SAS Anchors Enterprise SSD Deployment

- Harry Mason - Director, Industry Marketing, LSI Corporation

 - Marty Czekalski - Interface & Emerging Architecture Program Manager, Seagate
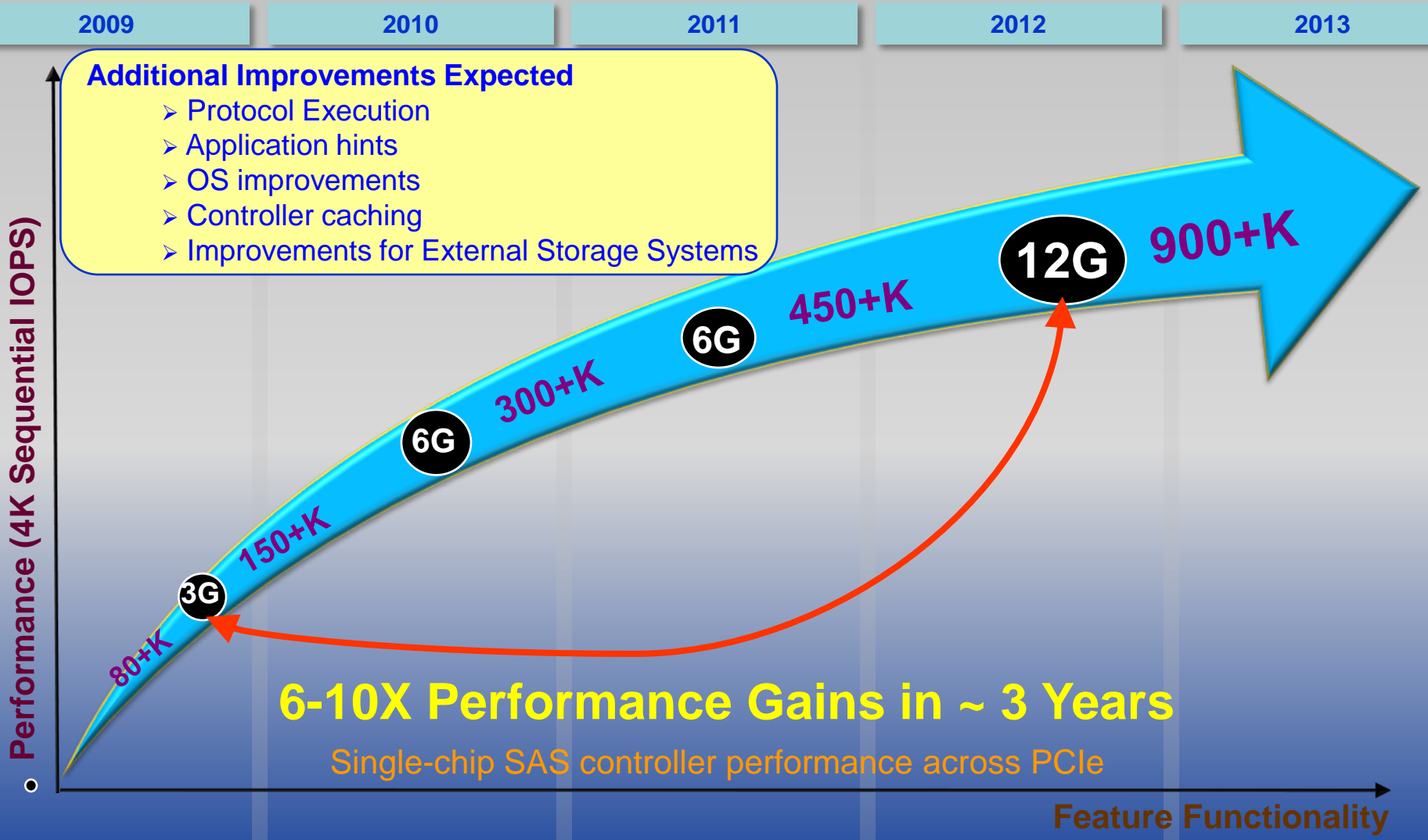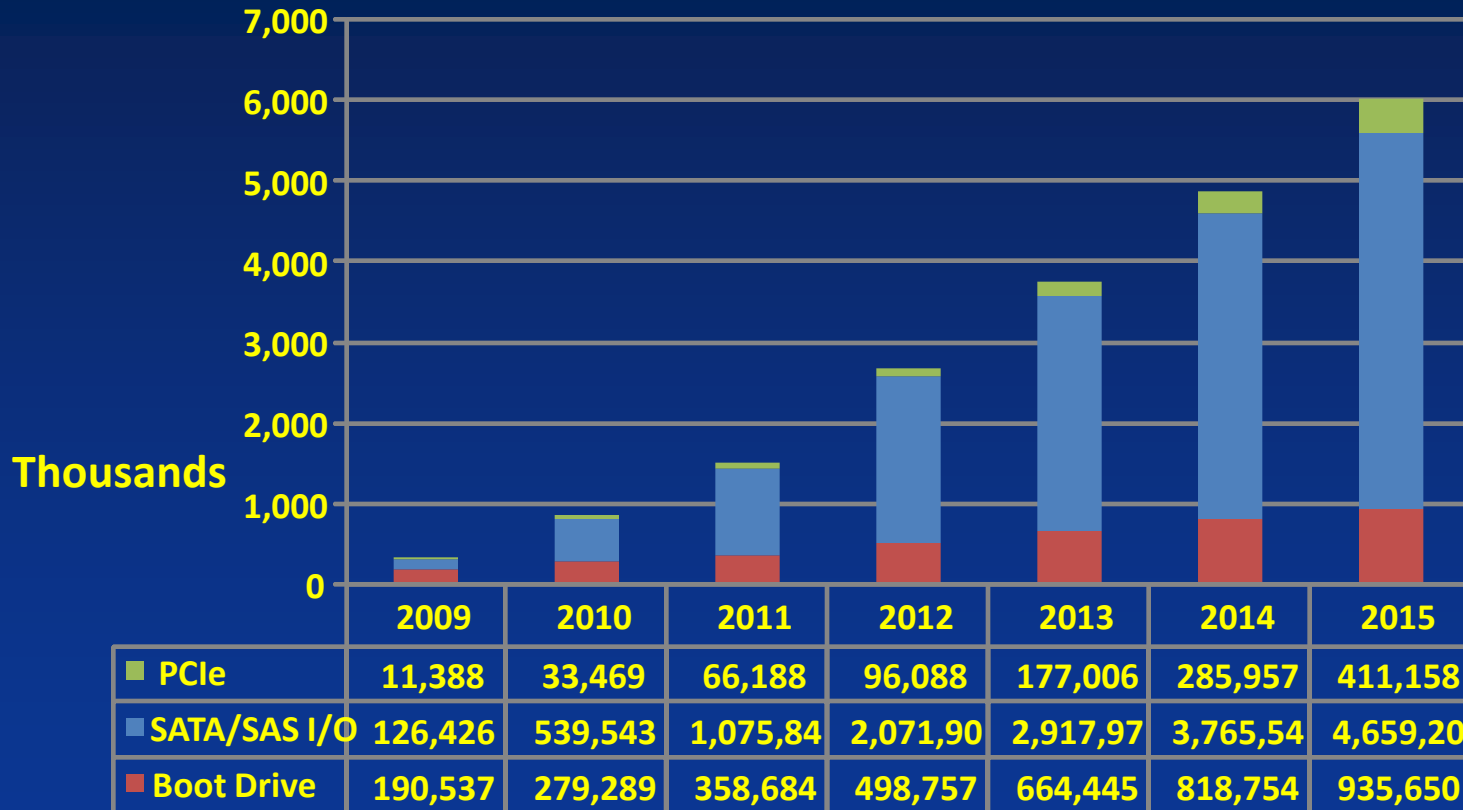
www.scsita.org

