



Session 303-B

SNIA SSSI PCIe SSD Round Table

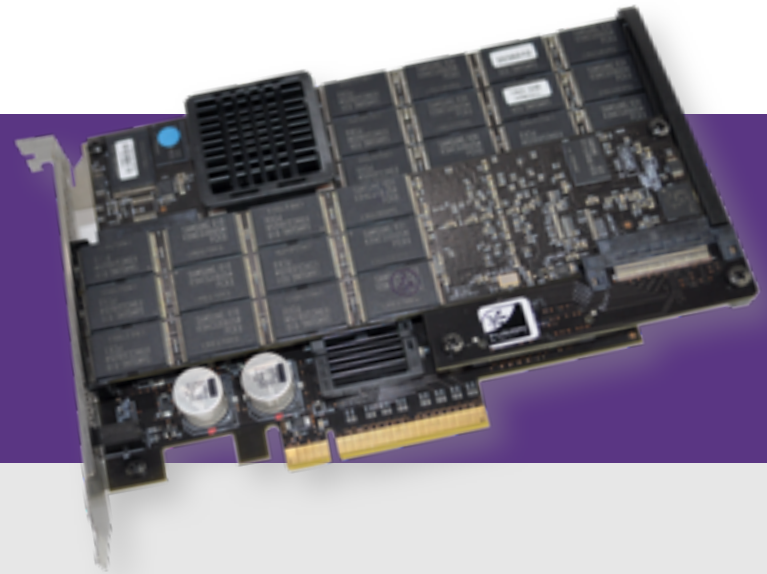
Thursday, August 23, 2012

3:10 pm- 4:25 pm

Eden Kim, Chair

Solid State Storage PCIe . . .

a Round Table



What are issues facing Adoption of PCIe Solid State Storage devices?

- Standards for PCIe Attached Storage
- Technology & Architectural Issues
- Mass Storage Ecosystem Adoption & Optimization
- Market & Product Positioning
- Deployment Strategies





Panelists - Session 303-B

Moderator: Eden Kim, Calypso Systems, Inc.

3:10 – 3:15: Eden Kim, Calypso - Introduction

Speakers

3:15 – 3:20: Easen Ho, Calypso – PCIe Performance Testing

3:20 – 3:25: Gary Orenstein, Fusion-io – Lessons from the Front Lines & A Look to the Future

3:25 – 3:30: Paul Wassenberg, Marvel – SATA Express

3:30 – 3:35: Mark Myers, Intel – SFF Working Group – PCIe 2.5” Form Factor

3:35 – 3:40: Marty Czekalski, Seagate – SCSI Express: Extending the SCSI Platform

3:45 – 3:50: Don Jeanette, Toshiba – PCIe Love & Stuff that needs to Happen

3:50 – 3:55: Tony Rogue, Virident – PCIe & Storage Class Memory

Question & Answer Session – 3:55 – 4:10

Audience Questions

Please fill out Questionnaire Cards



SNIA SSSI PCIe SSD Task Force



Paul Mitchell

Santa Clara, CA August 2012

Open Forum April – July 2012

62 Companies
130 Individuals

- SSD OEMs ODMs
- Controller Companies
- Semiconductor Fabs
- VARs, SAN, NAS
- Analysts, Blogs
- Industry Associations
- Standards Groups
- End Users



SNIA SSSI PCIe SSD Task Force

8 MEETINGS 24 PRESENTATIONS

PRESENTING COMPANIES

HP - Marvell - Micron - Toshiba - Seagate

STEC - Fusion-io - Virident - Intel

NVMe - SATA-IO - STA - PCI SIG

Calypso - Agilent - HyperIO - LeCroy - eASIC

Coughlin Associates - Objective Analysis

SNIA IOTTA - NVMP - SSS - Security TWGs

TOPICS

- **Standards**
- **Testing & Instrumentation**
- **PCIe Performance**
- **System Integration**
- **Form Factors**
- **System Architectures**
- **PCIe Driver Topics**
- **Analysts View**
- **Market Development**
- **Deployment Strategies**

Presentations can be downloaded in mtg minutes at www.snia.org/forums/sssi/pcie
2012

PCIe SSD Standards

ITEM	GROUP → STANDARD	CLIENT PCIe SSD	ENTERPRISE PCIe SSD
PHYSICAL INTERFACE	PCI-SIG → PCI Express SATA-IO → SATA Express SCSI Trade Assn → SCSI Express	up to 2 LANES Drive up to 4 LANES Card	Up to 4 LANES
REGISTER INTERFACE & COMMAND SET	Intel → AHCI NVM Express Group → NVMe T10 → SOP/PQI w/support for NVMe	AHCI or NVMe	SOP / PQI Or SOP / NVMe Or NVMe
HOST CONNECTOR	SATA-IO → SATA Express SSD FF WG → multifunction connector Small FF Committee → SFF-8639	SATA Express Host Connector Or SFF-8639	SFF-8639
DEVICE CONNECTOR	SATA-IO → SATA Express SSD FF WG → multifunction connector Small FF Committee → SFF-8639	SATA Express Drive & Card Connectors	SFF-8639

Advancing Solid State Storage 





Panelist Easen Ho, Calypso

Speaker Name	Company Title	Introductory Bio
Easen Ho	Calypso CTO eho@calypsotesters.com	<p>Dr. Ho is the CTO of Calypso Systems, Inc. and has been a principal architect of the recently released SNIA Solid State Storage Performance Test Specification. Dr. Ho has been intimately involved in the development of performance benchmarking for NAND Flash based solid state storage devices.</p> <p>Dr. Ho received his doctorate in laser physics from MIT and his BSEE from the Tokyo Institute of Technology. Dr. Ho previously was founder, CEO and CTO of digital papyrus,</p>



PCIe SSD Roundtable

Testing of PCIe SSDs

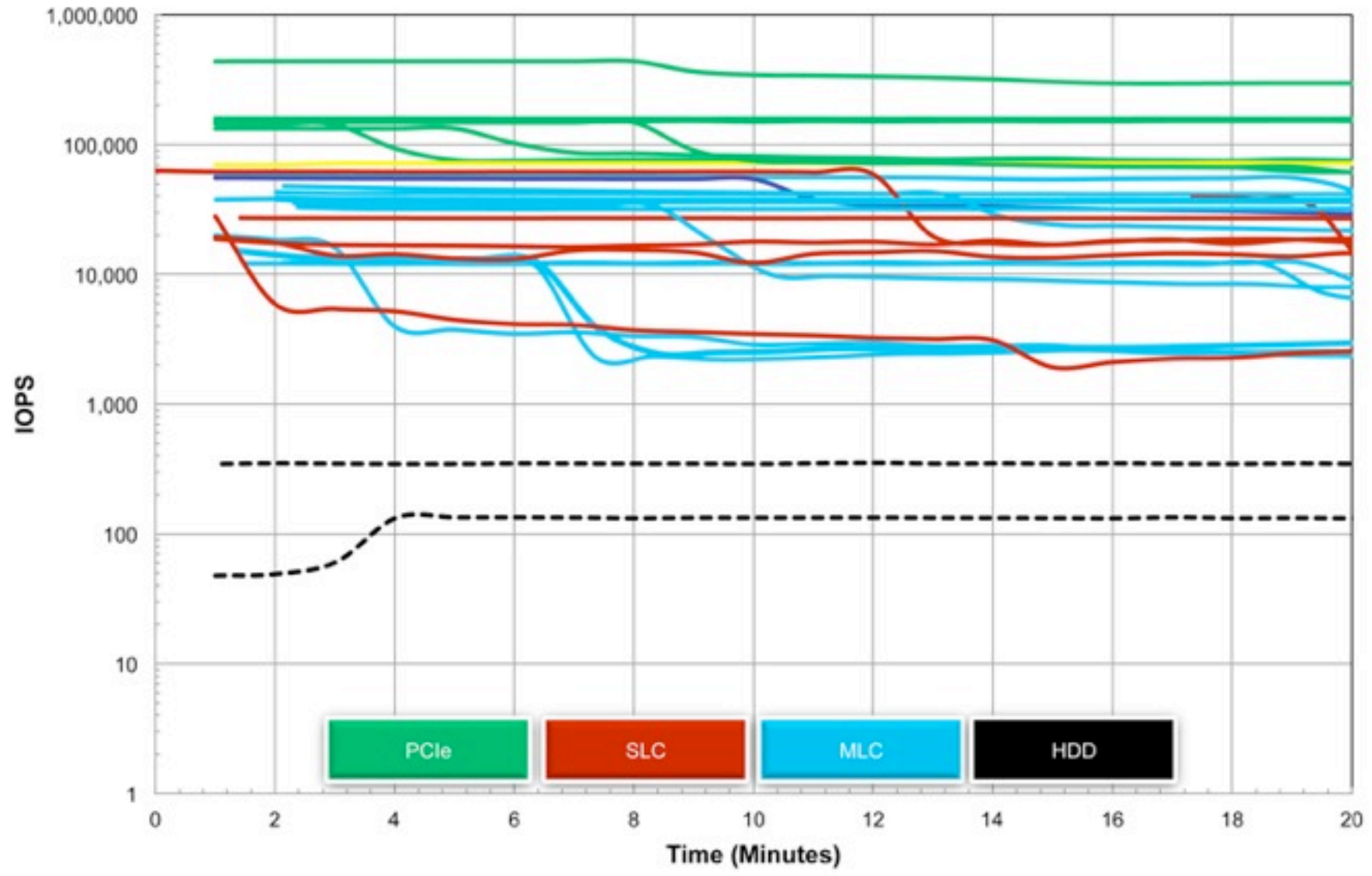
Easen Ho
CTO, Calypso Systems, Inc.

Santa Clara, CA August 2012

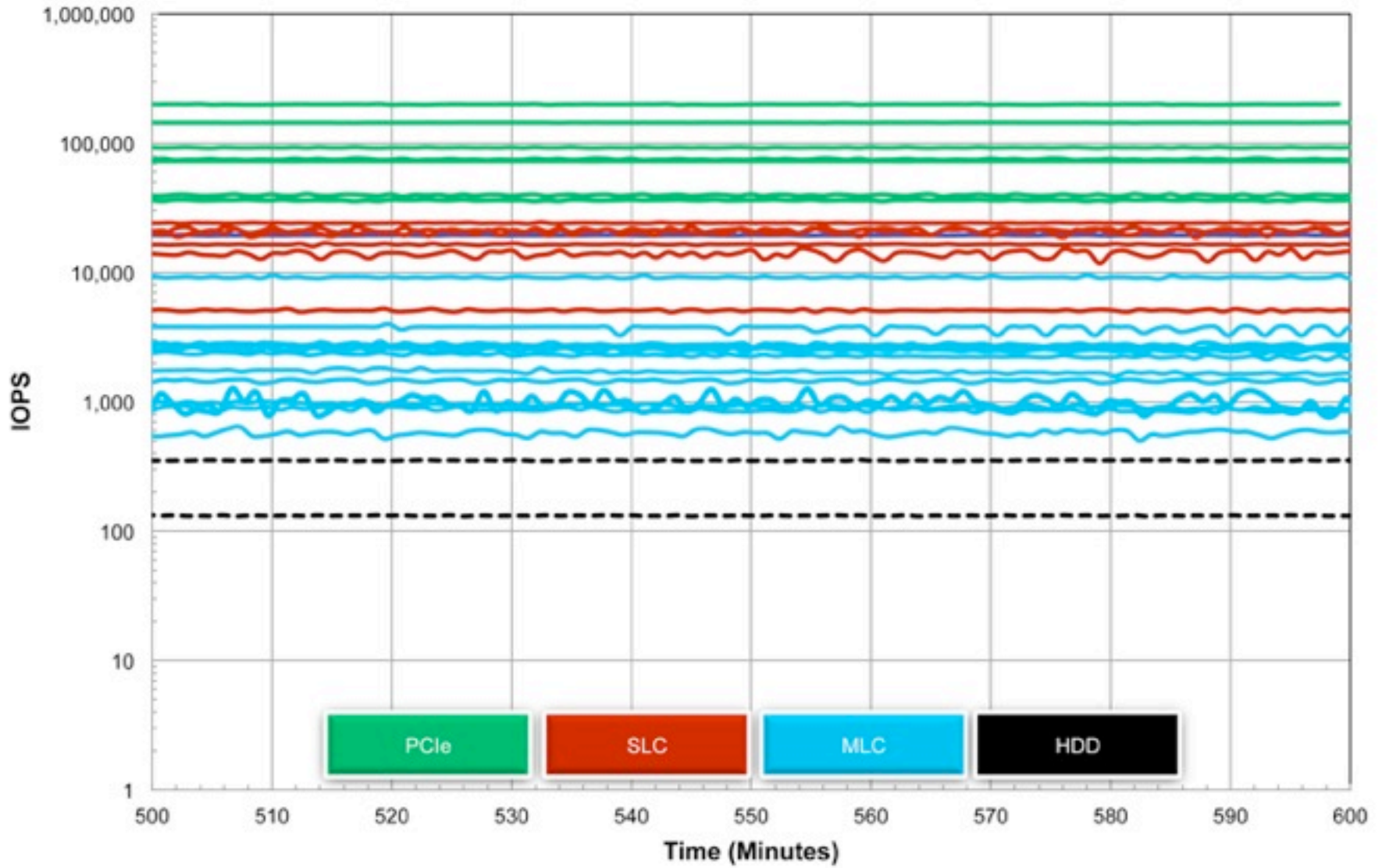
PCIe SSD Testing Issues

- From a testing perspective, PCIe SSDs looks just like another drive to the application, however...
 - Generally targeted at high end → faster TP, IOPS, and latencies
 - Wider variety of architectures possible:
 - no longer gated by a specific protocols such as SATA/SAS → possible to reduced IO latencies
 - host system can become part of the drive by design in some architectures → additional testing metrics needed; standardization becomes important
 - Variety of form factors → heat dissipation; power measurement issues
 - Variety of protocol standards → how do they affect performance?

