PCIe NVMe* SSD in Smaller Form Factors

Becky Loop, Sr. Principal Engineer, Intel

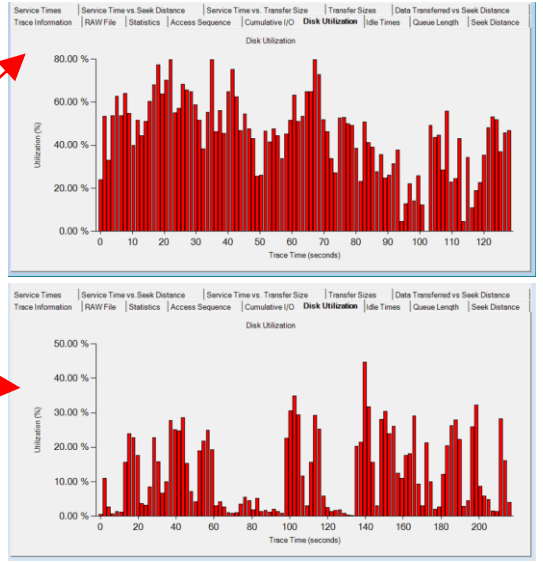
Zhiping Yang, Hardware Manager, Google

What is Small Form Factor Computing?

Problem with small form factor devices

Running Google Play Apps in Chrome with minimal DRAM, eMMC shows user experience issues:

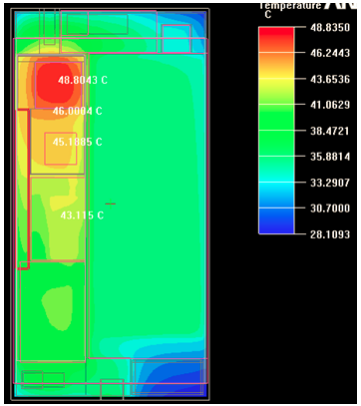
- Gameplay (Badland)
  - Install time measured for eMMC is 10s vs. 6s for NVMe
  - Outlier latencies of 0.15s for eMMC vs. 0.0015s for NVMe
  - Frame drops seen with eMMC. None seen with NVMe.
- Multitasking:
  - NVMe shows smooth transitions. eMMC transitions glitch
  - NVMe shows 50% less disk utilization than eMMC



Source: Storage Technologies Group, Intel. Setup: SKL-U reference platform,1GB DDR3, eMMC=HS400, NVMe=PCIe gen 3x4

eMMC with small capacity DRAM sees user experience issues that are fixed by adding DRAM (cost adder). NVMe does not see these issues.

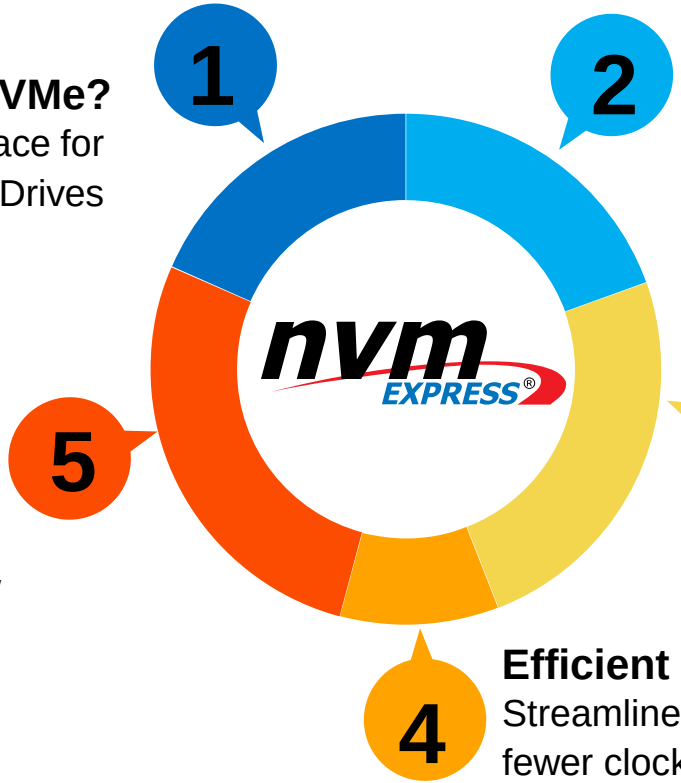
Small Form Factor Storage Needs



NVM Express Explained

What is NVMe?