SAS is the Enterprise

- Preserves Storage Investment – Logical SCSI
- Broad <u>Open</u> Industry Standards Support
- Dynamic Platform for Storage Innovation
- Enterprise Proven – RAS
- Depth and Breath of Infrastructure
- Class Drivers Personalize non-HDD devices
- Operates across numerous transports
  - iSCSI, ATAPI, FC, Firewire, USB, PCIe
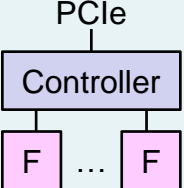
SAS Projections – Existing Standards

| 2009 | 2010 | 2011 | 2012 | 2013 |

Additional Improvements Expected
- ➢ Protocol Execution
- ➢ Application hints
- ➢ OS improvements
- ➢ Controller caching
- ➢ Improvements for External Storage Systems

Performance (4K Sequential IOPS)

900+K

12G

450+K

6G

300+K

6G

150+K

3G

80+K

6-10X Performance Gains in ~ 3 Years

Single-chip SAS controller performance across PCIe

Feature Functionality
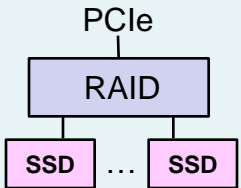
Server-attached SSD Forecast

Server-attached SSD Units

Thousands

	2009	2010	2011	2012	2013	2014	2015
■ PCIe	11,388	33,469	66,188	96,088	177,006	285,957	411,158
■ SATA/SAS I/O	126,426	539,543	1,075,84	2,071,90	2,917,97	3,765,54	4,659,20
■ Boot Drive	190,537	279,289	358,684	498,757	664,445	818,754	935,650
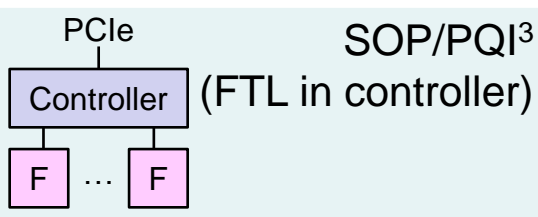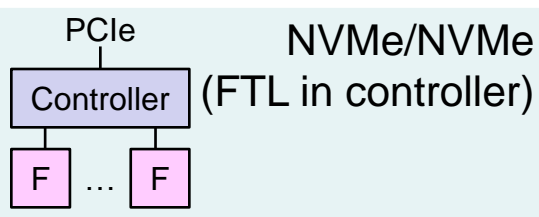
Source: Forward Insights – March 2011

4

	Native		Aggregator	
Commands/Transport	PCIe Controller F ... F	Proprietary (FTL[1] in host/ main memory)	PCIe RAID SSD ... SSD	SCSI or SATA (Multiple SSDs & controller on card)
Committee	None		None	
Standards Based	No		**Yes**	
Performance with Flash	High		High	
CPU Overhead	High		**Low**	
Latency with short queue	Very Low		Low	
Latency with deep queue	Moderate		**Low**	
Use Case Extensibility	No		**Yes** (RAID, HBA, etc)	
Maturity	Evolving		**Based on Proven Industry Architectures**	
Enterprise feature set (PI, Security, Mgmt, etc.)	No		Depends on implementation	