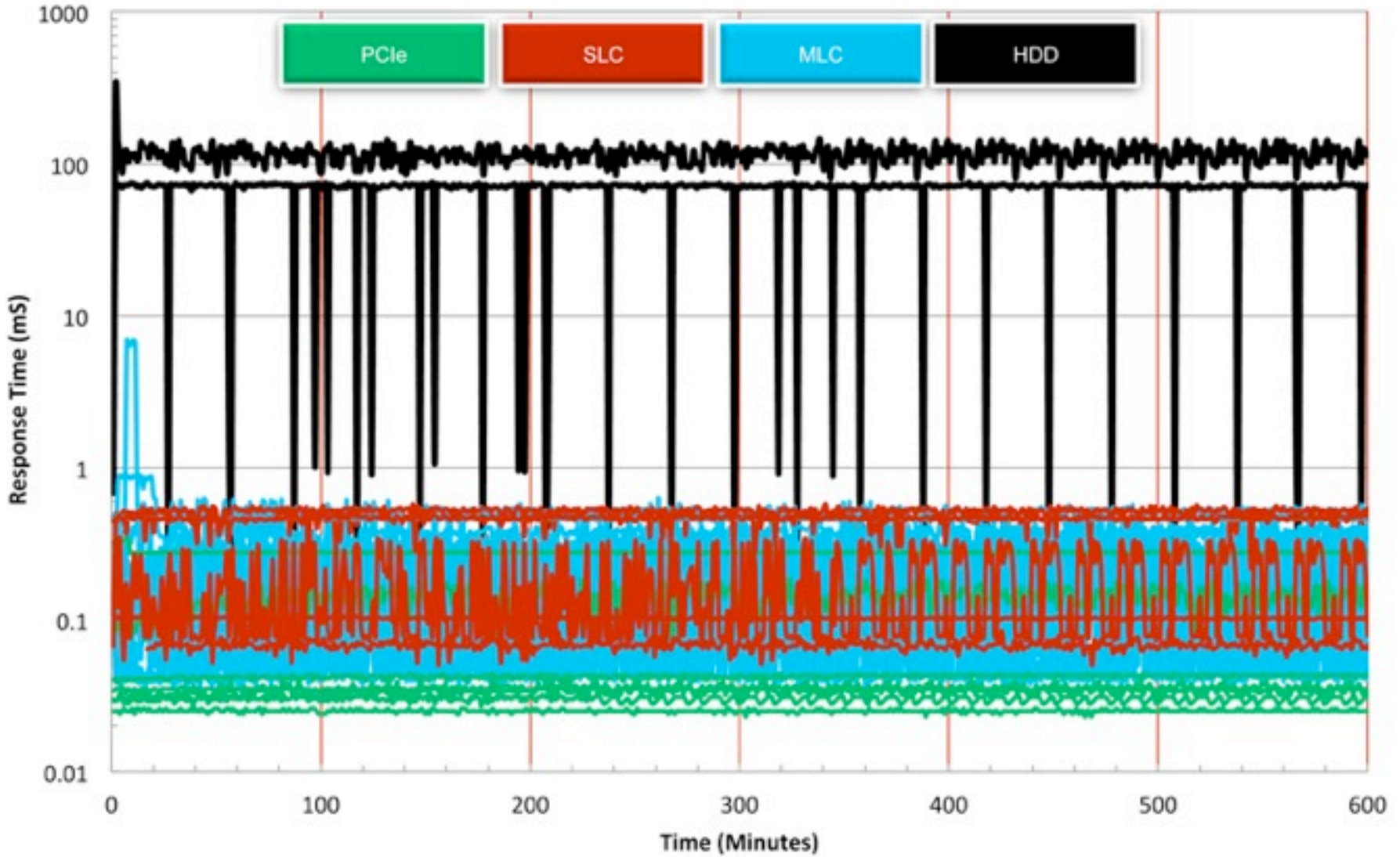
RND/4K Writes, Group By Classes



RND/4K Writes, Group By Classes

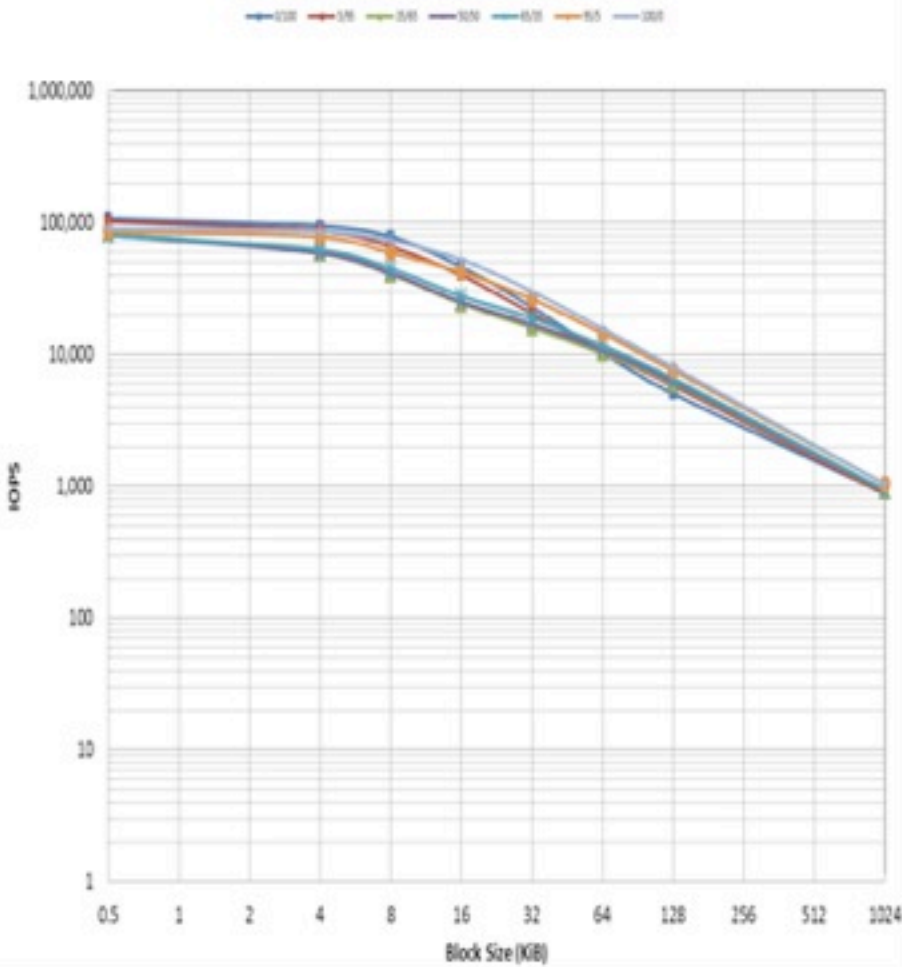


RND/4K Writes: Minimum Response Time – Group By Classes

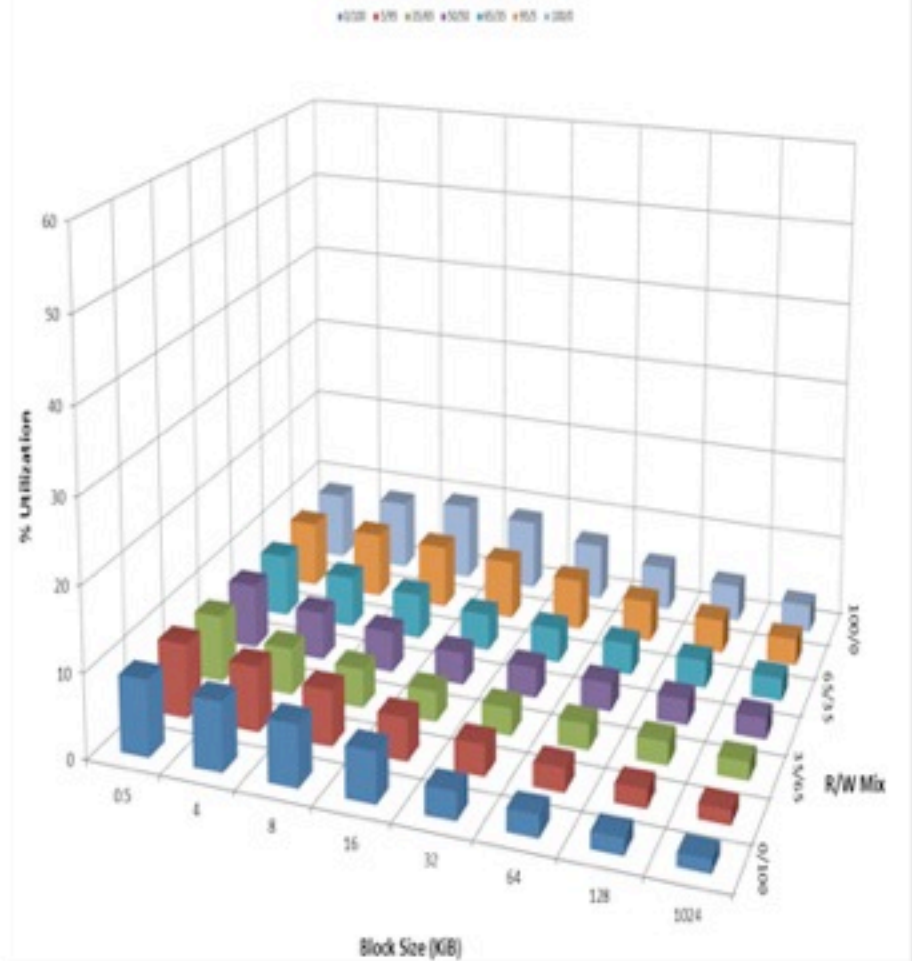


SNIA PTS-E 1.0 IOPS: T2/Q16

P4 2D SS IOPS vs BS-RW

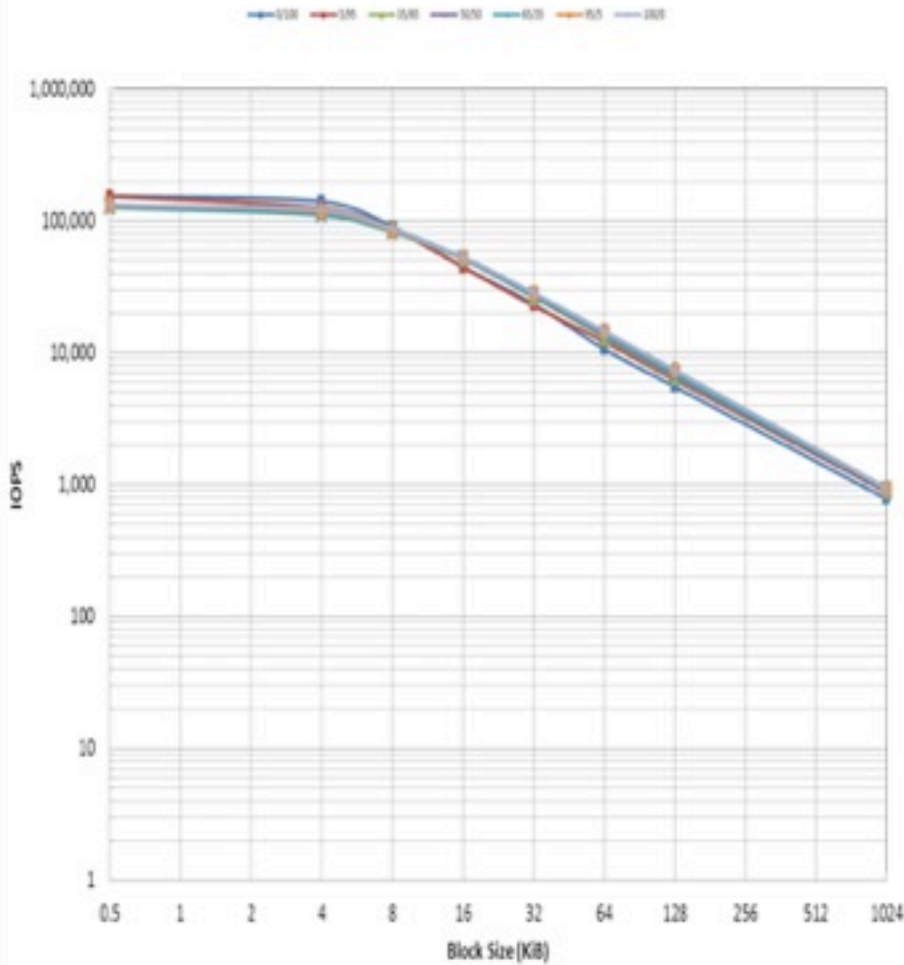


Steady State CPU_SYS vs BS & RW: T2/Q16

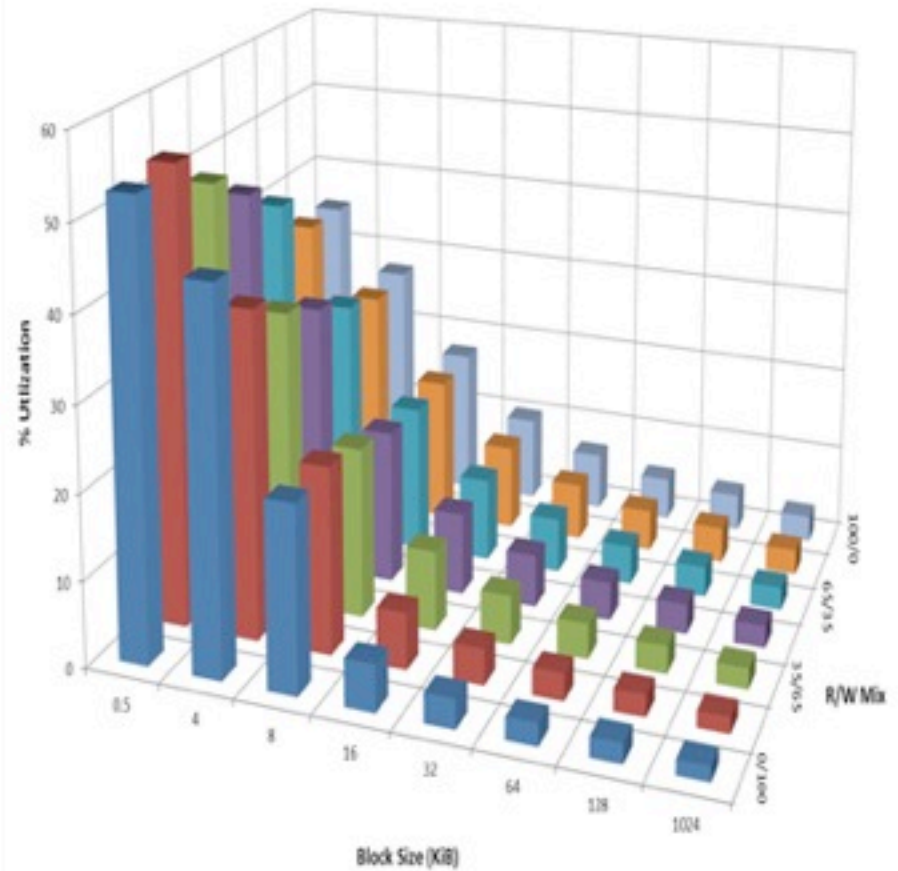


SNIA PTS-E 1.0 IOPS: T16/Q32

P4 2D SS IOPS vs BS-RW



Steady State CPU_SYS vs BS & RW: T16/Q32





Panelist Gary Orenstein, Fusion-io



Speaker Name	Company Title	Introductory Bio
Gary Orenstein	VP Products	<p>Gary has served in leadership roles at numerous data center infrastructure companies. Prior to Fusion-io he was the vice president of marketing at MaxiScale, focused on web scale file systems and acquired by Overland Storage.</p> <p>Prior to MaxiScale, he was the vice president of marketing and business development at Gear6, focusing on storage and web caching. He also served as vice president of marketing at Compellent which went public and 2007, and was a co-founder at Nishan Systems, acquired by McDATA/</p>