THE industry standard interface for PCI Express® Solid-State Drives



2 Architected for SSDs

Eliminates hard drive legacy that SATA and SAS are based upon.

3 Works Out of The Box

Drivers for all major OSes

4 Efficient & Scalable

Streamlined protocol & command set, fewer clock cycles per IO

5 Optimized for Next Gen NVM

New storage stack with low latency to take full advantage

*Other names and brands may be claimed as the property of others.

Why NVMe in Small Form Factor?

- Scalable software stack across all Intel® Core™ client platforms
- Low latency interface
- Drivers for all major OS's
- Host Memory Buffer provides DRAM-less solutions
- Scalability to next generation NVM
- One less platform interface to develop and validate

NVMe has clear benefits that are used today in the client and can also be realized in small form factor.

Where 16x20mm exists

Google pcie bga

All | Shopping | Images | Videos | News | More | Search tools

About 355,000 results (0.70 seconds)

[PDF] PCIe* BGA SSD - Flash Memory Summit
www.flashmemorysummit.com/English/Collaterals/.../20140805_101B_Loop.pdf
Aug 4, 2014 - PCIe* BGA SSD. Becky Loop. Principal Engineer. Intel PC Client Group. *Other names and brands may be claimed as the property of others.

Samsung Demos Its First BGA SSD: 1500 MB/s Read Speed and Tiny ...
www.anandtech.com › SSDs AnandTech
Mar 22, 2016 - Early last year Toshiba rolled-out its first BGA SSDs with PCIe interface, and in September several members of PCI SIG (the organization, which ...

Samsung Begins Mass Production of PM971: Tiny BGA SSDs with ...
www.anandtech.com › Computex_2016 AnandTech
Jun 1, 2016 - 512 GB. Form Factor, BGA 20 mm × 16 mm × 1.5 mm. Controller, Samsung Photon. Interface, PCIe 3.0 x2. Protocol, NVMe. DRAM, 512 MB.

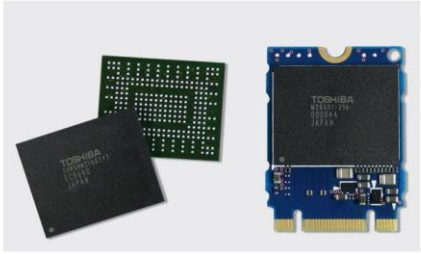
Toshiba debut's world's first single-package PCIe SSD | KitGuru
www.kitguru.net/components/.../toshiba-debuts-worlds-first-single-package-pcie-ssd/
Jan 8, 2015 - A good news is that Toshiba's BGA PCIe SSDs will be used inside tablets powered by the latest Intel Atom "Cherry Trail" or Core M "Broadwell" ...

Samsung Mass Producing Industry's First 512-Gigabyte NVMe SSD in ...
https://news.samsung.com/.../samsung-mass-producing-industrys-first-512-gigabyte-n...
May 31, 2016 - ... Producing Industry's First 512-Gigabyte NVMe SSD in a Single BGA ... NVMe* PCIe solid state drive (SSD) in a single ball grid array (BGA) ...

[PDF] Supporting PCIe and SATA BGA form factor for SSDs ECN - PCI-SIG
https://pcisig.com/sites/default/files/specification.../BGA_SSD_ECN.pdf PCI-SIG
Sep 25, 2015 - will allow PCIe and SATA to be delivered using a BGA package, ... BGA pinout supports additional pins than defined for Socket-3, for ...

Toshiba to Showcase World's First PCI Express Single Package SSD ...
www.businesswire.com/.../Toshiba-Showcase-Worlds-PCI-Express-Single-Package
Jan 7, 2015 - Toshiba will showcase a reference display of the world's first PCI Express (PCIe) single package SSD, integrating up to 256 GB in a single BGA ...

Toshiba Corporation's Semiconductor & Storage Products Company announced it will showcase a reference display of the world's first PCI Express (PCIe) single package SSD, integrating up to 256 GB in a single BGA package, at the 2015 International Consumer Electronics Show (CES), to be held from January 6 to January 9 in Las Vegas, U.S.A.