[1] FTL : Flash Translation Layer

Enterprise Interfaces: The Future of PCIe SSDs

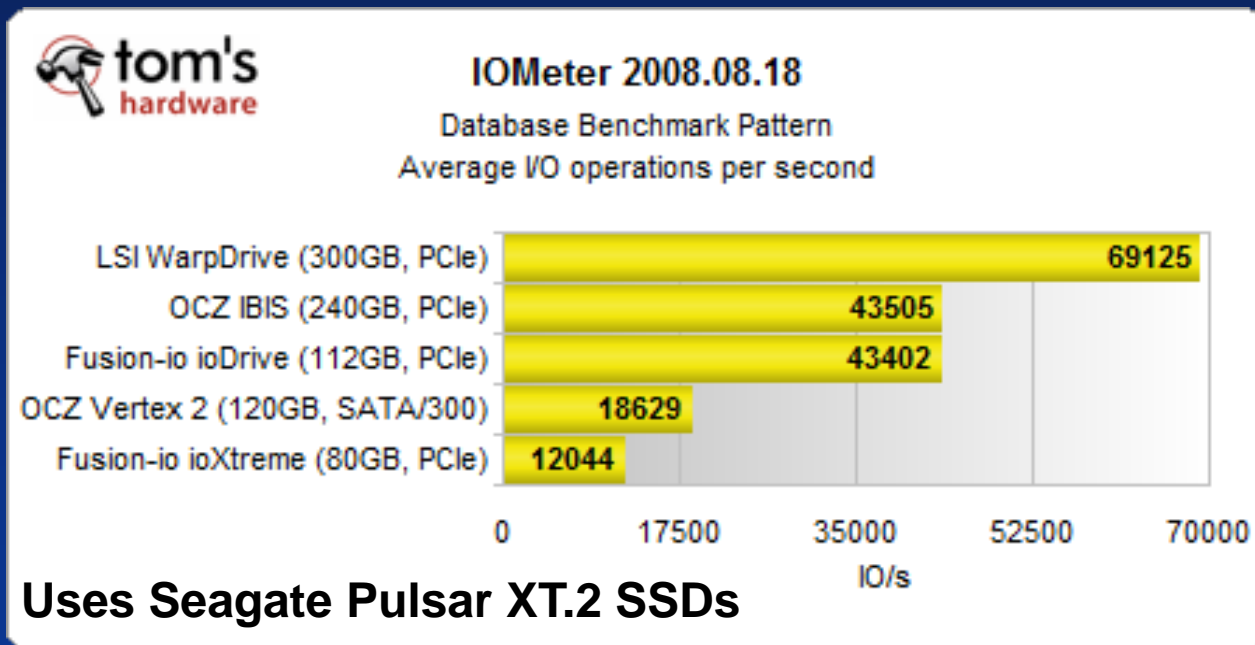|	SOP/PQI[1]	NVMe[2]
Commands/Transport [1]SOP : SCSI over PCI Express [2]NVMe : Non- Volatile Memory Express [3]PCIe Queuing Interface [4]INCITS : International Committee for Information Technology Standards	PCIe — Controller — F … F SOP/PQI[3] (FTL in controller)	PCIe — Controller — F … F NVMe/NVMe (FTL in controller)
Committee	**T10/INCITS[4]**	Industry Working Group
Standards Based	**Yes (ANSI/ISO)**	No
Performance with Flash	Very High	Very High
CPU Overhead	Low	Low
Latency with short queue	Very Low	Very Low
Latency with deep queue	Low	Low
Use Case Extensibility	**Yes** (RAID, HBA, etc.)	No (NVM only)
Maturity	**Investment Protection**	TBD
Enterprise feature set (PI, Security, Mgmt, etc.)	**Full Support**	Limited

www.scsita.org

Flash SSD Latency Facts

- Dominant contributor to SSD latency – Flash Devices
  - SLC access > 25us, MLC access > 50us, assuming no contention
  - Write/erase times are considerably longer, especially with enterprise versions, and getting longer with each flash generation

- Lager queue depths increase latency
  - Once a flash part begins access, other requests to that part must wait
  - Up to 8 flash die share bus access, causing each die to wait its turn

- Housekeeping activities add additional latency (address translation, garbage collection, wear leveling, etc.)

