



Flash Memory Summit

PCI Express[®]: What's Next for Flash Storage

Justin Wenck

Senior Technical Marketing Engineer



Flash Memory Summit

PCI-SIG® Snapshot

Organization that defines the PCI Express® (PCIe®) I/O bus specifications and related form factors.

- 750+ member companies located worldwide

PCI-SIG continues its solid reputation of delivering **low cost, high-performance, low-power specifications** to **support compliance** and **interoperability** across **multiple applications** and **markets**.



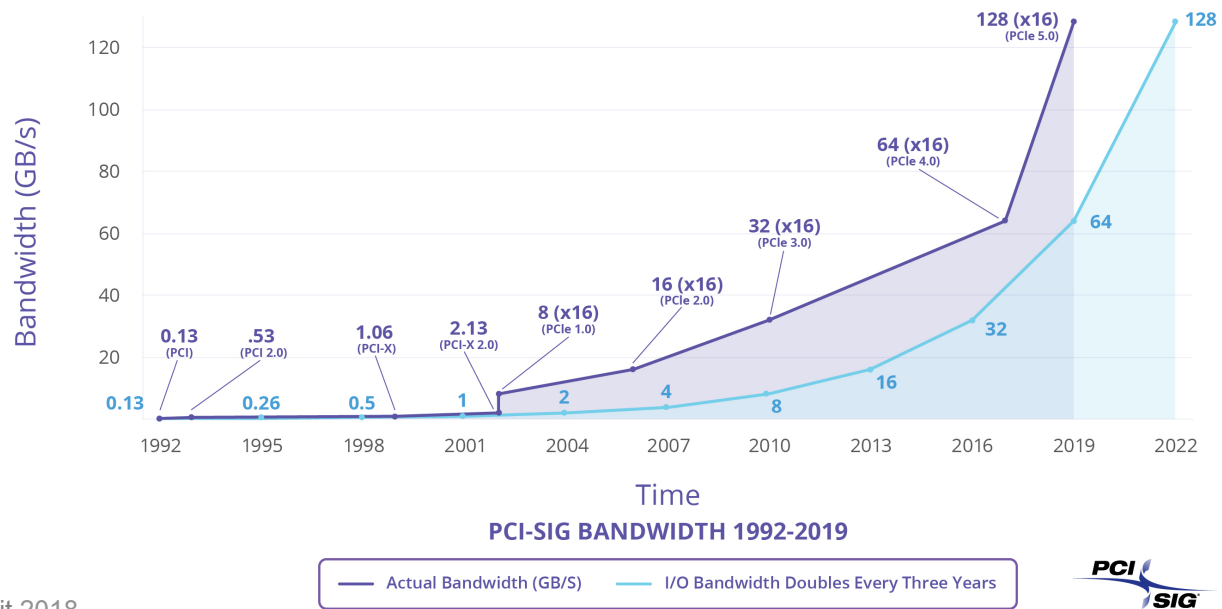
Flash Memory Summit 2018
Santa Clara, CA