August 23, 2012

Lessons from the Front Lines and a Look to the Future

Gary Orenstein, SVP of Products, Fusion-io, @garyorenstein

Initial Stage of NAND Flash deployment

FUSION-iO

Applications

Block I/O



NAND FLASH



Architectural Differences



Applications

Block I/O

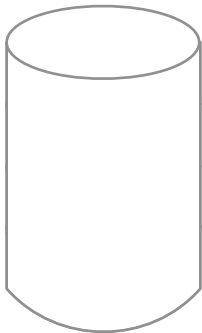
**Architected like
a disk**

**Architected like
memory**

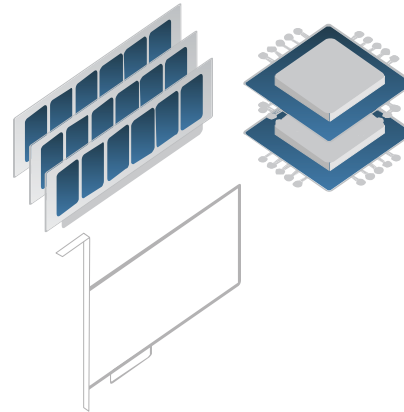
Memory-like Architecture

Applications

Block I/O



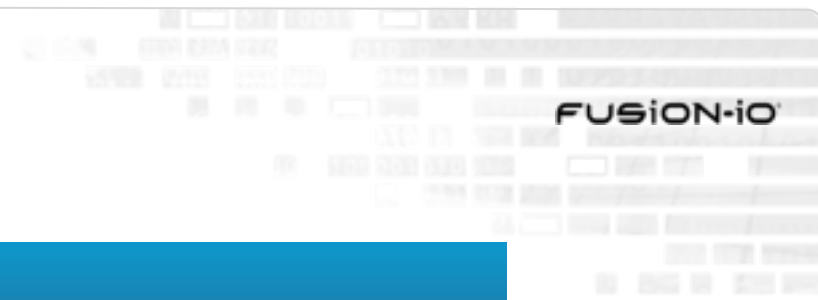
SSD approach



Flash as Memory approach



Beyond Block I/O Access



Applications

Block I/O

**SSD
SAS
SATA**

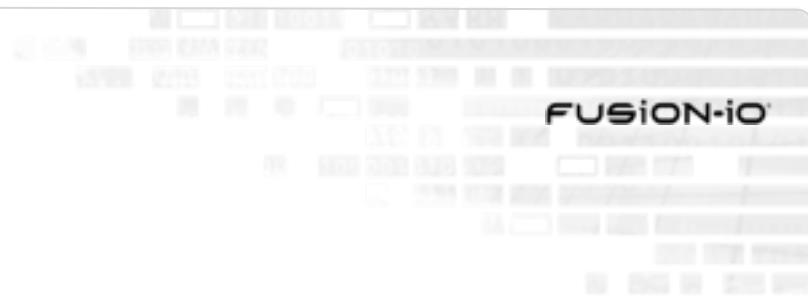
**Block and
Enhanced I/O**

Memory

Flash as Memory

A cartoon-style rearview mirror with a black frame and a white interior. The word "Stores" is written inside the mirror in a red, bubbly, sans-serif font. The mirror is set against a background of a bright blue sky with soft, white clouds.

Stores



amazon web services™



S3 API – Bucket and Object operations

DELETE Bucket

- DELETE Bucket lifecycle
- DELETE Bucket policy
- DELETE Bucket website

GET Bucket (List Objects)

- GET Bucket acl
- GET Bucket lifecycle
- GET Bucket policy
- GET Bucket location
- GET Bucket logging
- GET Bucket notification
- GET Bucket Object versions
- GET Bucket requestPayment
- GET Bucket versioning
- GET Bucket website
- HEAD Bucket
- List Multipart Uploads

PUT Bucket

- PUT Bucket acl; PUT Bucket lifecycle
- PUT Bucket policy; PUT Bucket logging
- PUT Bucket notification; PUT Bucket requestPayment
- PUT Bucket versioning; PUT Bucket website

DELETE Object

- Delete Multiple Objects

GET Object

- GET Object ACL
- GET Object torrent
- HEAD Object
- POST Object

PUT Object

- PUT Object acl
- PUT Object – Copy
- Initiate Multipart Upload

Upload Part

- Upload Part – Copy
- Complete Multipart Upload
- Abort Multipart Upload
- List Parts



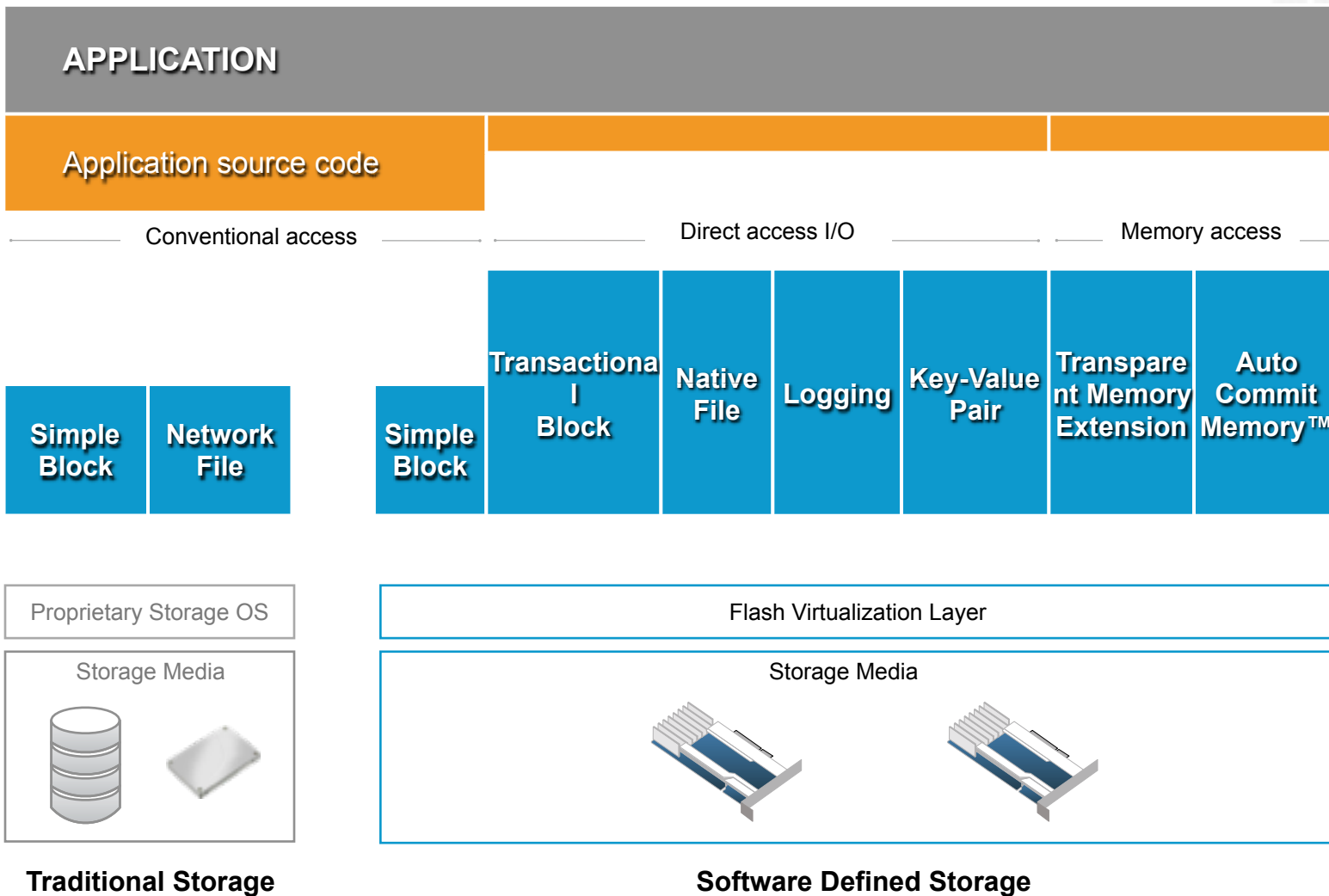
Flash Disk

The background of the slide is a dark blue/black field filled with numerous thin, purple and blue lines radiating outwards from a central point. Each line ends in a small, glowing circular dot, creating a starburst or fiber optic effect. The lines are most dense and bright in the center, fading as they move towards the edges. The text 'Flash Memory' is centered horizontally and vertically over this background.