> **Protocols & interconnects have negligible effects on latency.**

www.scsita.org

"… the I/O capabilities of its LSISAS2008 controller, as it dominates the I/O test patterns for database, file server, Web server, and workstation scenarios."
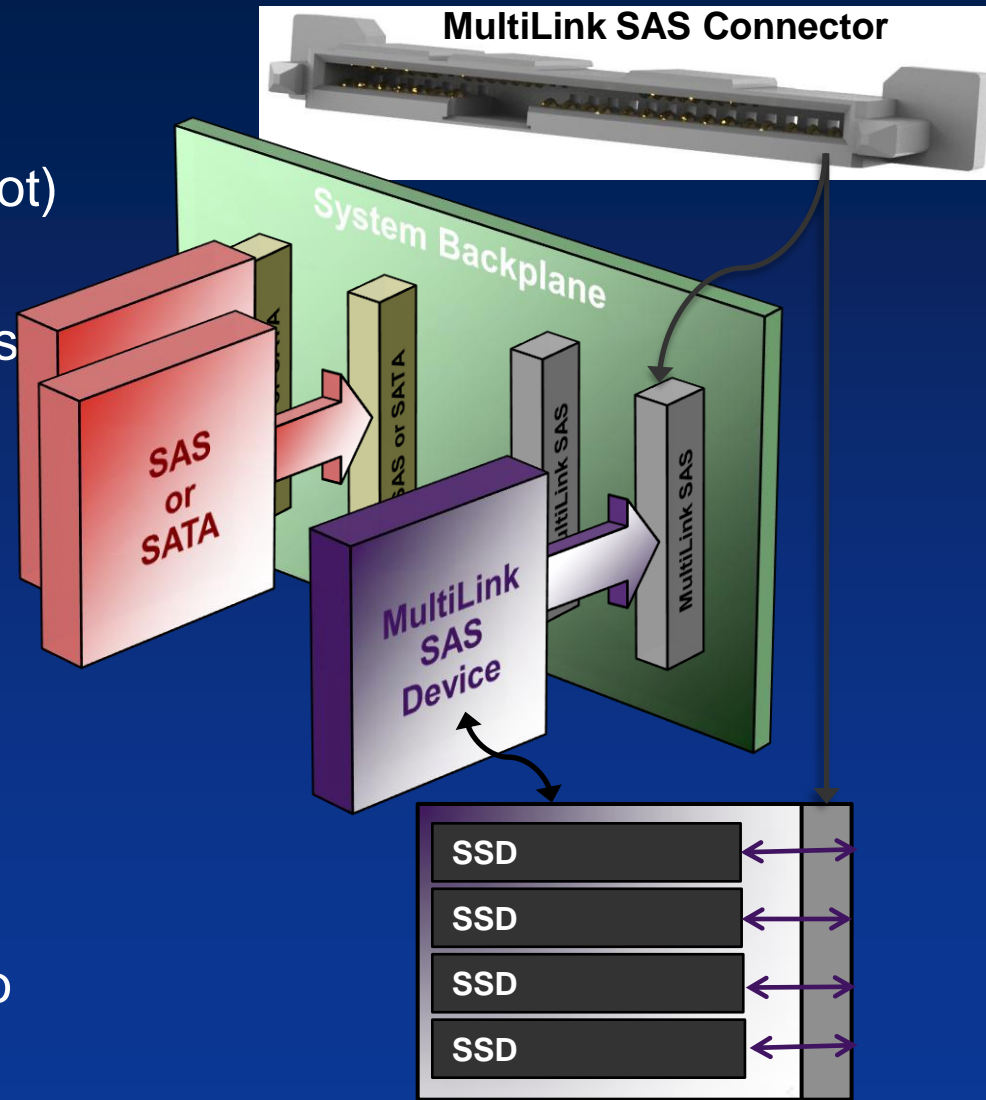

IOMeter 2008.08.18
Database Benchmark Pattern
Average I/O operations per second

	IO/s
LSI WarpDrive (300GB, PCIe)	69125
OCZ IBIS (240GB, PCIe)	43505
Fusion-io ioDrive (112GB, PCIe)	43402
OCZ Vertex 2 (120GB, SATA/300)	18629
Fusion-io ioXtreme (80GB, PCIe)	12044