BOARD OF DIRECTORS 2018-2019



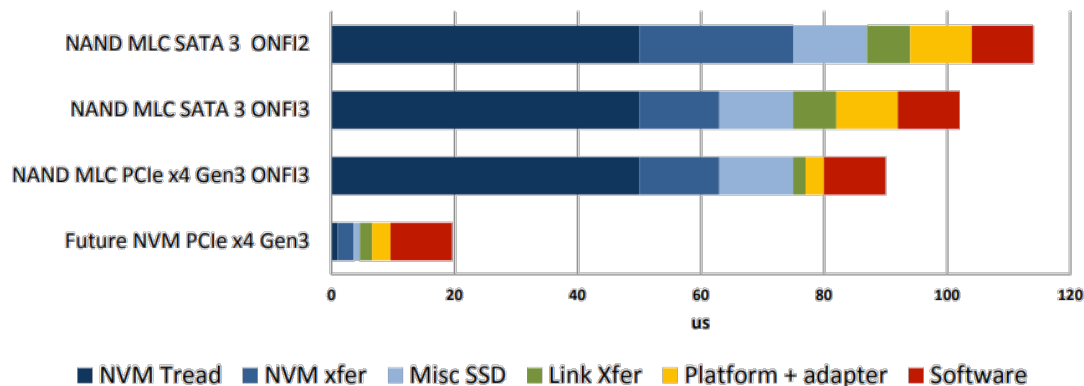
PCI-SIG Continues to Increase Bandwidth

 I/O BANDWIDTH DOUBLES
Every 3 Years



PCIe SSDs for Storage

App to SSD IO Read Latency (QD=1, 4KB)



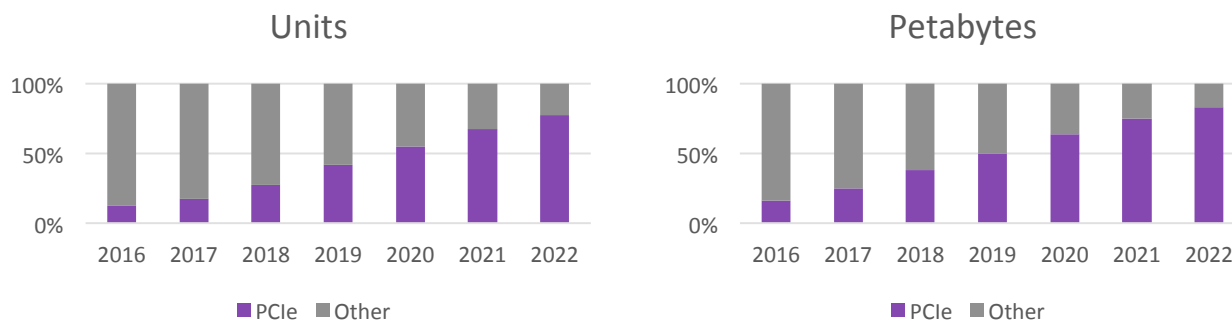
- PCI Express is a great interface for SSDs
 - Stunning performance
 - Lane scalability
 - Lower latency
 - Lower power
 - Lower cost
 - CPU-integrated PCIe lanes
- With Next Gen NVM, the NVM is no longer the bottleneck
 - 1GB/s per lane (PCIe 3.0 x1)
 - 4GB/s per device (PCIe 3.0 x4) or more
 - Platform + Adapter: 10 μ sec down to 3 μ sec
 - No external SAS IOC saves 7-10 W
 - No external SAS IOC saves \$
 - Up to 128 PCIe 3.0

Source: FMS
2013 “
[NVMe Express
Overview &
Ecosystem
Update](#)”



Growth of PCIe Technology in Storage

- Data explosion is driving SSD adoption
 - SSD market CAGR of 14.8% during 2016-2021 *source: IDC*
 - PCIe SSD market to surpass a CAGR of 33% during 2016-2020 *source: Technavio*
- PCIe technology is outpacing other interconnect technologies in both units and bandwidth/capacity



