPS at 1ms latency and 1.4PB effective capacity from their Flash Storage Platform 7700. Bitmicro claims 560K IOPS at $0.11 per I/O. Smart Storage Systems has implemented a fully MCS-enabled system in a IBM chassis with ½ a Petabyte as of 2013.

TRENTON Systems, Inc.
1725 MacLeod Drive • Lawrenceville, Georgia 30043
Sales (800) 875-6031 • Phone (770) 287-3100 • Fax (770) 287-3150

As a Trenton Product

An example x4 M.2 NVMe SSD, the Samsung 950 Pro, is currently $187.75 on Amazon for a 256GB drive at 2,200MB/s read, 900MB/s write, 270k read IOPS, 85k write IOPS. Using Sales Engineering's preferred PCIe to M.2 adapter at a cost of $20 means that the retail cost to deploy 256GB of x4 NVMe into one of our systems is ~$210, less labor. This means that for $3,000 in option cards and M.2 devices, ($2730) a Trenton Systems chassis with a BPG8194 (assuming 13 PCIe) could deploy 3.3 raw, unformatted TB of storage. Similarly, the 512GB version of the 950 Pro at $316 would offer 6.6 unformatted TB of storage for ~$4500. As the NVMe technology improves, speed and capacity are sure to go up.

Approximations of NVMe Systems (Typical Values for Single Systems)			
Technology	*HDEC*	*PICMG*	*Motherboard*
Max. # Slots	18	13	7
Max. Storage Capacity	~9TB @ 512GB/ea	~6.5TB @ 512GB/ea	~3.5TB @ 512GB/ea
Approximate Cost	~$15,000	~$9,500	~$4,500

A custom Trenton solution for NVMe could theoretically leverage up to 20 NVMe devices assuming Haswell-EP HDEC numbers of PCIe lanes, without the need for switches. (80 lanes/4 lanes per NVMe=20 devices) With switches, many more could be implemented, but latency will increase. Such a project would require at least custom board work and possibly custom metal work.

MCS is prohibitive due to available board space/DIMM capacity on our traditional SBC/SHBs. Future Trenton products which are able to fully utilize all the memory channels provided by high-end dual-socket processor designs will afford system designers far more leeway when deciding to implement MCS-enabled compute solutions.

Final Thoughts

The M.2 interface has industry backing, a consumer mandate and represents a legitimate paradigm shift in storage technology. Prices for these drives will surely fall as they become ubiquitous in all but the most value-oriented consumer devices. Commercially, U.2 has a handhold. Something "has to give" as SATA and SAS simply cannot compete with newer interconnect technologies. U.2 seems best positioned to capitalize on this untapped market. In the enterprise arena, Memory Channel Storage looks poised to become a major market force. The initial cost barrier is sure to be high, but this market generally shows willingness to endure longer time to values. Regardless of which form factor wins out in the high performance/server world, how well the benefits translate, sales-wise into the embedded, industrial and military computing markets are less clear, but the performance, power consumption, form factor and vibration-resistance benefits speak for themselves: these new storage technologies will be making their way not only into the datacenter but also into the embedded and rugged computing market.