Flash Memory

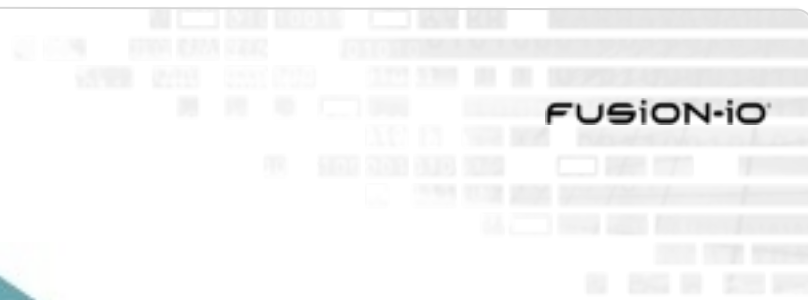


Software Development Kit ADVANTAGES





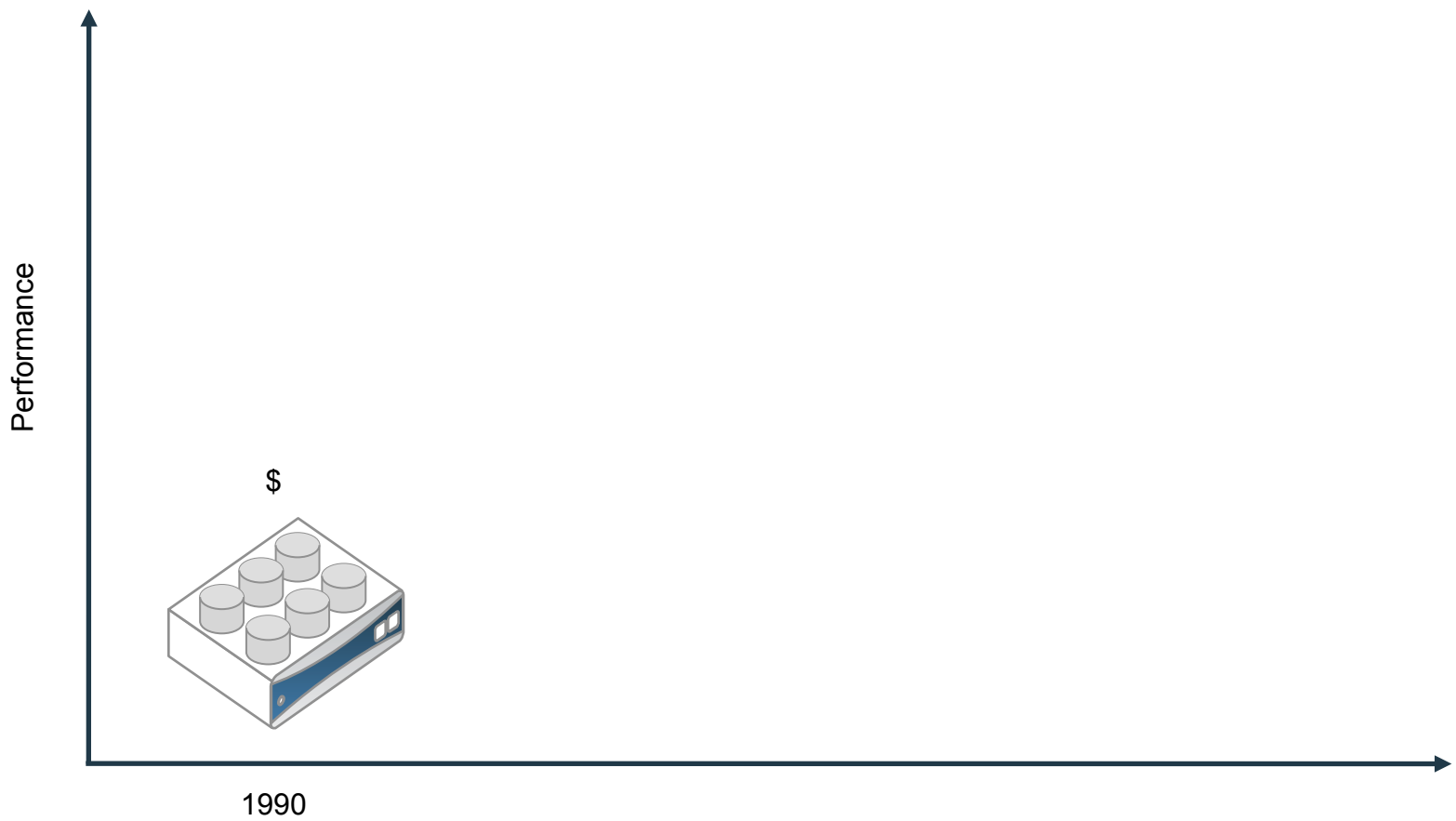
Flash Deployment Architectures





A NEW APPROACH IS NEEDED

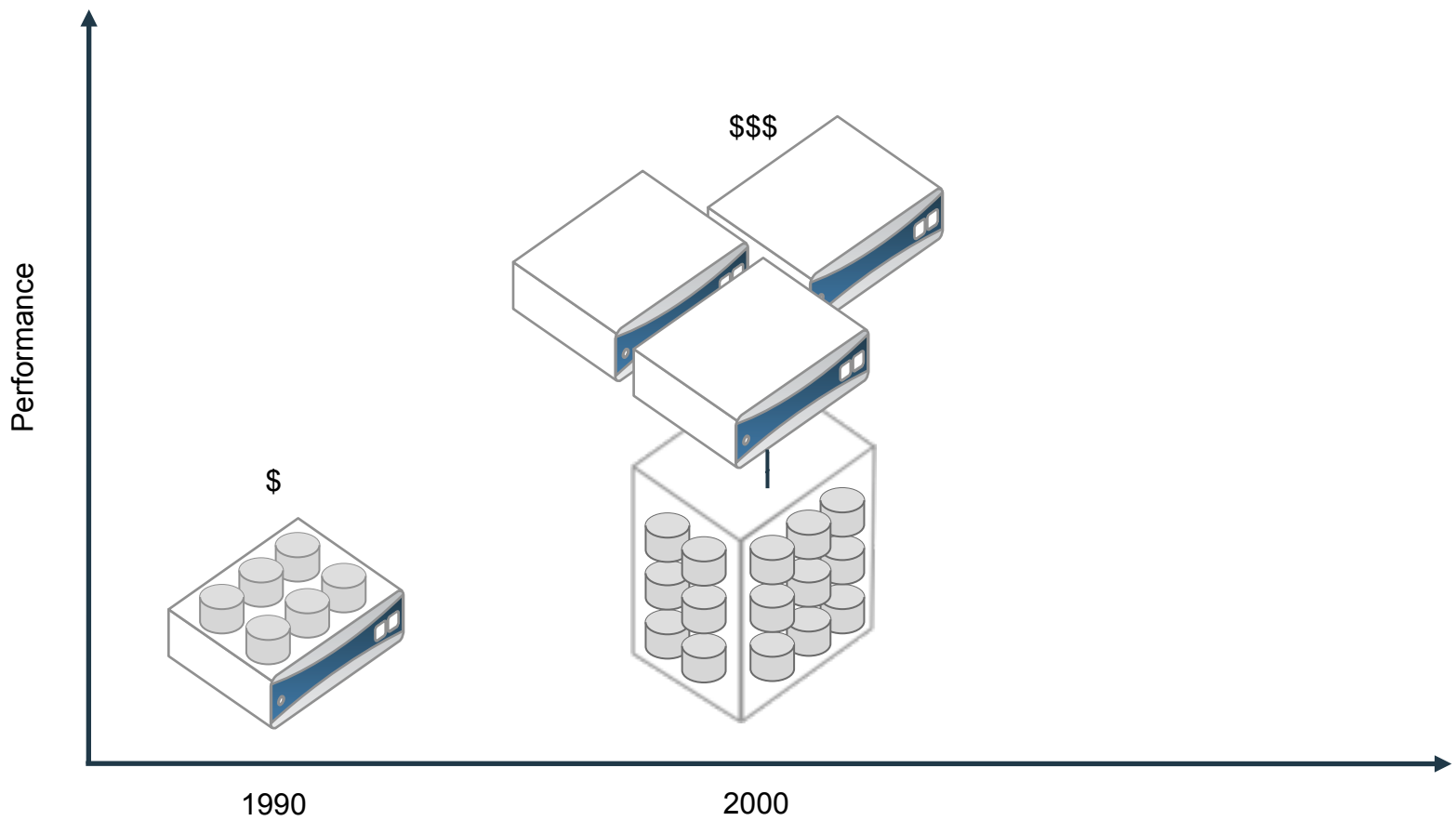
FUSION-io





A NEW APPROACH IS NEEDED

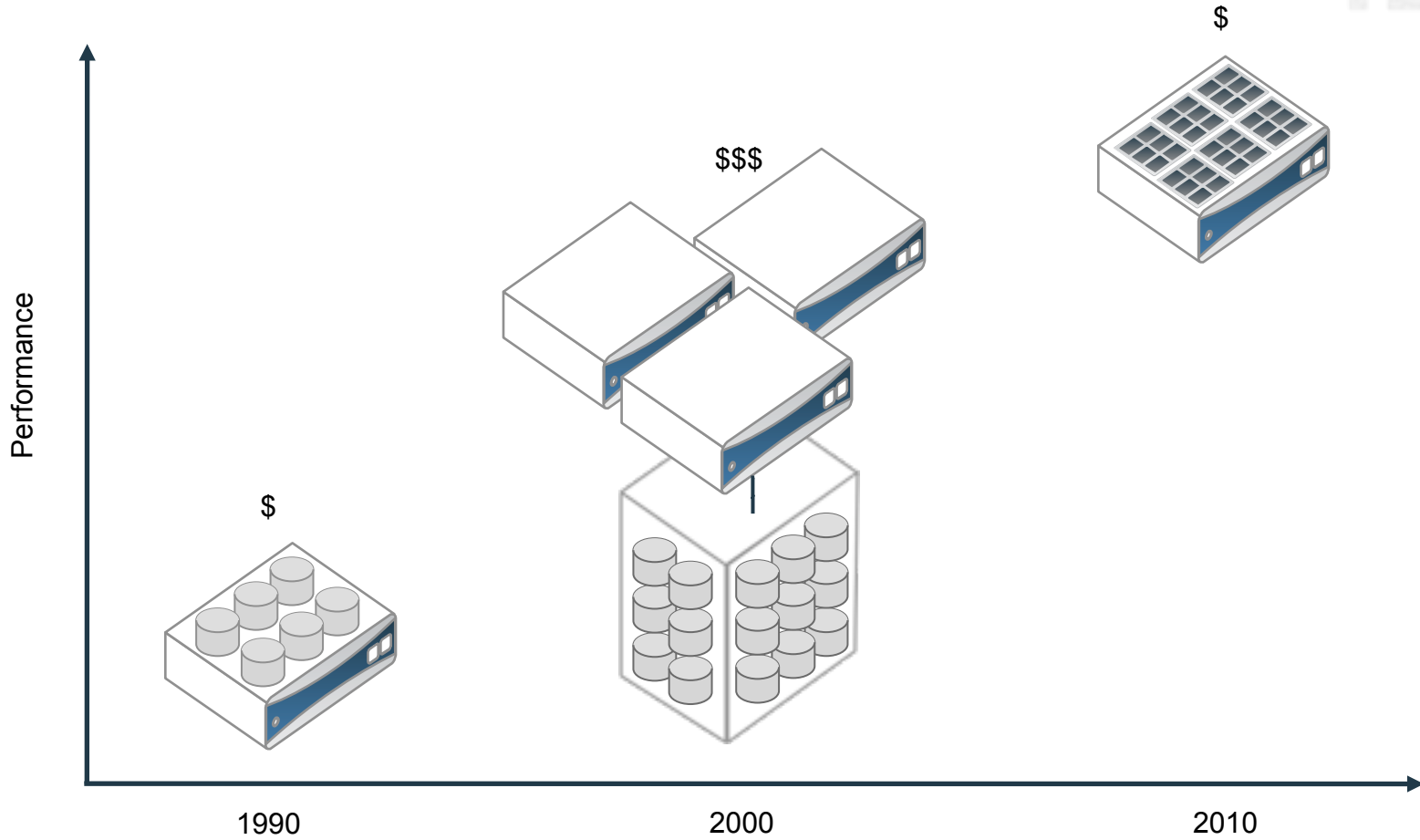
FUSION-io





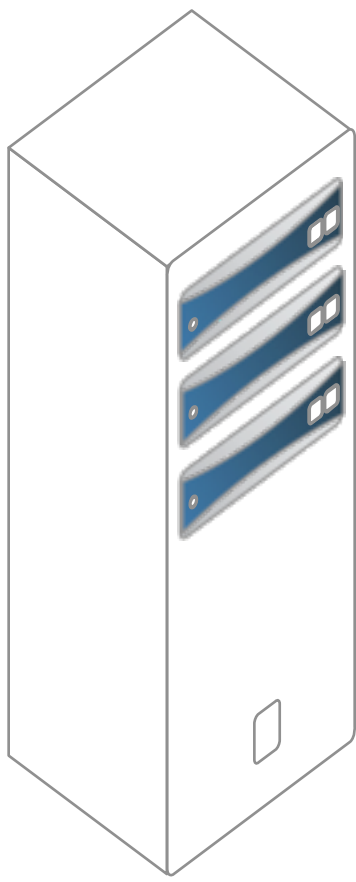
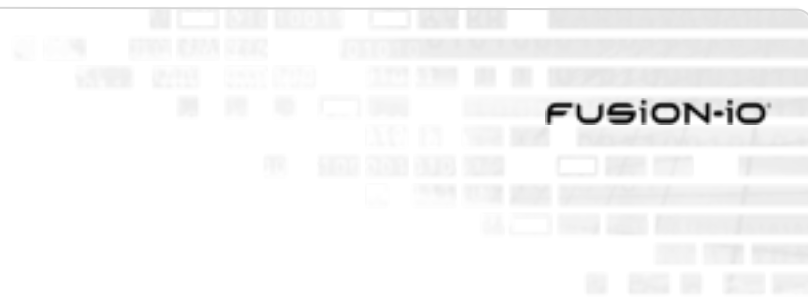
A NEW APPROACH IS NEEDED

FUSION-io





The Power of A SAN

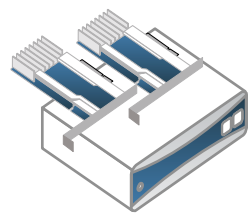
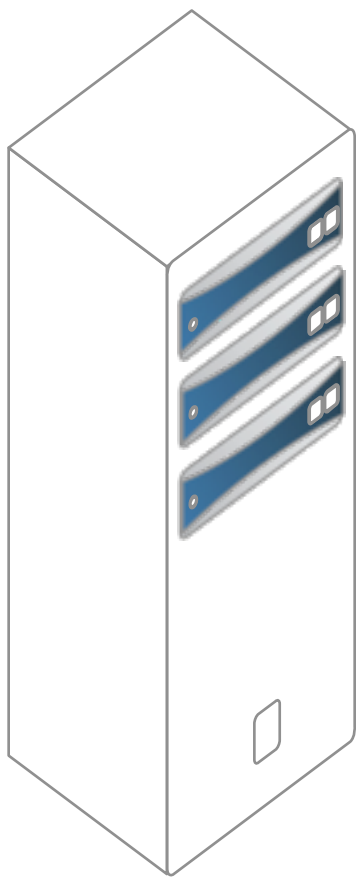




The Power of A SAN

IN A 1 U SERVER

FUSION-io





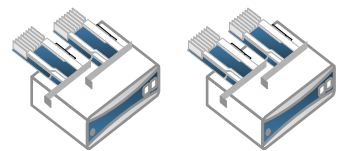
acceleration solutions

FUSION-io

Direct

Max Acceleration

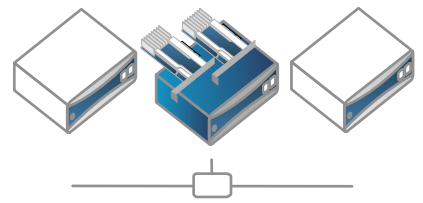
- Lowest latency
- Smallest footprint
- For I/O intensive applications



Shared

Max Scalability

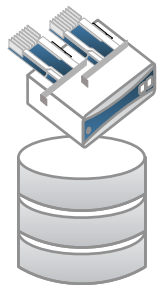
- Multi-protocol
- Platform Independent
- For clustered architectures



Caching

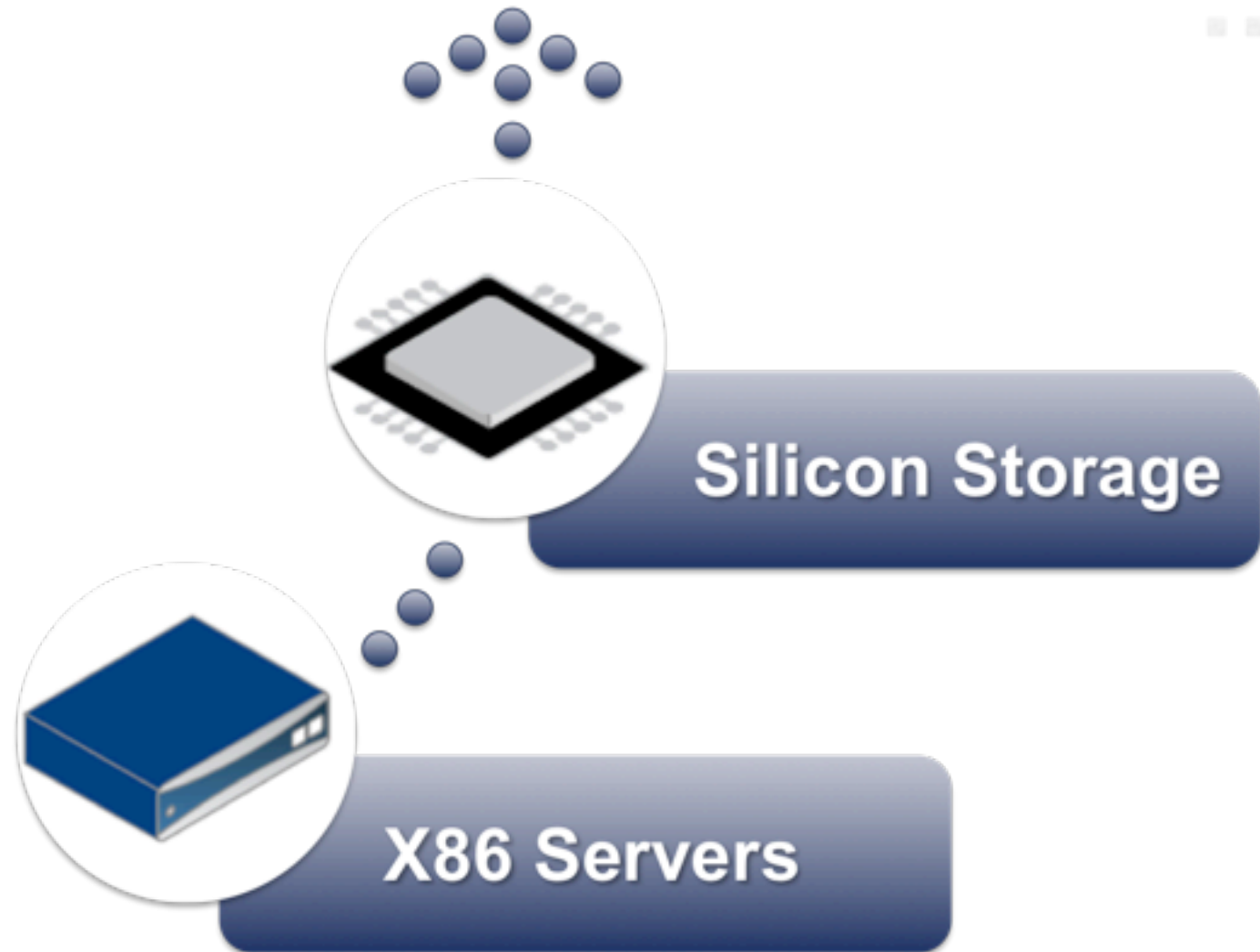
Max Interoperability

- Drop-in SAN/NAS acceleration
- Storage workload reduction
- Greater VM density





Market Growth Accelerators





Thank You

@garyorenstein

Panelist Paul Wassenberg, Marvell



Speaker Name	Company Title	Introductory Bio
Paul Wassenberg	Marvell	<p>Paul Wassenberg has over 20 years of experience in data storage and has been deeply involved with storage interface technology, including SATA since its inception. Early in his career, he was a storage controller designer, before moving into Marketing in the HDD industry, and eventually into storage semiconductors.</p> <p>Paul currently holds the position of Director, Product Marketing with Marvell Semiconductor. In that role, he has responsibility for transceiver technology and HDD/SSD storage standards. He is on the SATA-IO board of directors and chairs the SNIA Solid State Storage Initiative. Paul holds</p>



SATA Express

Evolving SATA for High-Speed Storage

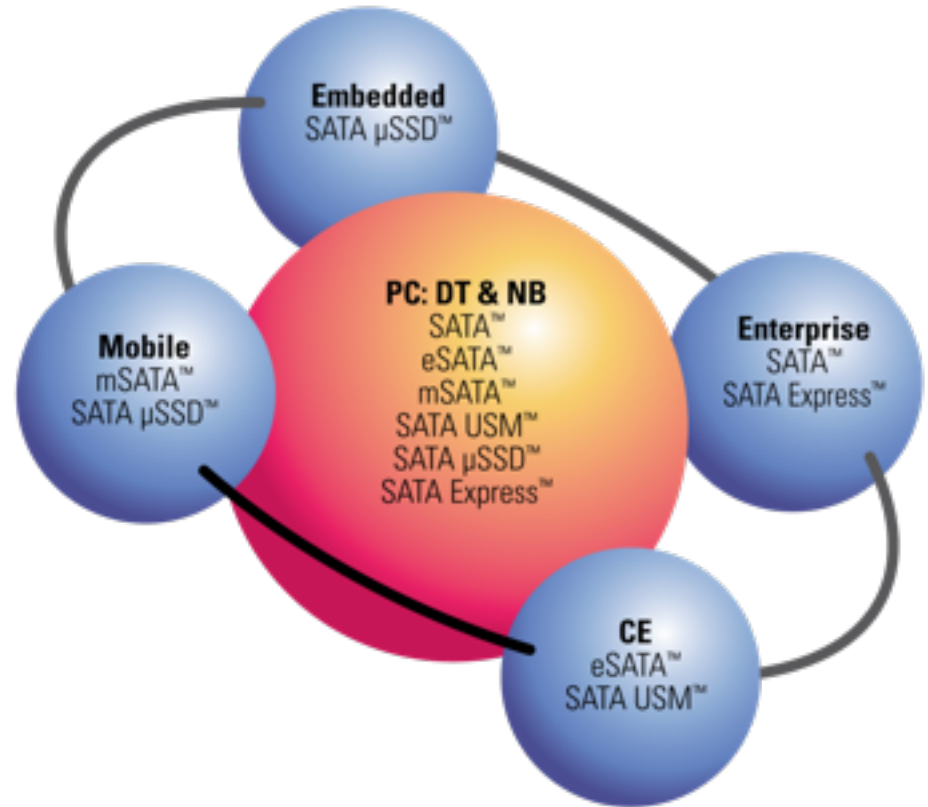
Paul Wassenberg
SATA-IO Marketing Chair

Santa Clara, CA August 2012

SATA is Everywhere

SATA is the de facto standard for PC storage

Since its introduction, SATA has evolved into new application spaces and now provides storage interface solutions for HDDs, ODDs, SSDs, and Hybrid HDDs in client, mobile, enterprise, CE, and embedded storage markets



A Need for Speed

Today, most applications are well-served by SATA 6Gb/s and will be for the foreseeable future

However, some client SSDs and Hybrid HDDs (HHDD) will soon require greater speeds than those enabled by the current generation of SATA

SATA-IO is developing SATA Express which utilizes PCI Express® (PCIe®) as the physical interface



Why SATA Express?