Source: SSD Insights Q1/18, Forward Insights



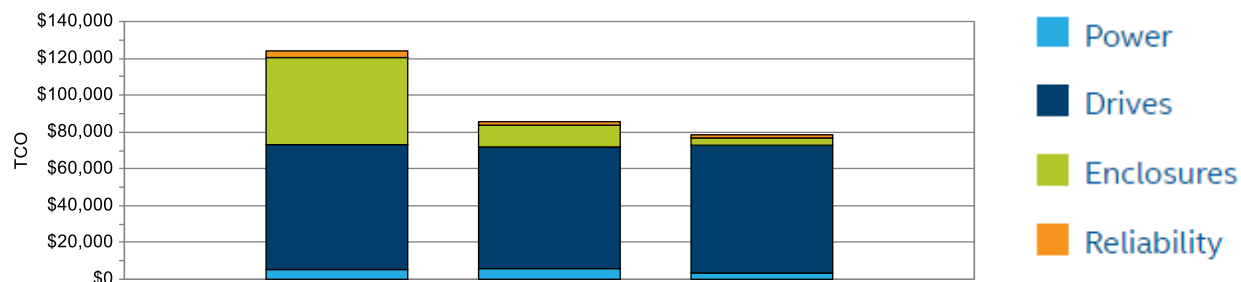
PCIe: Power Efficient Performance

- **Scalable Performance**
 - Width scaling: x1, x2, x4, x8, x12, x16,
 - Frequency scaling: Five generations
 - 2.5 and 5 GT/s with 8b/10b encoding
 - 8 and 32 GT/s with 128b/130b encoding
- **Low Power (Active/Idle)**
 - Rich set of Link and Device States
 - L0s, L1, L1-substates, L2/L3
 - D0, D1, D2, D3_hot/cold
 - Platform-level power optimization hooks: Dynamic Power Allocation, Optimized Buffer, Flush Fill, Latency Tolerance Reporting
 - Active power –5pJ/b, Standby power: 10uW/Lane
- **Vibrant ecosystem with IP Providers**



It's Not All About \$/GB – Total Cost of Ownership (TCO) Matters

TCO Savings Over 5 Years for Equal Write Performance Capture 60 Uncompressed UHD-1 4K (10-b) RGB



	SATA- Intel S3520	PCIe 3.0 – Intel P4510	PCIe 4.0 – FutureNVM*
TCO Savings	-	\$38,461	\$45,569
Total Drives	300	57	23
Sequential Read/Write (MB/s)	135000 / 114000	182400 / 114000 +35% / +0%	57500 / 115000 -57% / +1%
Power Reduction	-	~ 10%	~ 35%
Total Usable Capacity	144,000 GB	114,000 GB	8,625 GB

Source: <http://test-estimator.intel.com/ssddc/>

*Hypothetical PCIe 4.0 Future NVM SSD: 375GB, \$3017.00, 5000MB/s Seq Wr. 6Watts Idle, 25Watts Full Load, 0.44%AFR



Flash Memory Summit

Range of Form Factors

Current
PCIe
Form
Factors



Low Power NVMe
M.2 80mm and 110mm
U.2 2.5in x 7mm



Server Performance NVMe
Low profile HHHL x4 AIC
U.2 2.5in x 15mm



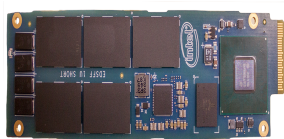
Server Performance NVMe
Low profile HHHL x8 AIC

Power (W)  Low

High

EDSFF
family

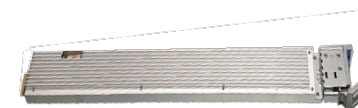
EDSFF 1U Short



EDSFF 1U Long x4, x8 (ruler)