Additional Reading

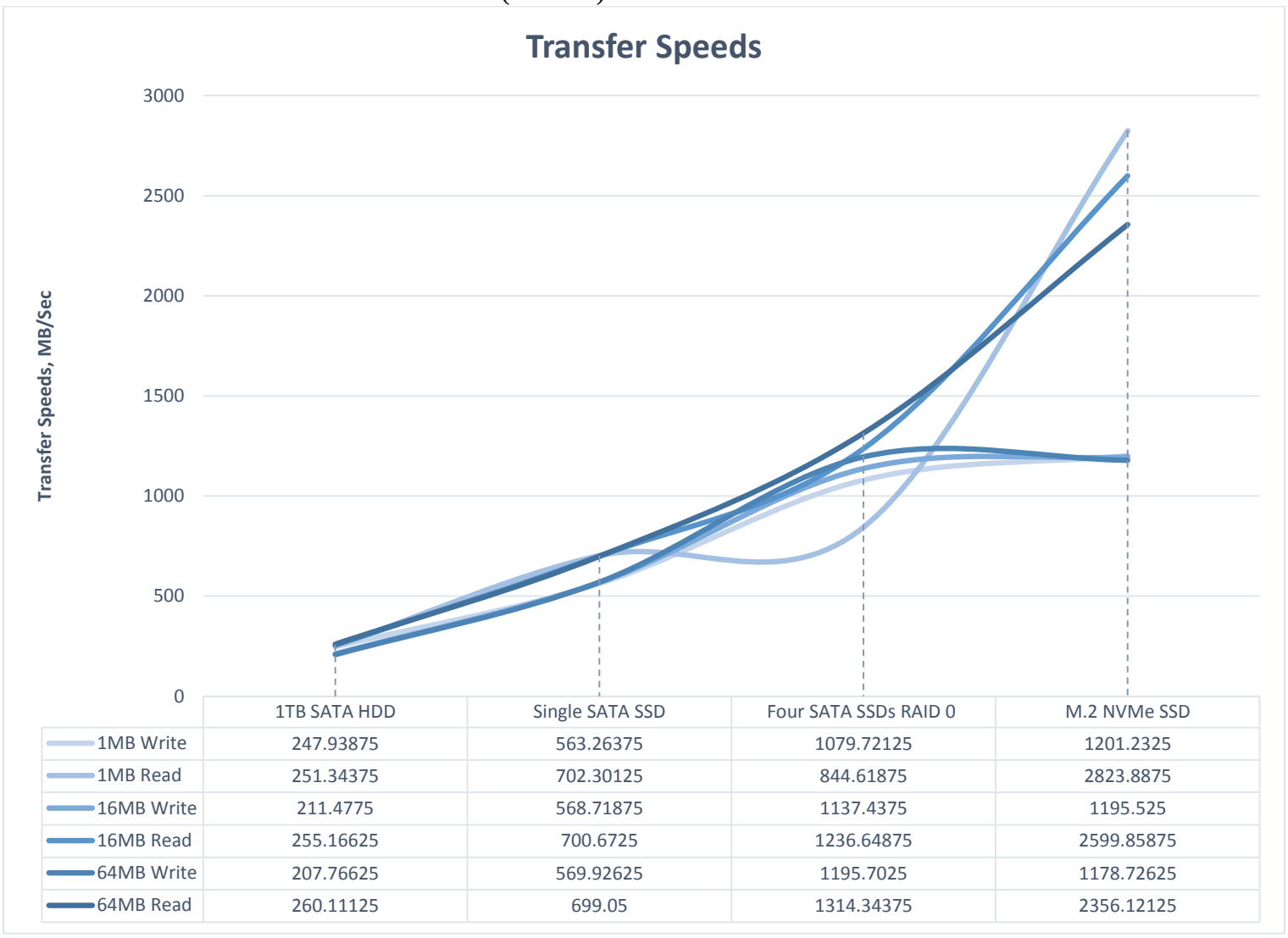
http://www.tomshardware.com/reviews/intel-750-series-ssd,4096-2.html
http://arstechnica.com/gadgets/2015/02/understanding-m-2-the-interface-that-will-speed-up-your-next-ssd/
https://www.sata-io.org/sata-m2-card
http://www.extremetech.com/computing/162944-diablos-memory-channel-storage-tech-will-deliver-terabytes-of-ram-using-nand-flash
http://www.diablo-technologies.com/benchmarking-ibm-exflash-dimm-sysbench-fileio/

Synthetic Performance Tests of NVMe x4 PCIe 3.0 on TKL8255

The following is a table and graph of four different standard system storage paradigms, a traditional, 1TB magnetic rotating 7200RPM SATA600 drive, a single SATA Solid State Disk (SSD), four SATA SSDs in a RAID0 and a M.2 x4 PCIe NVMe SSD. As you can see, the performance increase the NVMe M.2 drive provides is dramatic, even over the RAID0 of SATA SSDs, on all data block sizes.

Percent read increase of M.2 over other drives (64MB block):

- 89% faster than a 1TB SATA HDD
- 70% faster than a SATA SSD
- 44% faster than a SATA SSD RAID0 (4 drives)



	1TB SATA HDD	Single SATA SSD	Four SATA SSDs RAID 0	M.2 NVMe SSD
1MB Write	247.93875	563.26375	1079.72125	1201.2325
1MB Read	251.34375	702.30125	844.61875	2823.8875
16MB Write	211.4775	568.71875	1137.4375	1195.525
16MB Read	255.16625	700.6725	1236.64875	2599.85875
64MB Write	207.76625	569.92625	1195.7025	1178.72625
64MB Read	260.11125	699.05	1314.34375	2356.12125