- Some client SSDs & HHDDs will soon require more than 6Gb/s
 - 12Gb/s SATA would take too long; PCIe is here now
- With the next speed increase, the client storage infrastructure has to change in any case
 - Whether 12Gb/s or PCIe
- SATA Express must be low cost
 - 6Gb/s SATA will be more than adequate for HDDs for the foreseeable future
 - Portion of client SSDs that will require greater than 6Gb/s is fairly small
 - Difficult for 12Gb/s to achieve low cost within a reasonable timeframe – PCIe is already widely used

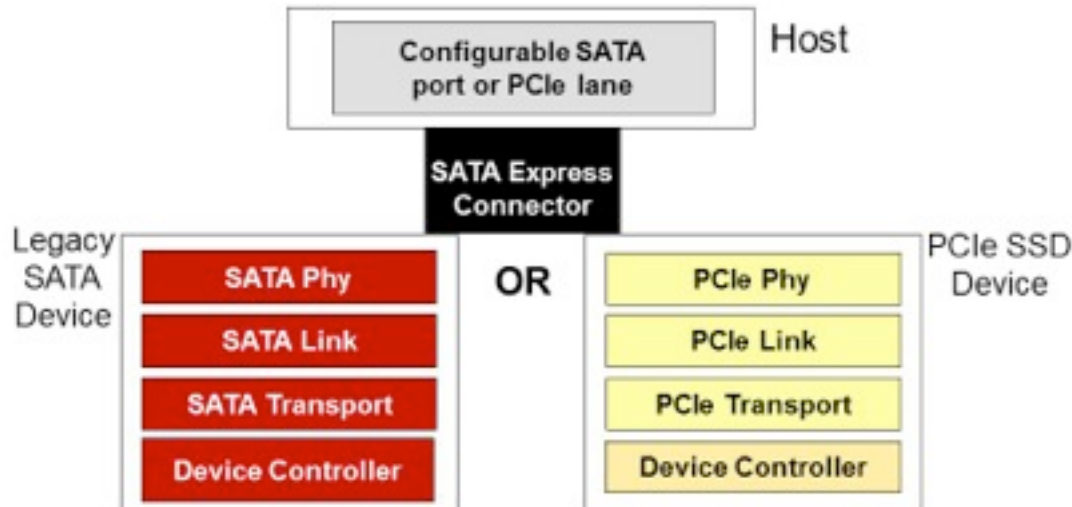
What is SATA Express?

- Client storage with a PCIe interface, utilizing SATA Express connectors
- Provides up to 8Gb/s and 16Gb/s
 - One lane or two lanes of PCIe Gen 3, 2 or 1
- Defines new device and host connectors to support both new SATA Express and current SATA devices



SATA Express Connectivity

- The SATA Express host connector can mate with a SATA Express device or a SATA device
 - A signal driven by the device tells the host whether it is connected to a SATA Express or a SATA device



- A SATA Express device can also mate with the SFF-8639 connector for enterprise applications



SATA Express Software Architecture

- Although not defined by the specification, there are two choices for register interface / command set:
 1. AHCI, which is used for SATA, would enable a SATA Express device to be compatible with SATA software environments
 - AHCI is supported in most major O/Ses
 - But AHCI is not optimized for SSD performance
 2. NVMe Express is architected for high performance SSDs
 - But NVMe does not provide SATA software compatibility
 - Drivers for Windows, Linux, and VMWare are available at



Next Steps and Timeline

- SATA Express is currently under development within SATA-IO
- Completed specification expected by late 2012
- In the meantime, SATA-IO will continue to optimize the existing SATA infrastructure for a wide variety of applications
- SATA will continue to be the mainstream storage interface for the foreseeable future



For More Information...

- Go to the SATA Express page on the SATA-IO site www.sata-io.org/technology/sataexpress.asp
- Check out the NVM Express site at www.nvmexpress.org
- Download the SFF-8639 connector specification at <ftp://ftp.seagate.com/sff/SFF-8639.PDF>

Panelist Mark Meyers, Intel



Speaker Name	Company Title	Introductory Bio
Andrew Ku	Intel	<p>Mark is a Server Platform Architect working in Intel's Datacenter and Connected System group.</p> <p>Mark is technical chair of the Enterprise SSD Form Factor WG which includes definition of proposed SFF-8639 connector.</p> <p>Mark has been at Intel for 12 years in various server and IO architecture projects.</p>



PCIe SSD 2.5" Form Factor

Mark Myers
Datacenter Platform Architect
Intel

Santa Clara, CA August 2012

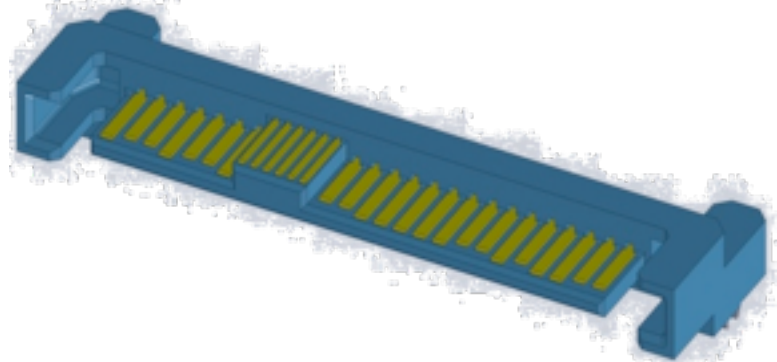
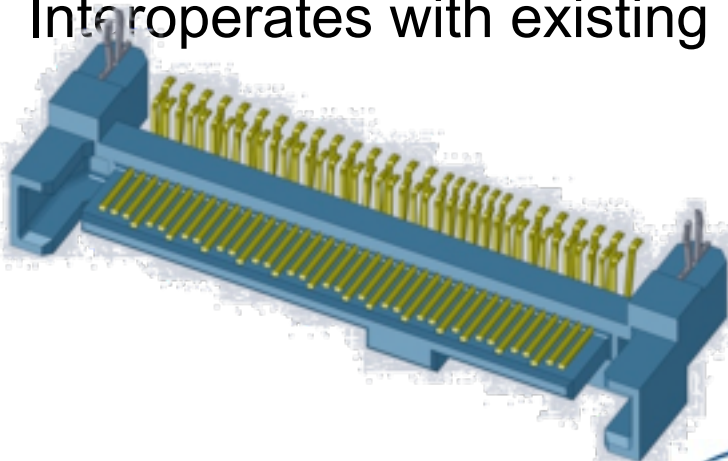


Some Notes about this template

- **PCIe value**
 - Industry standard, high BW, multilane, low latency
 - Flexible attach models, discoverable, and supports many form factors
 - Our work adds a classic 2.5” disk form factor
- **PCIe for high performance; Coexisting with many other**
 - Hard Disks stay on SATA/SAS for long time
 - High performance SSD will 1st to move to PCIe
 - higher BW & low latency
 - PCIe supports multiple device types:
 - NVM-Express, SOP, proprietary, expect interface models to evolve as devices improve
 - Many form factors: Client’s NGFF, 2.5” drives, PCIe Cards

Enterprise SSD Form Factor WG Status

- Defined usages and requirement and connector
 - 5 promoters; >50 contributor companies
- Rev 1.0 Specification Approved <http://www.ssdformfactor.org/>
 - Mechanical piece is SFF-8639 <ftp://ftp.seagate.com/sff/SFF-8639.PDF>
- Extends SAS connector with pins all across both sides
 - Interoperates with existing SATA/SAS connector