EDSFF 1U Long 18mm heatsink





PCIe in Storage

Performance and user benefits for current and future storage applications

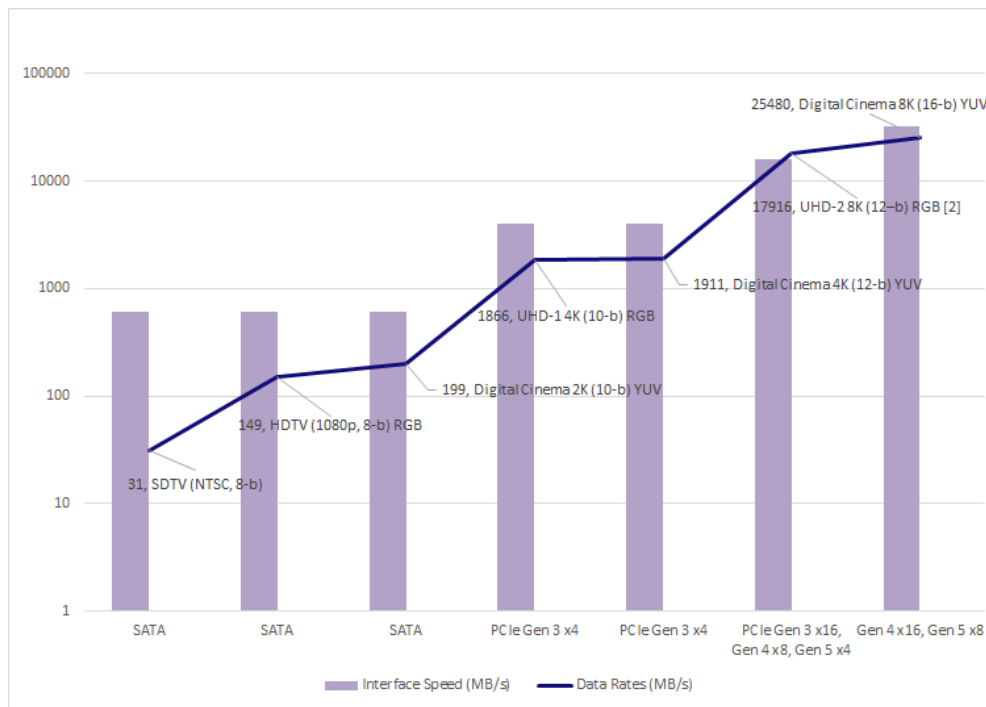
Faster data transfer:

- **PCI Express 3.0 Specification (8GT/s) published in 2010**
 - Low power with high performance
 - Wide breadth of solutions available from numerous vendors
 - Provides the cost effective performance required for storage today
- **PCI Express 4.0 Specification (16GT/s) finalized and published in October 2017**
 - Numerous vendors confirmed with 16GT/s PHYs in silicon
 - Major IP vendors offering 16GT/s controllers
- **PCI Express 5.0 (32GT/s) Specification targeted for release in Q1 2019**
 - Protocol already supports higher speed via extended tags and credits and additional changes target speed transition
 - Existing PHYs in the industry already run at 28GT/s / 56GT/s

Better user experience:

- Client and enterprise storage applications using PCIe technology helps keep data closer to CPU

Required BW for Uncompressed Video



Source: T. Coughlin, "How Big Are Your Dreams? Gauging the Size of Future Content [The Art of Storage]," in *IEEE Consumer Electronics Magazine*, vol. 6, no. 2, pp. 108-124, April 2017.



PCI-SIG and SDA Liaison

- PCI-SIG and the SD Association have formed a liaison to advance SD Express as a component in the PCIe ecosystem
 - Collaborate on a technical interchange related to SD Express and SD Express Test Guidelines, as well as information related to PCIe electrical certification of SD Express products
 - Form the PCI-SDA Advisory Team and the SD-PCIe Technical Group whose members are from companies that belong to both the SDA and PCI-SIG
- New SD Express Card leverages PCIe 3.0 interface to deliver up to 985 MB/s transfer rate
 - Maintains backward compatibility with existing SD hosts in the market
 - Meets changing performance levels of mobile and client computing, imaging, gaming, IoT and automotive applications



SD Association





Summary

- PCIe 3.0 provides the cost effective performance required for storage today
 - Will continue to outpace other SSD storage interconnect technologies in both units and bandwidth/capacity
- PCIe 4.0 (16GT/s) and 5.0 (32GT/s) will continue to deliver performance and key features
 - Enabling emerging NVM