ANANDTECH

PC COMPONENTS | SMARTPHONES & TABLETS | SYSTEMS | ENTERPRIS

TRENDING TOPICS | GPUS | CPUS | INTEL | AMD | SSDS | SMARTPHONES | MOBILE | NVIDIA | STORAG

Home > Computex_2016

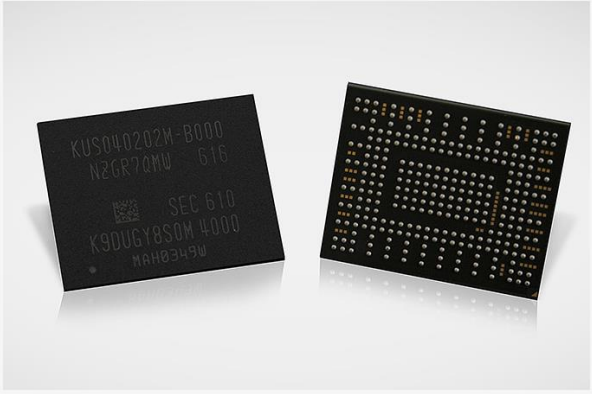
Samsung Begins Mass Production of PM971: Tiny BGA SSDs with 1500 MB/s Read Speed

39 Comments

+ Add A Comment

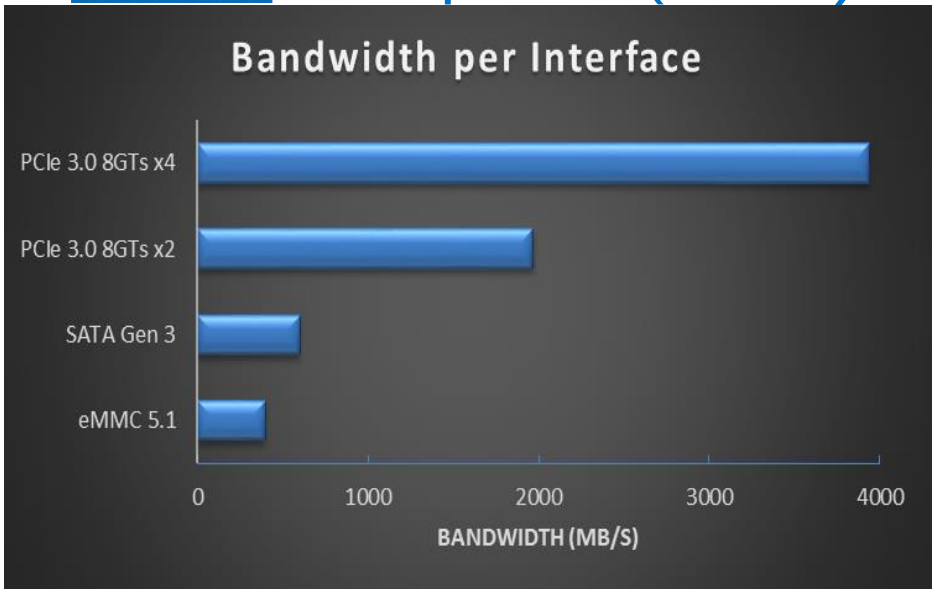
by Anton Shilov on June 1, 2016 10:00 AM EST