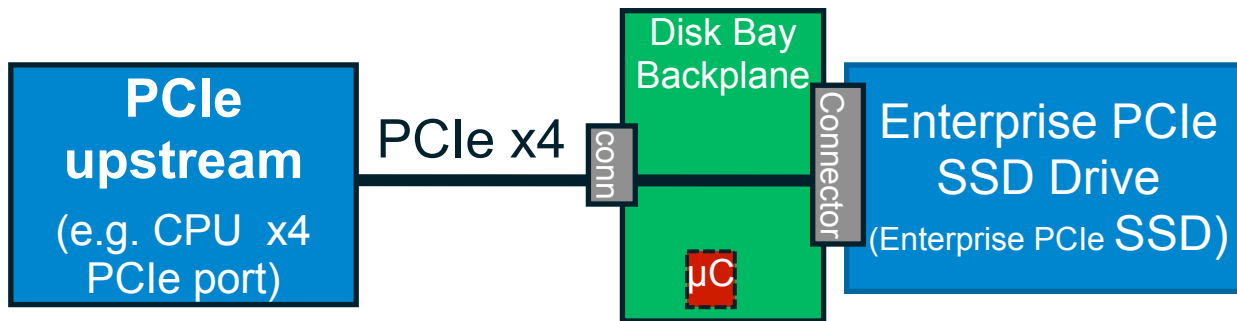
Santa Clara, CA August 2012



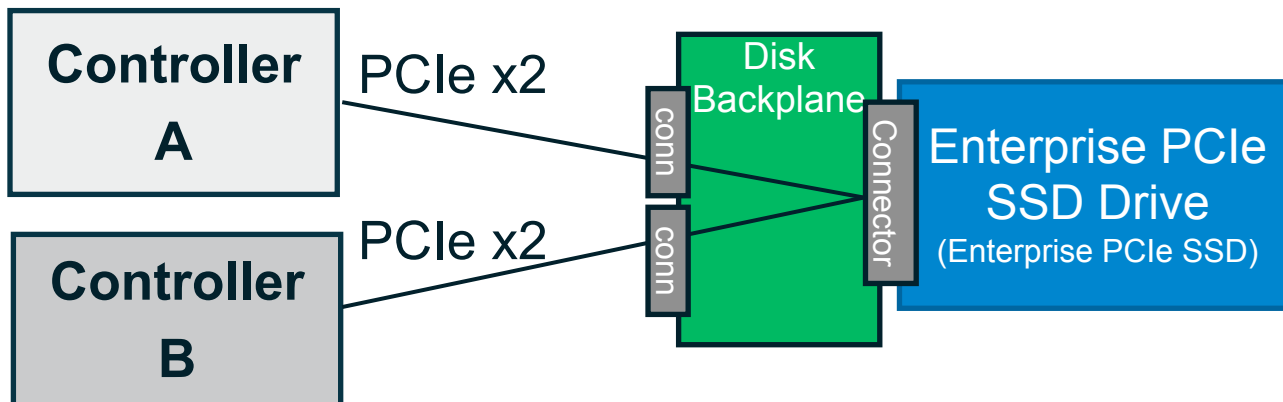


Common Usages: Server x4, Dual Port for Storage

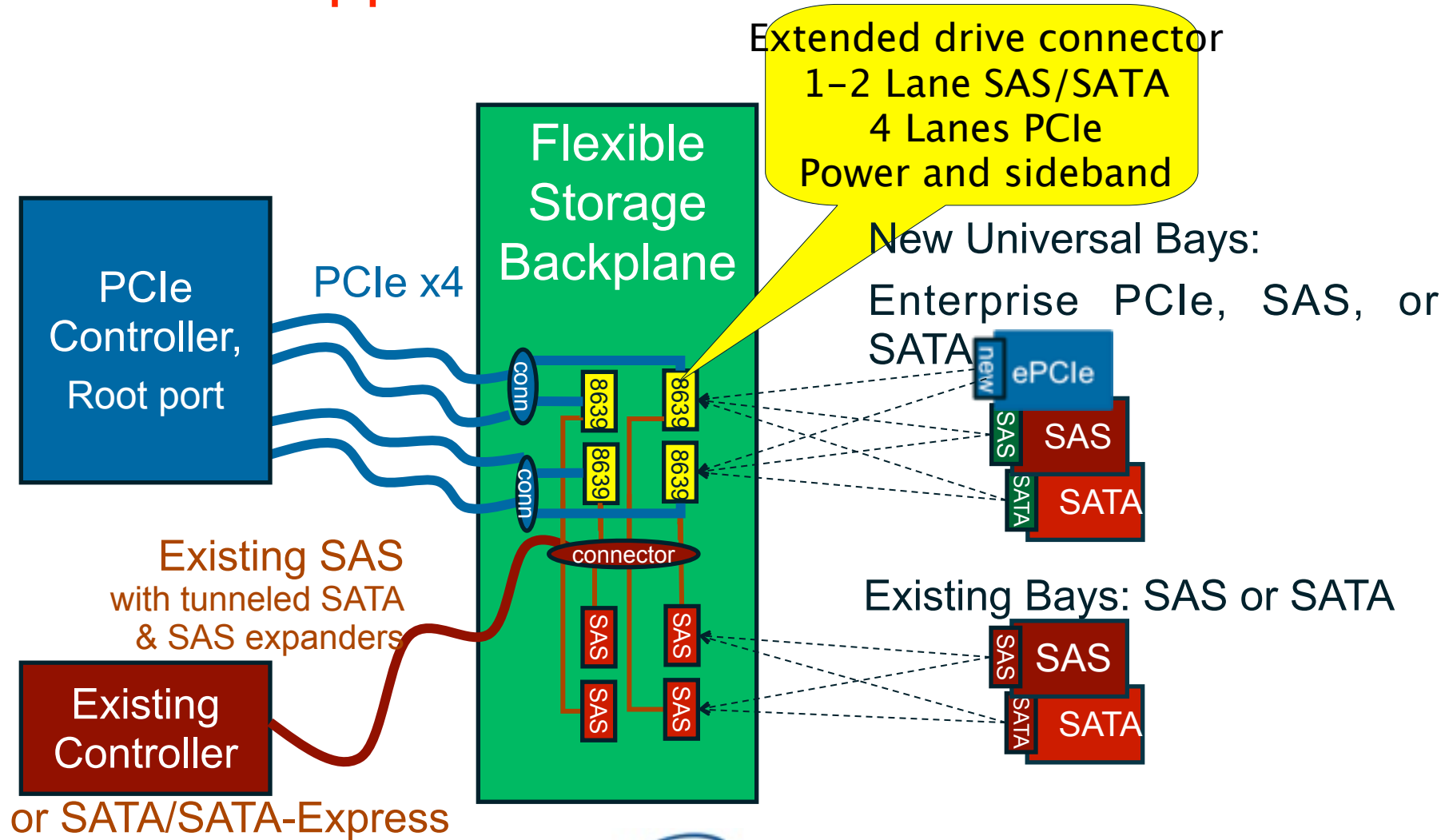
Typical Server configuration



Typical High Availability Storage configuration



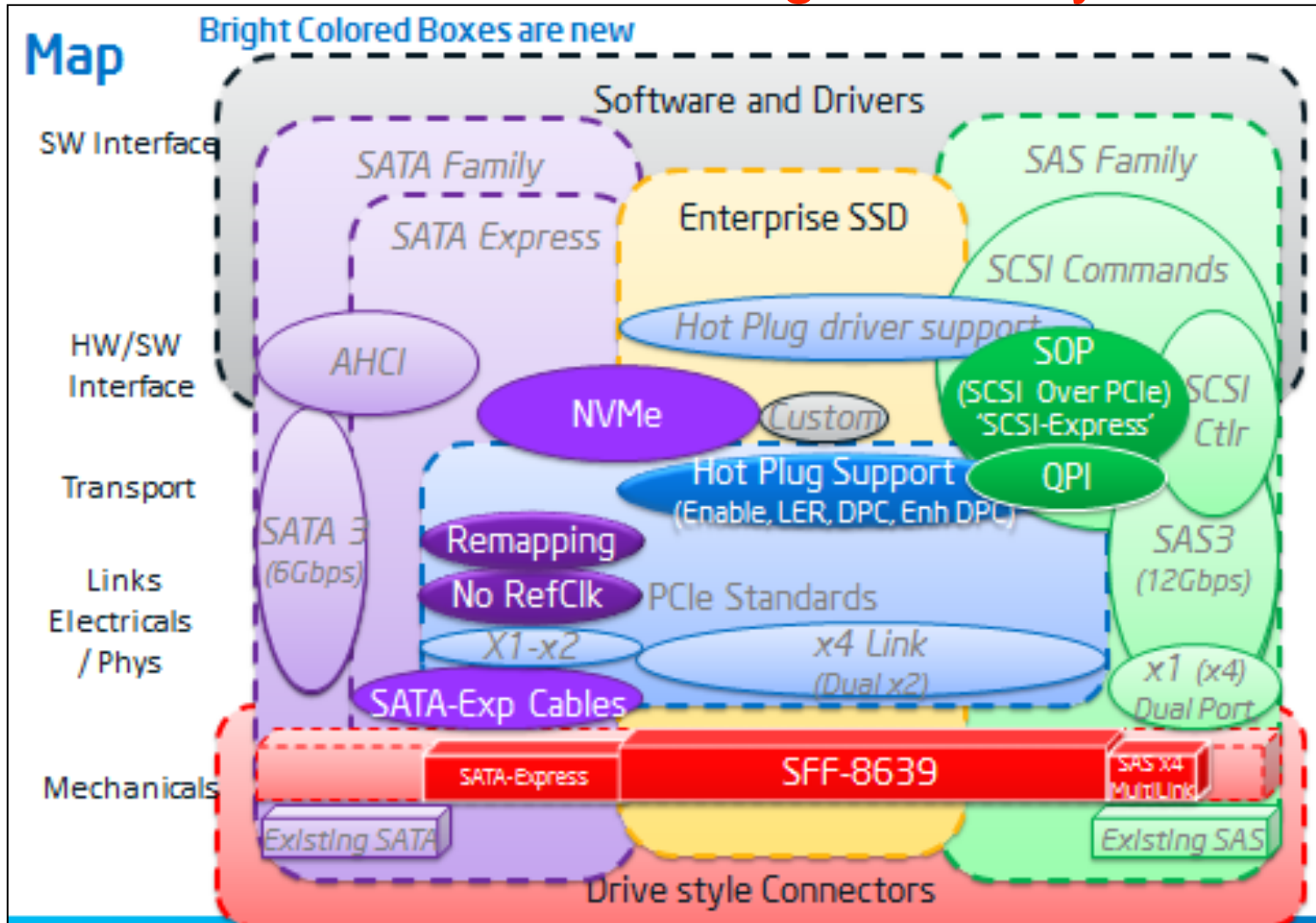
Flexible Backplane Supports PCIe and SAS/SATA



Santa Clara, CA August 2012



Layered and Completing Standards - Common in the Storage industry



Santa Clara, CA August 2012



- Enterprise PCIe SSD Form Factor Specification
 - PCIe 2.5” drives are in the market
- 2.5” supports Flexible Storage Backplanes
 - High Performance Enterprise x4 PCIe SSDs
 - Existing SAS/SATA drives
 - Emerging SATA-Express and x4 SAS
- Many Standards, provide choices and innovation opportunities



Panelist Marty Czekalski, Seagate



Speaker Name	Company Title	Introductory Bio
Marty Czekalski	Seagate	<p>Marty Czekalski brings over thirty years of senior engineering management experience in advanced architecture development for Storage and IO subsystem design, ASIC, and Solid State Storage Systems.</p> <p>He is currently Sr. Staff Program Manager within Seagate's Strategic</p>



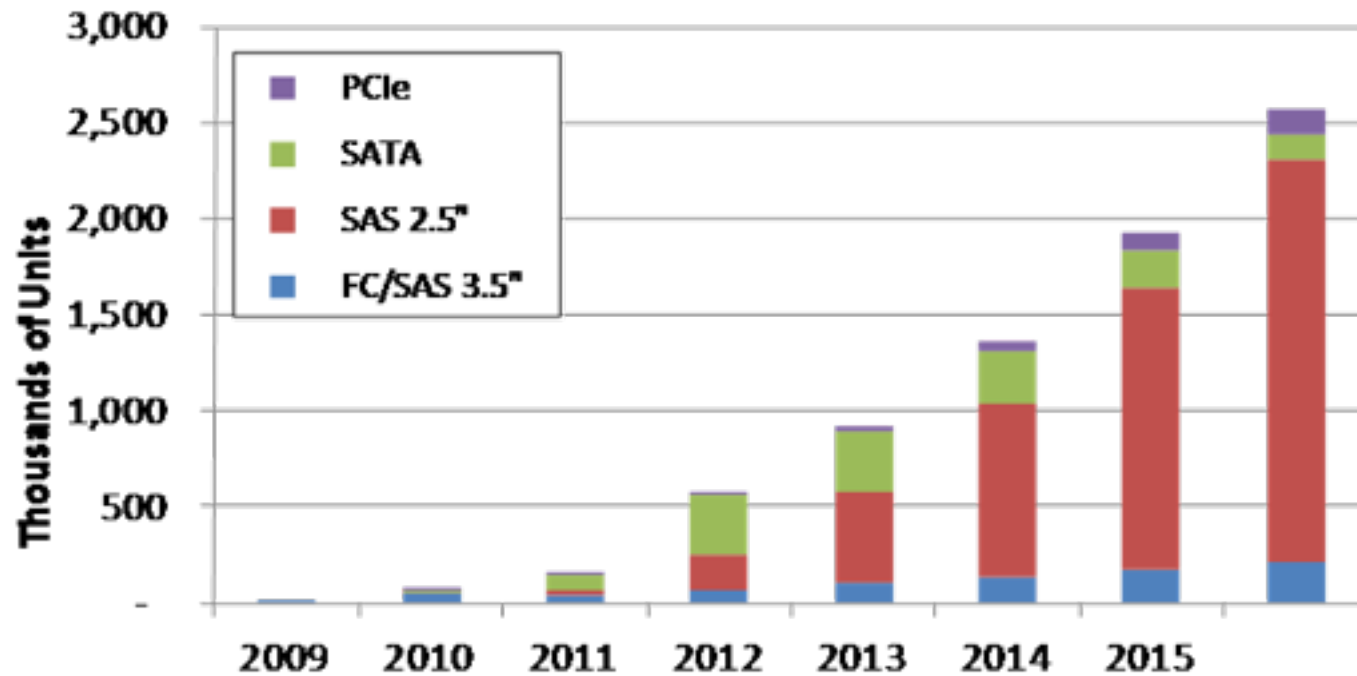
Extending the SCSI Platform of Innovation



Santa Clara, CA August 2012

SAS is the preferred SSD Interface for Storage Systems

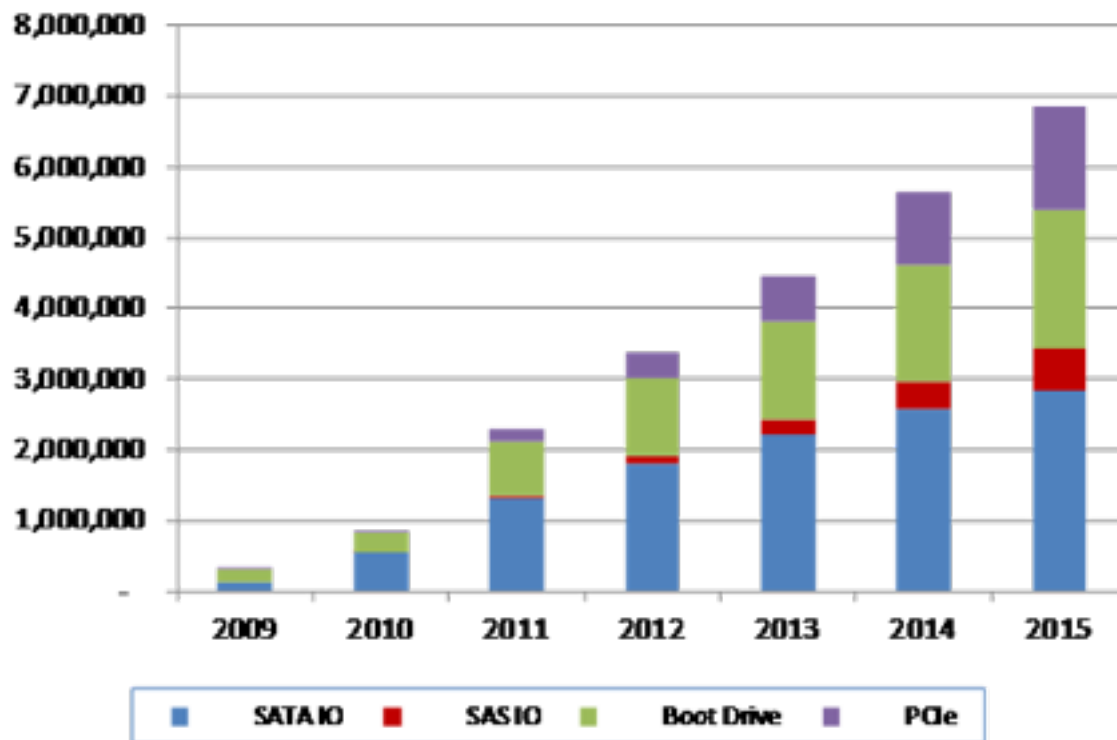
Storage-attached SSD Units



Forward-Insights 11-2011

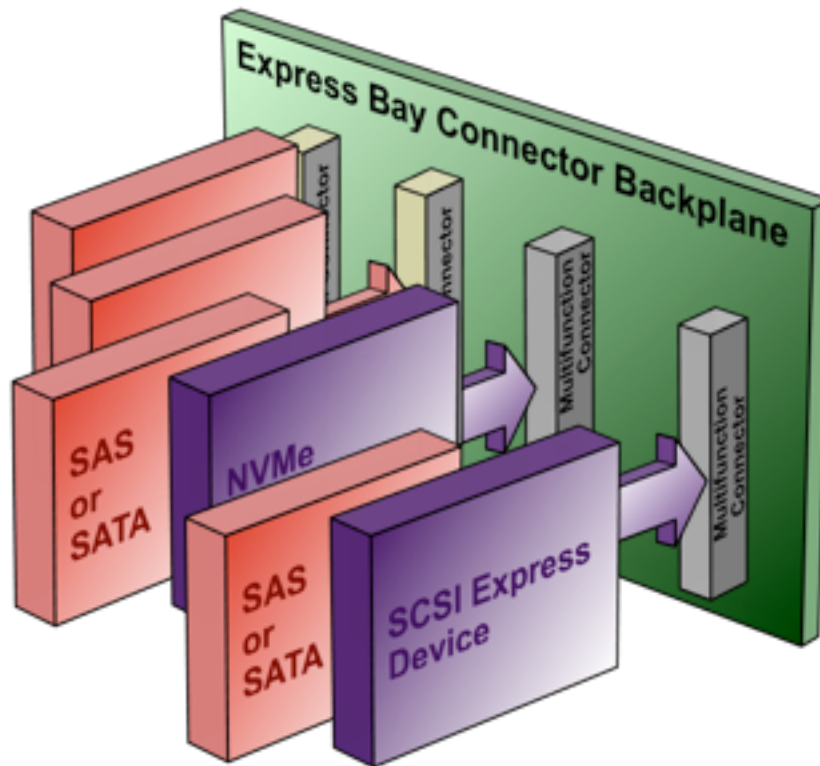
Server Attached SSDs

Server-attached SSD Units



Forward-Insights 11-2011

Express Bay



Express Bay

- Up to 25 Watts
- SFF-8639 connector
- PCI-SIG electrical specification

Objectives

- Preserve the enterprise storage experience for PCI Express storage
- Meet SSD performance demands
- Serviceable, hot-pluggable Express Bay opens up new possibilities...

SCSI Express Overview



Proven SCSI protocol combined with PCIe creating an industry standard path to PCIe-based storage

- Enterprise storage for PCIe based storage devices
- Increased performance through lower latency
- Coexistence with SAS via Express Bay and common command set
- Unified management and programming interface

STA Member Companies



SCSI Express Components

Existing industry initiatives delivering enterprise storage using PCI Express

Technology

Description

SCSI

The storage command set

SCSI Over PCIe (SOP)

Packages SCSI for a PQI queuing layer

PCIe Queuing Interface (PQI)

Flexible, high-performance queuing layer

Express Bay connector (SFF 8639)

Accommodates PCIe, SAS, and SATA drives

PCI Express

Leading server I/O interconnect

SCSI/SAS – Looking to the Future



- 12Gb/s, 24Gb/s SAS
- Extended Copy Feature
- Power Limit Control - up to 25W SAS devices
- Atomic Writes
- Hinting & other NVM features
- SCSI Express (SOP/PQI spec out for letter ballot)

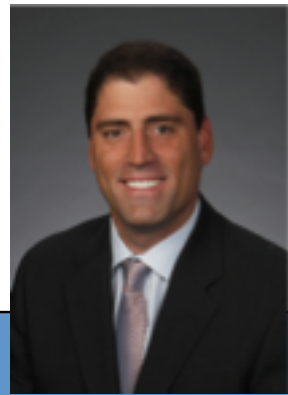


So where do we go from here?