Flash Memory Summit

Back-up Slides

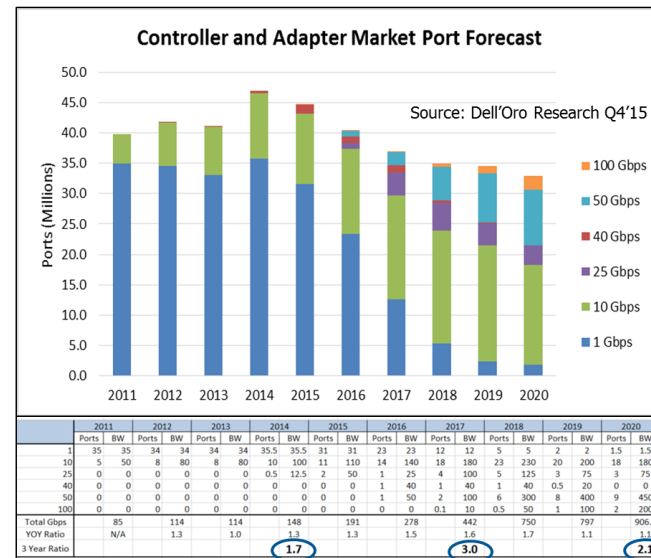
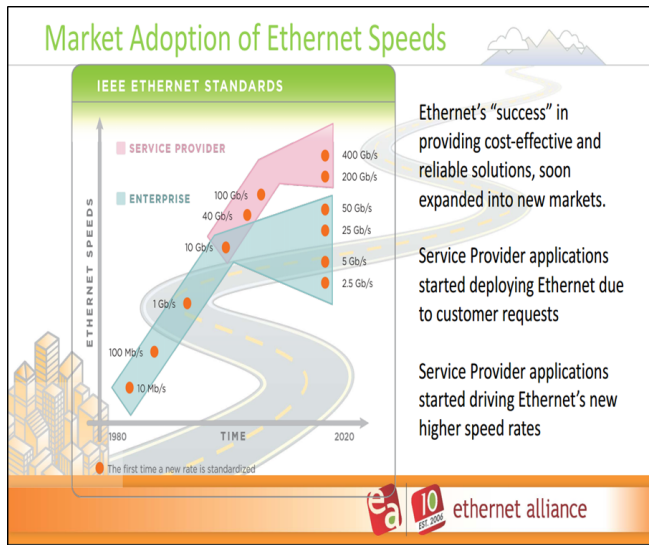


PCI Express 4.0 and 5.0

- **PCI Express 4.0 Specification (16GT/s) finalized and published in October 2017**
 - Numerous vendors confirmed with 16GT/s PHYs in silicon
 - Major IP vendors offering 16GT/s controllers
- **PCI Express 5.0 Specification targeted for release in Q2 2019**
 - Changes targeted to primarily speed upgrade
 - Protocol already supports higher speed via extended tags and credits
 - Existing PHYs in the industry already run at 28GHz / 56GHz
- **PCI Express 5.0 Specification (32GT/s) – ideal for:**
 - Applications such as artificial intelligence, machine learning, gaming, visual computing, storage and networking
 - High-end networking solutions (i.e. 400Gb Ethernet and dual 200Gb/s InfiniBand solutions)
 - Accelerator and GPU attachments for high-bandwidth solutions
 - Constricted form factor applications that cannot increase width and need higher frequency to achieve performance
 - Continued use of L1 Sub-states to constrain power consumption during transmission idle periods

	RAW BIT RATE	LINK BW	BW/ LANE/WAY	TOTAL BW X16
PCIe 1.x	2.5GT/s	2Gb/s	250MB/s	8GB/s
PCIe 2.x	5.0GT/s	4Gb/s	500MB/s	16GB/s
PCIe 3.x	8.0GT/s	8Gb/s	~1GB/s	~32GB/s
PCIe 4.0	16GT/s	16Gb/s	~2GB/s	~64GB/s
PCIe 5.0	32GT/s	32Gb/s	~4GB/s	~128GB/s

Ethernet Evolution





Compressed Streams

Hulu Quality	Internet Rate (Mbps)	MB/s	# SATA Streams	# SAS Streams	PCIe 3.0 x4 Streams	PCIe 4.0 x4 Streams	PCIe 5.0 x4 Streams
SD	1.5	0.1875	3200	6400	21333.3333	42666.6667	85333.33
720p HD	3	0.375	1600	3200	10666.6667	21333.3333	42666.67
1080p HD	6	0.75	800	1600	5333.33333	10666.6667	21333.33
4k Ultra HD	13	1.625	369.2308	738.461538	2461.53846	4923.07692	9846.154
8k Estimate	50	6.25	96	192	640	1280	2560