Uses Seagate Pulsar XT.2 SSDs

Selected by ICC for National Center for Supercomputing Applications (NCSA) "Dark Energy" project
http://www.lsi.com/downloads/Public/Direct%20Assets/LSI/casestudy_ICC-NSCA.pdf
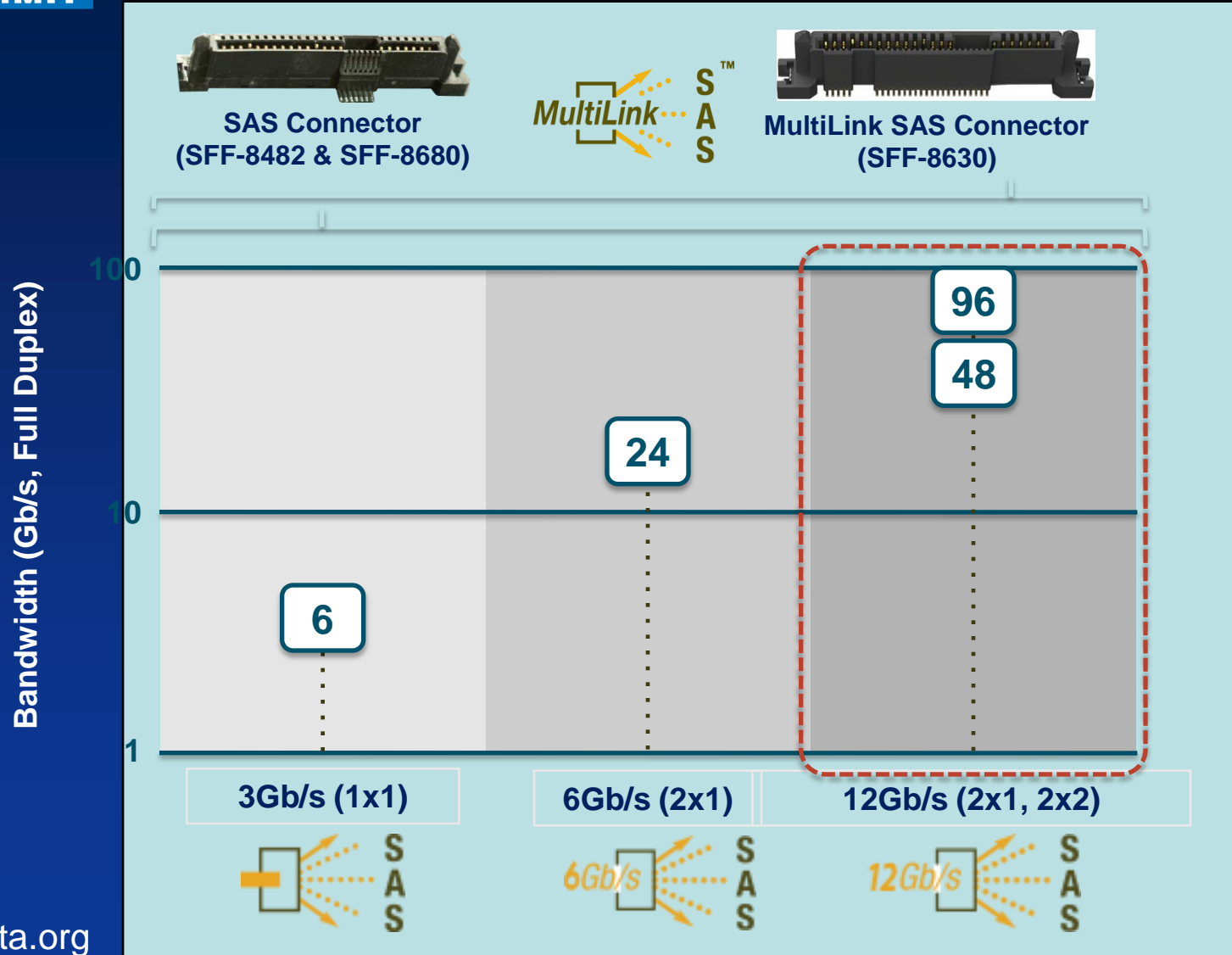
8