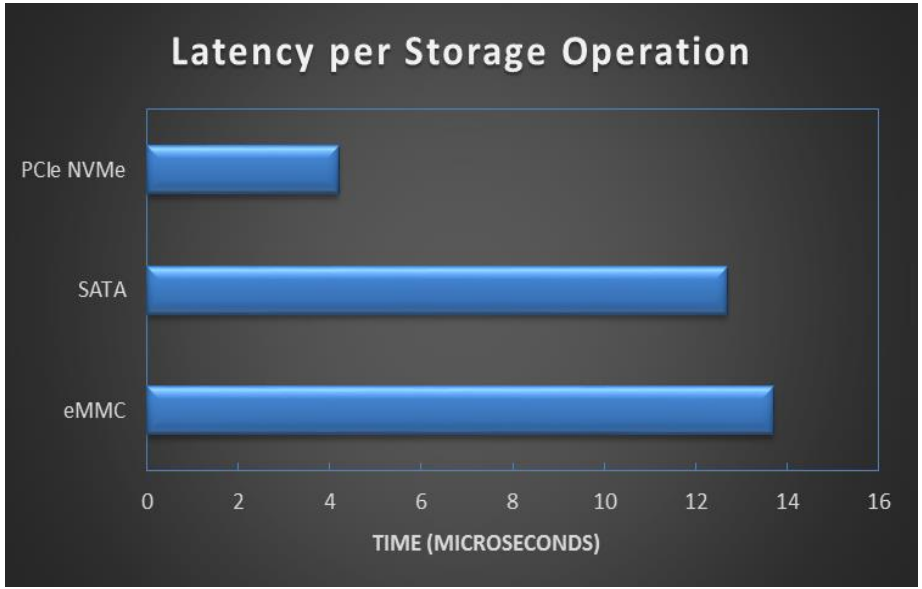
Posted in Computex_2016 | SSDs | Samsung | Trade Shows | V-NAND | BGA



Samsung this week said that it had begun mass production of its first SSDs in BGA form-factor. The PM971-NVMe tiny drive weighs about one gram, can store up to 512 GB of data and offers up to 1500 MB/s read speed. The SSDs will be used inside 2-in-1 hybrid computers, high-end tablets as well as ultra-thin notebooks later this year.