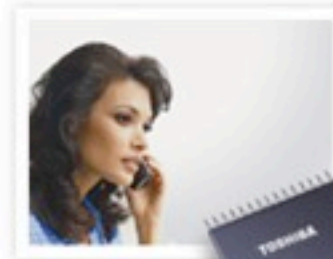
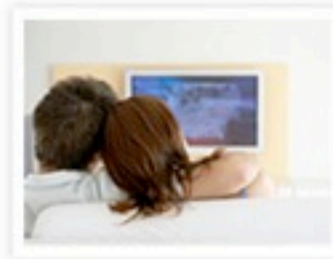
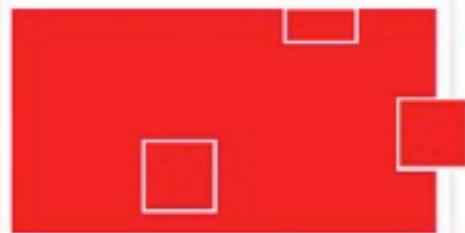
- SAS controllers > 1 Million IOPS and increased power for SAS drives diminish PCIe SSD differentiation
 - Standardized SAS solutions exist today and will continue to be deployed
 - Increased power in Express Bays allow SAS devices to achieve similar performance levels to PCIe devices
- **PCIe SSD Storage Call to Action**
 - Once the PCIe capable bays are available, any PCIe device can be packaged in a 2.5" FF and used, in as long as a driver exists.
 - SCSI Express, NVMe, proprietary, non-storage devices, etc.
 - Interoperability – Electrical spec for SFF-8639 (Express Bay) started
 - Compliance tests
 - Hot plug – work underway (DPC and enhanced DPC)
- **New form factors will emerge**
 - How will they effect the market?



Panelist Don Jeanette, Toshiba



Speaker Name	Company Title	Introductory Bio
Don Jeanette	Director, Product Marketing Toshiba America Electroni	Don Jeanette is Toshiba's Director of Product Marketing. He is responsible for Storage Products Marketing and New Business Development for Toshiba's Hard Disk Drives and Solid State Drives. He brings to the



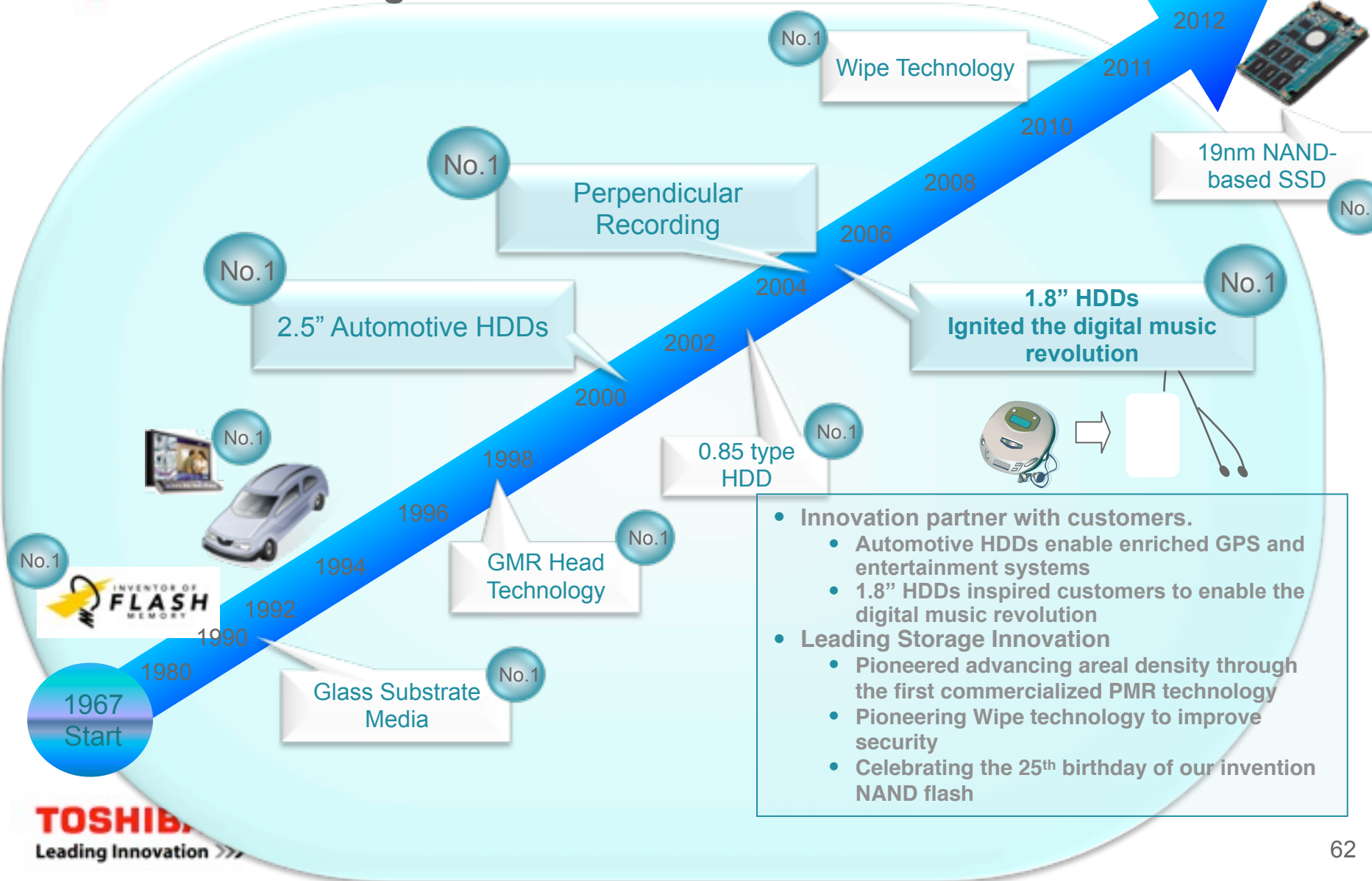
ELECTRONIC COMPONENTS

Flash Memory Summit

**Toshiba America Electronic Components, Inc.
Storage Products Business Unit**

August 2012

Toshiba Storage – 44 Years Of Innovation



- Innovation partner with customers.
 - Automotive HDDs enable enriched GPS and entertainment systems
 - 1.8" HDDs inspired customers to enable the digital music revolution
- Leading Storage Innovation
 - Pioneered advancing areal density through the first commercialized PMR technology
 - Pioneering Wipe technology to improve security
 - Celebrating the 25th birthday of our invention NAND flash

TOSHIBA
Leading Innovation >>>

- PCIe Love and Stuff That Has To Happen



- PCIe Love and Stuff That Has To Happen
 - Hottest issues in SSD



■ PCIe Love and Stuff That Has To Happen

- Hottest issues in SSD
 - Performance



■ PCIe Love and Stuff That Has To Happen

- Hottest issues in SSD
 - Performance
 - Cost



■ PCIe Love and Stuff That Has To Happen

- Hottest issues in SSD
 - Performance
 - Cost
 - Reliability



■ PCIe Love and Stuff That Has To Happen

- Hottest issues in SSD
 - Performance
 - Cost
 - Reliability
 - NAND management



■ PCIe Love and Stuff That Has To Happen

- Hottest issues in SSD
 - Performance
 - Cost
 - Reliability
 - NAND management
 - Retention



■ PCIe Love and Stuff That Has To Happen

- **Hottest issues in SSD**
 - Performance
 - Cost
 - Reliability
 - NAND management
 - Retention
- **PCIe - Why we love it**
 - Least amount of latency!
 - High bandwidth!
 - Maximum performance between the host and the SSD!

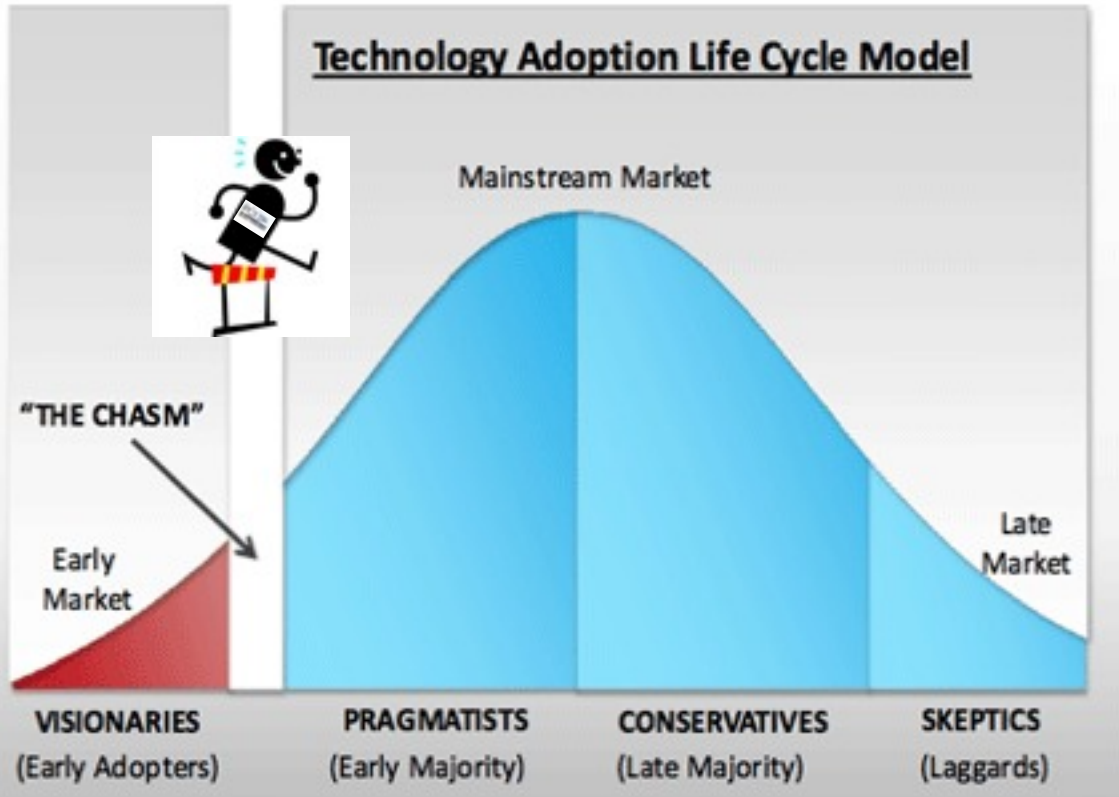


■ PCIe Love and Stuff That Has To Happen

- **Hottest issues in SSD**
 - Performance
 - Cost
 - Reliability
 - NAND management
 - Retention
- **PCIe - Why we love it**
 - Least amount of latency!
 - High bandwidth!
 - Maximum performance between the host and the SSD!
- **Today... internet data centers... tomorrow the World!**
 - But hold on... not so fast.....



■ Crossing the Chasm – What Will It Take?



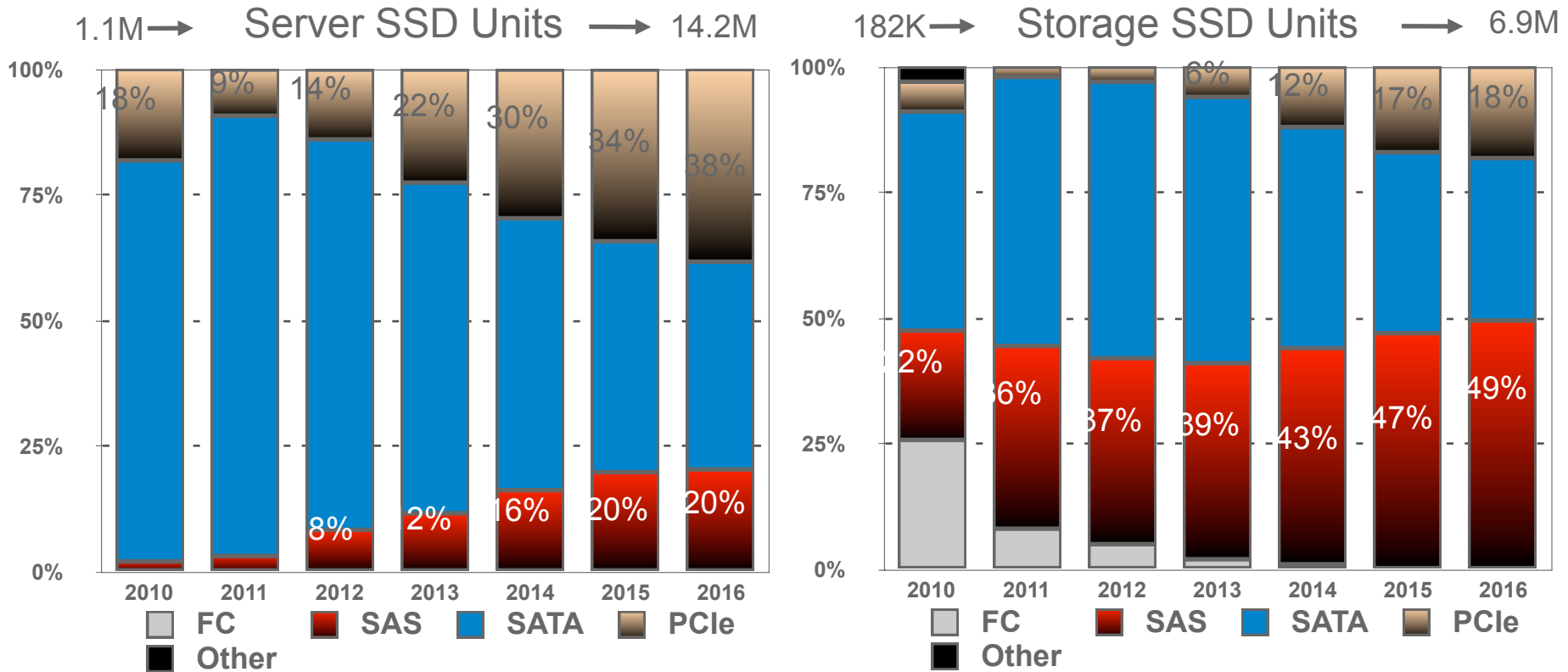
Geoffrey Moore

- What will it take?
 - Unified standards
 - Driver, OS and software development
 - Hot pluggability
 - Physical access to PCIe SSD
 - Familiarity with technology
 - Acceptance of technology
 - Continuity of supply
- And Remember....
 - Other offerings may be ‘good enough’

TOSHIBA
Leading Innovation >>>

Image source: http://www.dshen.com/blogs/business/archives/when_an_investment_thesis_moves_beyond_you.shtml

Evolution of Enterprise Grade SSD Units by Interface



- PCIe steady adoption reaches >1M units in 2013; 31% of Total Enterprise Shipments in 2016
- SATA dominates Total Enterprise shipments with 7.7M units in 2015 with SAS at 6.0M units

TOSHIBA
Leading Innovation >>>

Source: Gartner June 2012 "Q2 '12 NAND: Sinking Before Upwelling"