MultiLink SAS™ Slot

- ## MultiLink SAS™

  - High performance (20+W per slot)

  - Hot swap, serviceability

  - High availability (2 fault domains possible)

  - Low implementation risk:
    Standard SAS drivers
    - Fully hardened protocol stack
    - Common management stack

  - Low investment (repackaging)

  - Flexible: Independent SSDs or wide port SSDs

  - Able to isolate tier-tier traffic w/o accessing system memory

MultiLink SAS Connector

SAS Connector
(SFF-8482 & SFF-8680)

MultiLink SAS™

MultiLink SAS Connector
(SFF-8630)

Bandwidth (Gb/s, Full Duplex)

100

96

48

24

10

6

1

3Gb/s (1x1)

6Gb/s (2x1)

12Gb/s (2x1, 2x2)

SAS

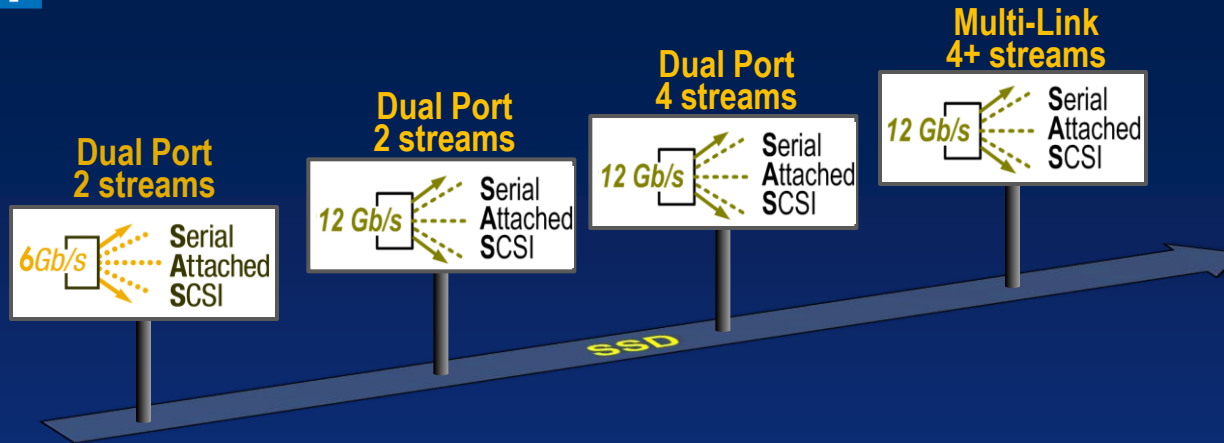6Gb/s SAS

12Gb/s SAS

www.scsita.org

10

Innate SAS Advantages

- Persistent & non-persistent usage models
- Concurrency – Performs best with large queues and high work loads
- Low CPU Utilization – Supports Multi-tenant Cloud service platforms
- Flexible Deployment Model – Cabled, PCIe, SAS slots, Motherboard, etc.
- Solid legacy, rich infrastructure, & viable roadmap
- "No compromise" performance

Innovating for Bandwidth Growth

- **Today: Leverage 6Gb/s SAS infrastructure for rapid deployment**

- **Tomorrow: Enhance the SAS to support SSD unique requirements:**

  - **SAS Roadmap Progression:**

    - **A) 6Gb/s SAS dual port, 2 streams**
    - **B) 12Gb/s SAS dual port, 2 streams**
    - **C) 12Gb/s SAS dual port, 4 streams**
    - **D): 12Gb/s SAS multi-link, 4+ streams**

  - **SSD unique commands: Trim, Unmap, others…**

Benefits of SAS & MultiLink SAStm Enhancements	
Multiple Links (BW)	X4 (4x600MB/s)
Power Available	25W (2.5")
Total Latency	>26 us
Multi host protocol	**Yes**
High availability	**Yes (Dual Port)**
Scalability	**Excellent**
Robust proven protocol stack	**Yes**
Hot Swap serviceable	**Yes**
Compatible with existing management SW	**Yes**