Controller + S/W Latency Improvement With PCI Express* (PCIe*) in Client

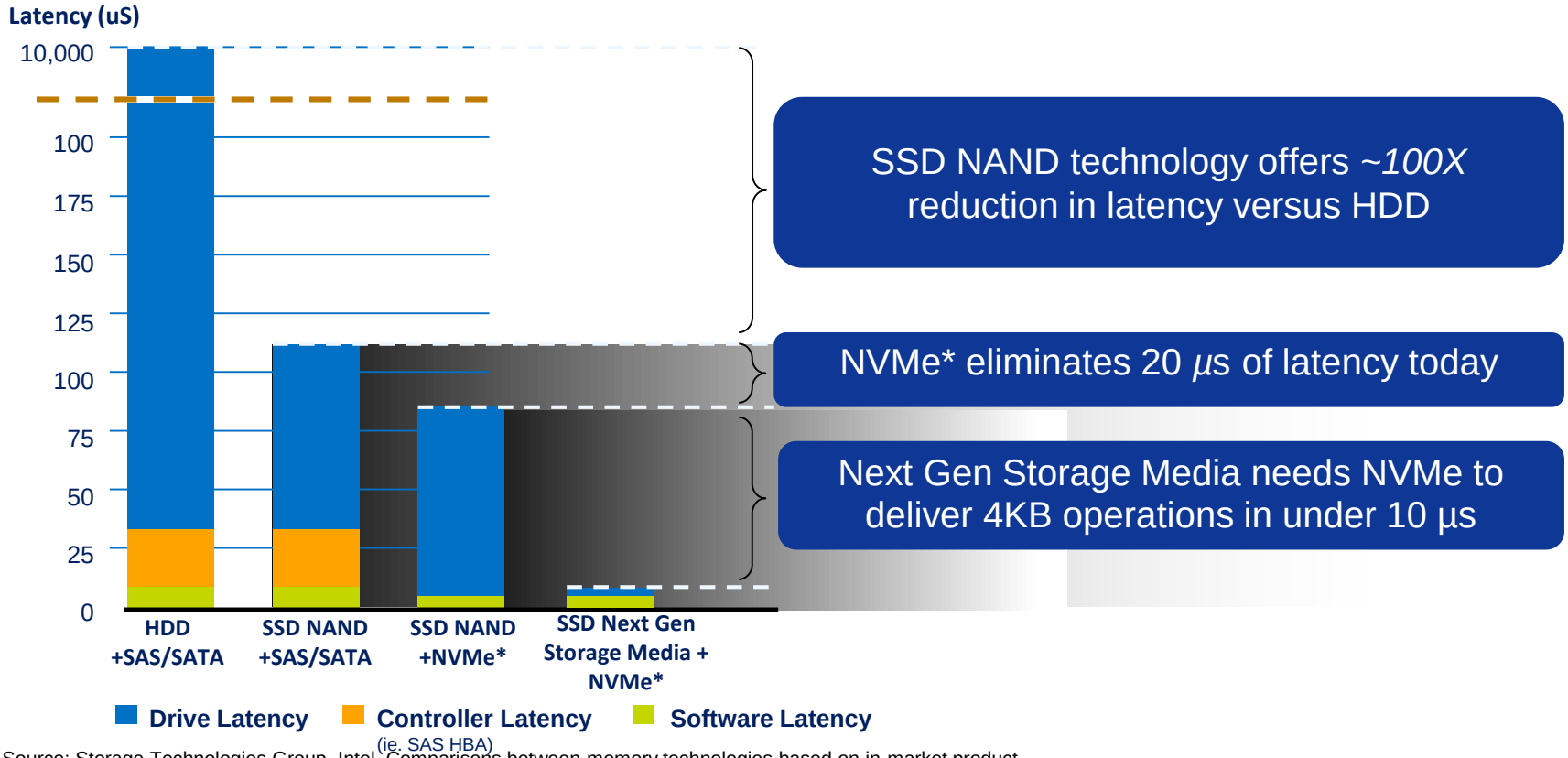




Source: Intel. Paper calculations of Bandwidth and latency. Bandwidth from each interface are max theoretical that account for encoding overhead. Latency values are estimated based on storage access to and from memory. eMMC and names associated with it are property of JEDEC.

PCIe* SSDs using NVMe* provides the higher throughput and lower latency

Delivering Benefit of Next Gen NVM



Source: Storage Technologies Group, Intel. Comparisons between memory technologies based on in-market product specifications and internal Intel specifications. *Other names and brands may be claimed as the property of others.

NVMe SSD Power

- Importance of active and standby power
  - Active Power: Power consumed when doing something
  - Standby Power: Power consumed when doing nothing
- Active Power consumption is about energy consumed. PCIe is competitive.
- L1 sub-states (L1.1 and L1.2) introduced in PCIe reduce standby power to mobile acceptable ranges.

Item	PCIe Gen3	PCIe Gen2	M-PHY Gear3
Line Speed [Gbps]	8	5	5.83
PHY Overhead	128/130,1GB/s	8/10, 500MB/s	8/10, 583MB/s
Active Power [mW]	60 (L0)	46 (L0)	58 (HS)
Standby Power [mW]	0.11 (L1.2)	0.11 (L1.2)	0.2 (Hibernate)
MB/mJ (higher better)	14-18	8-12	8-12

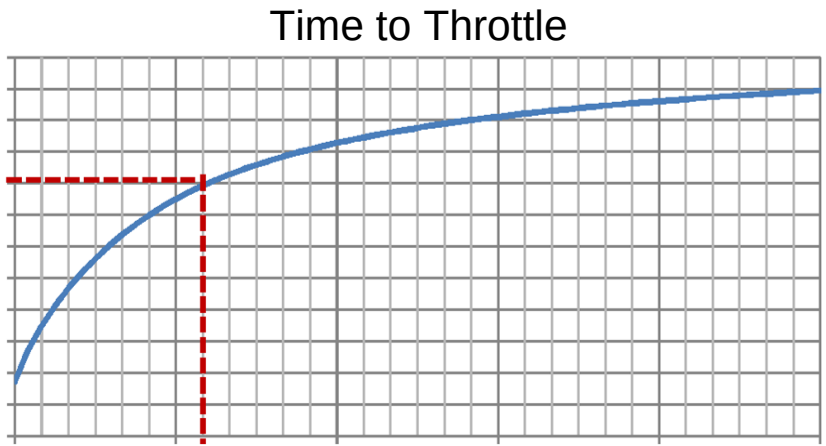
*UFS and names associated with it are property of JEDEC.

1. pci-sig.com: "L1 PM Substates with CLKREQ, Revision 1.0a"
2. Source: SanDisk. Data based on PHY power estimates of PCIe vs. MPHY.