TOSHIBA

TOSHIBA
Leading Innovation >>>



Panelist Tony Rogue, Toshiba

Speaker Name	Company Title	Introductory Bio
Tony Rogue	Virident	Tony Rogue

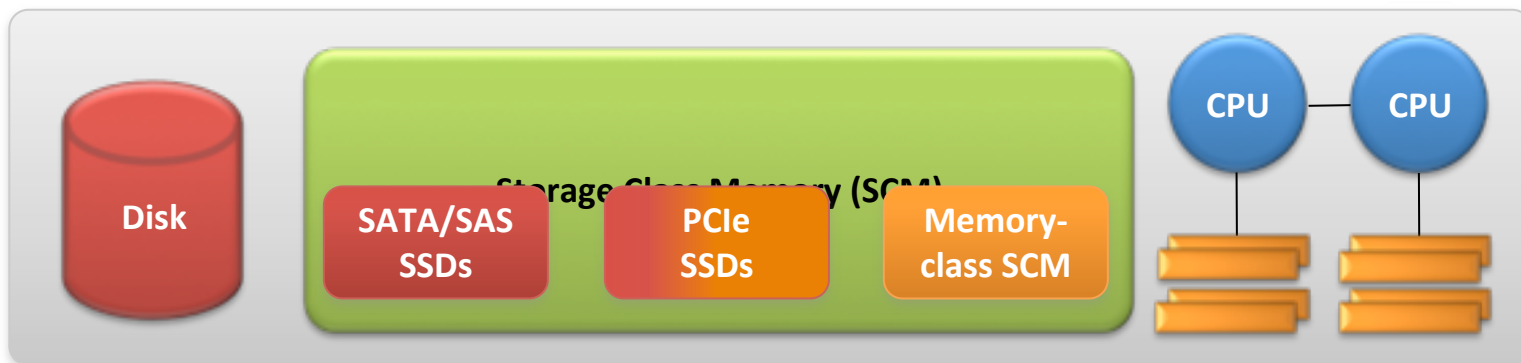


VIRIDENT

SNIA PCIe Roundtable Flash Memory Summit 2012

Storage Class Memory (SCM) in Datacenter

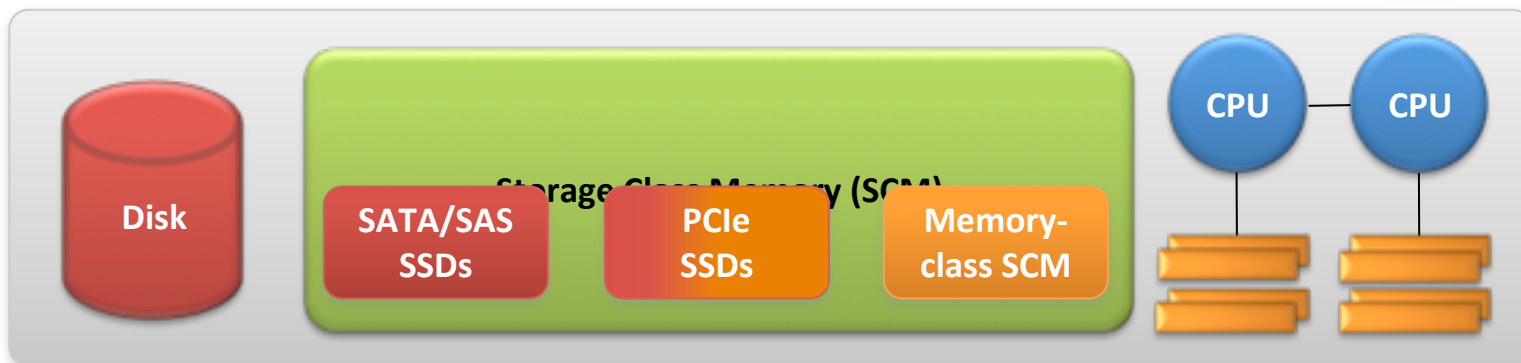
New Layer In Storage Hierarchy



PCIe SSDs established, Memory-class emerging

Storage Class Memory (SCM) in Datacenter

New Layer In Storage Hierarchy

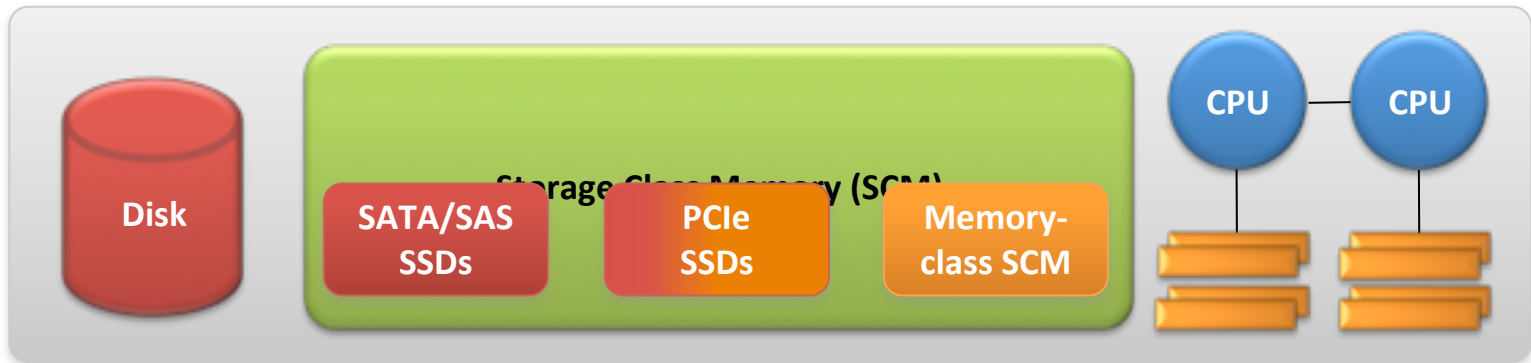


Attribute	Disk	SATA/SAS SSDs	PCIe SSD	Memory-class SCM	DRAM
Capacity (GB)	100's - 1000's	100's	100's - 1000's	100's-1000's	10's-100's
Read performance	10's ms, ~100 MB/s	100's us, ~100's MB/s	10's us, 1's GB/s	100's ns, 1's GB/s	~100 ns, 10's GB/s
Write performance	10's ms, ~100 MB/s	100's us, ~100 MB/s	~10's us, ~1's GB/s	1's us, 1's GB/s	

PCIe SSDs established, Memory-class emerging

Storage Class Memory (SCM) in Datacenter

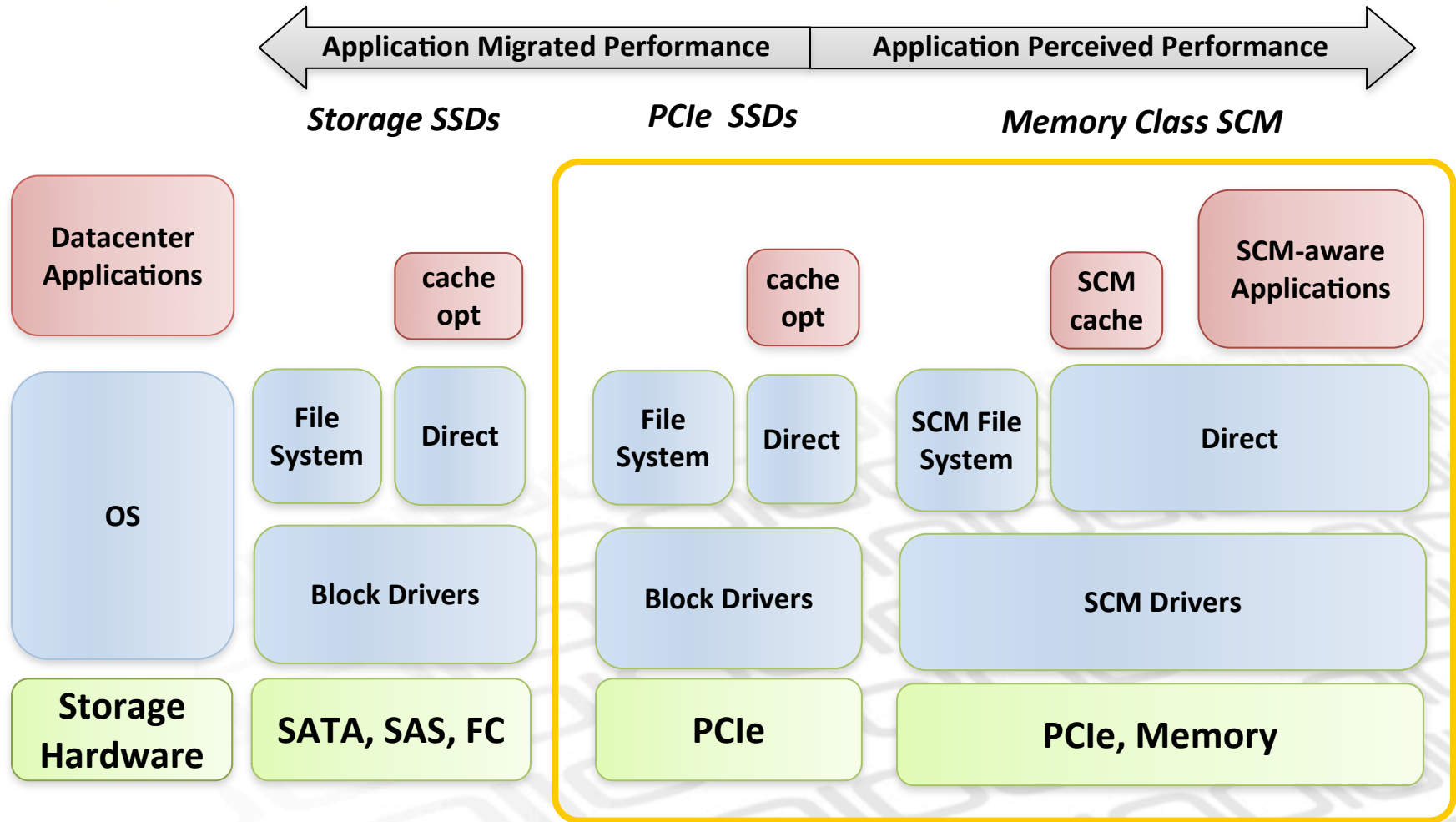
New Layer In Storage Hierarchy



Attribute	Disk	SATA/SAS SSDs	PCIe SSD	Memory-class SCM	DRAM
Capacity (GB)	100's - 1000's	100's	100's - 1000's	100's-1000's	10's-100's
Read performance	10's ms, ~100 MB/s	100's us, ~100's MB/s	10's us, 1's GB/s	100's ns, 1's GB/s	~100 ns, 10's GB/s
Write performance	10's ms, ~100 MB/s	100's us, ~100 MB/s	~10's us, ~1's GB/s	1's us, 1's GB/s	~100 ns, 10's GB/s

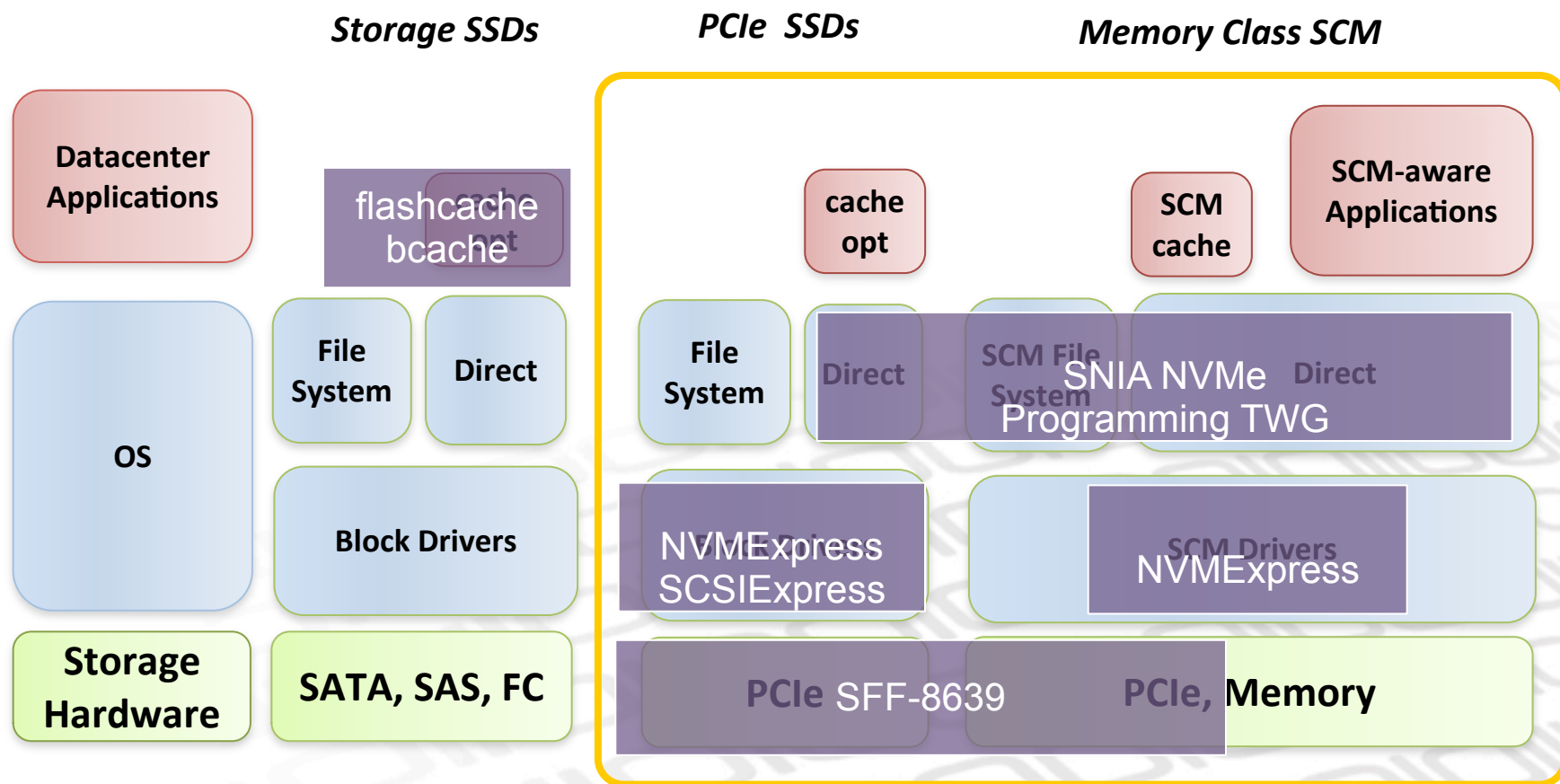
PCIe SSDs established, Memory-class emerging

Storage Class Memory (SCM) in Datacenter: From Block Applications to SCM Applications



Optimization required for applications to realize memory class benefit

Storage Class Memory (SCM) in Datacenter: An industry activity mapping...



Industry agreement for optimized architecture in place



VIRIDENT

Thank you!



Questions for Q & A

Questions:

- 1. Please explain the roles of NVM and SOP/PQI – will SOP/PQI play on NVM? What is the future of the inter relationships of these, and other, specs?**
- 2. What issues do you see needing to be solved before PCIe becomes a standard mass storage bus? Please comment on hot plugging, scatter gather, heat and performance and other issues.**
- 3. What are some of the key issues facing test & qualification of PCIe Storage?**
- 4. Please comment on the pros/cons of adding a Flash Translation Layer and its future impact to current system design architecture and performance.**
- 5. PCIe SSDs already produce prodigious throughput which is usually measured in GB/s, and IOPs on the order of hundreds of thousands. With the advent of PCIe Gen3, that may double. How does one get all that performance to move outside of the system that houses the SSD and into the system requesting the data?**



Thank You Very Much!