NVMe SSDs can be power competitive in Small Form Factor

NVMe* SSD Thermals (16x20)

Workload	Perf / Power	Time to Throttle
Worst case	1600MB/s @ 4W	~60s to heavy throttle
Sustained	1000MB/s @ 2W	Light throttle only
TDP	500MB/s @ <1W	Never



Source: NVM Solutions Group, Intel. Data based on the simulated thermal capabilities of the 16x20mm PCIe* SSD package

Sustained power is 2x higher than published SSD TDP workload power. 30-60 second throttle means you can burst 10GBs+ before throttling

NVMe* SSDs provides burst performance while being able to throttle back if needed.

NVMe 1.2 Improvements

Host Memory Buffer

- Allows the host driver to allocate system memory for the SSD's exclusive use
- Platform Value
  - Enables DRAM savings & smaller BGA packages
  - E.g., Allocate translation tables in host DRAM
- Specification Details:
  - Device indicates preferred HMB size
  - Host enables/disables via Set Features

Composite Temperature

- Allows host to monitor temperature of the SSD
- Platform Value
  - Allows a temperature point that can be monitored.
  - If the host believes the temperature is out of its limits, it can set a lower power state on the NVMe device
- Specification Details
  - Device indicates temperature in SMART
  - Power State can be changed in power management

NVMe innovations enable scaling into smaller form factors delivering new differentiated platforms.

Key Focus Areas for 11.5x13

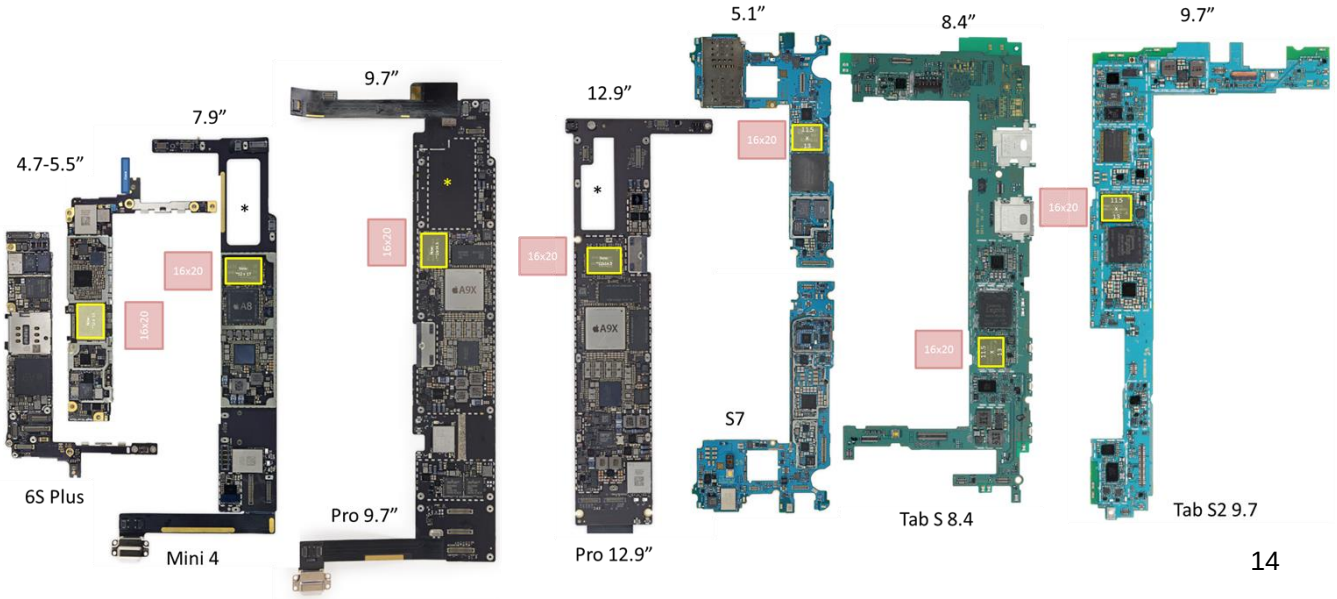
- Reduces the platform area in half compared to 16x20
  - 150mm2 vs. 320mm2
- 11.5x13 only supports x2 PCIe
  - no x4 or SATA support
- Dual footprint with other mobile storage devices
  - eMMC[1] and UFS[1]
- Optional SPI interface
- Multiple power rail ranges added
  - Platform still supports only three power rails (for power scalability)

1. eMMC and UFS specs are managed by JEDEC

NVMe innovations enable scaling into smaller form factors delivering new differentiated platforms.

Platform picture

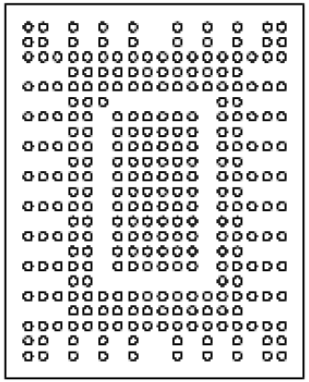
- 16x20mm does not fit in smaller form factor devices.
  - Need for a smaller performance SSD
- M.2 storage ~15% of platform area
- Allows for smaller motherboard.

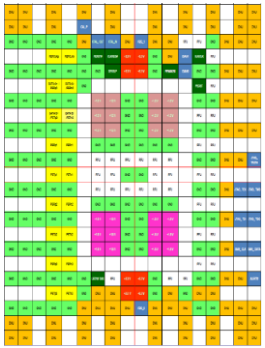


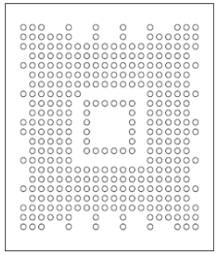
Credit: iFixit

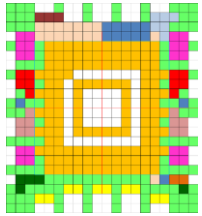
NVMe SSD's Size

- 16x20mm PCIe BGA device
  - Spec released in 2015[1]
- 11.5x13mm PCIe BGA device
  - Intel is working with PCI-SIG for spec under review





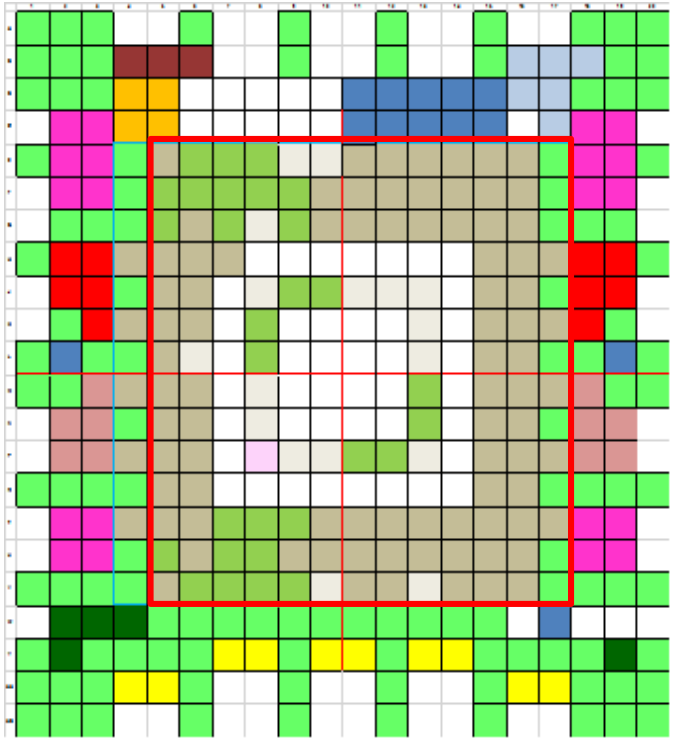




1. pci-sig.com: "Supporting PCIe and SATA BGA form factor for SSDs ECN"

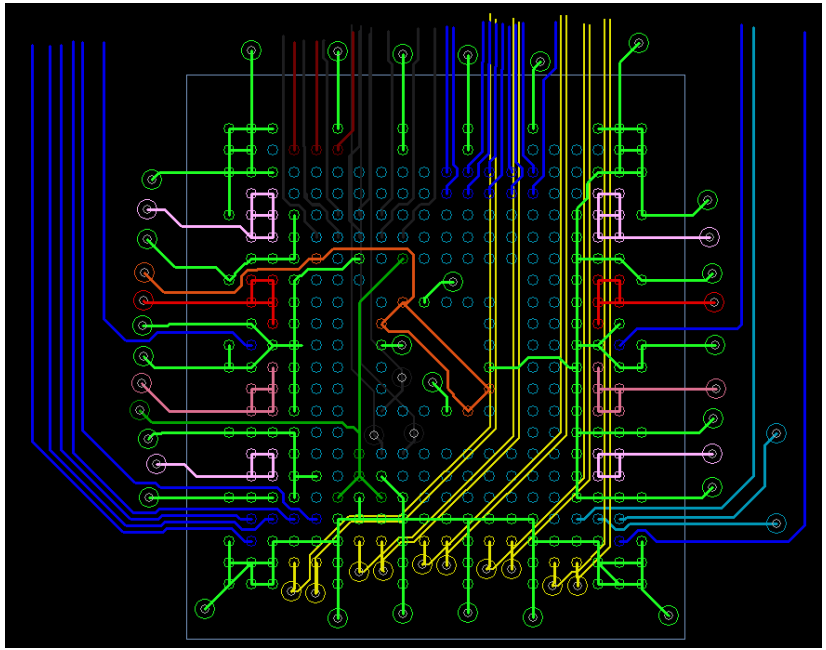
NVMe SSDs can scale to small form factors

MLB Dual Footprint for NVMe & eMMC



- The same MLB can be stuffed with either newly defined NVMe or existing eMMC or UFS.
  - The signals within the red square are for eMMC.
- It provides a good transition from eMMC based storage to NVMe storage.
- eMMC and NVMe together provide good **performance, price, feature, and capacity** coverage to meet different customers' needs.

Routing Studies for 11.5x13



The ballout is friendly to Type 3 or Type 4 motherboard designs enabling scalability across low cost and small form factor platform.

SPI ROM Interface for Secure Boot

Interface	Signal Name	I/O	Description	IO Voltage
SPI ROM[1]	WP_L	I	Write protect signal to prevent writes from occurring to SPI NOR. Active low.	1.8 V
	SPI_CLK	I	SPI clock. Max frequency is 50MHz.	1.8 V
	SPI_MOSI	I	Master Out Slave In signal for SPI NOR.	1.8 V
	SPI_MISO	O	Master In Slave Out signal for SPI NOR	1.8 V
	SPI_CS_L	I	Chip select for SPI NOR. Active low.	1.8 V
	SPI_18	I	+1.8 V supply. Optional voltage supply if SPI NOR included in package.	1.8 V

[1] **Optional.**

Cost

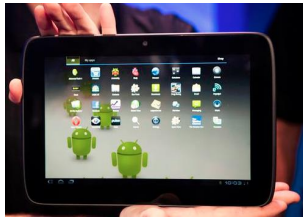
- Cost is a function of NAND, controller, and package/assembly/test.
  - The only major difference between a SiP PCIe SSD and eMMC is the controller difference (fixed cost)
    - •Higher capacity SSD means a lower percentage of cost due to controller
  - Controller cost has impact on performance.
    - •Lower performance means lower cost.

PCIe BGA SSD cost can be competitive to eMMC depending on capacity and performance targets

Building a Small Form Factor System

- Example: Detachable
  - Requirements: 256GB, High performance, in a thin and light system
  - Solution: 16x20mm or 11.5x13mm NVMe BGA, PCIe 8GTs x4 (x2 for 11.5x13mm)

- Example: Handheld Device
  - Requirements: 64GB, Small, secure.
  - Solution: 11.5x13mm NVMe BGA, PCIe 5GTs x2.



- Example: In-Dash Display
  - Requirements: 128GB, High performance but thermally limited
  - Solution: 11.5x13mm NVMe BGA, PCIe 8GTs x2, automotive grade

Call To Action

- Consider 11.5x13mm PCIe BGA devices for your small form factor and mobile usages.

Legal Disclaimers

▪Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com.

▪Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.

▪Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/performance.

▪Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

▪This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

▪No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

▪Statements in this document that refer to Intel's plans and expectations for the quarter, the year, and the future, are forward-looking statements that involve a number of risks and uncertainties. A detailed discussion of the factors that could affect Intel's results and plans is included in Intel's SEC filings, including the annual report on Form 10-K.

▪All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice. The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

▪Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

▪*Other names and brands may be claimed as the property of others.