Enterprise SSDs with Unrivaled Performance

A Case for PCIe® SSDs

Kam Eshghi
Sr. Director of Marketing
Integrated Device Technology, Inc.

Enterprise Server & Storage Trends

- CPU performance growth outpacing storage performance growth
- Multi-core CPUs and virtualization are increasing randomness of disk I/Os
- Key metrics: $$/IOPS and IOPS/W
- OEMs will use SSDs for performance and HDDs for capacity
- Flash filling the price/performance gap between DRAM and HDD

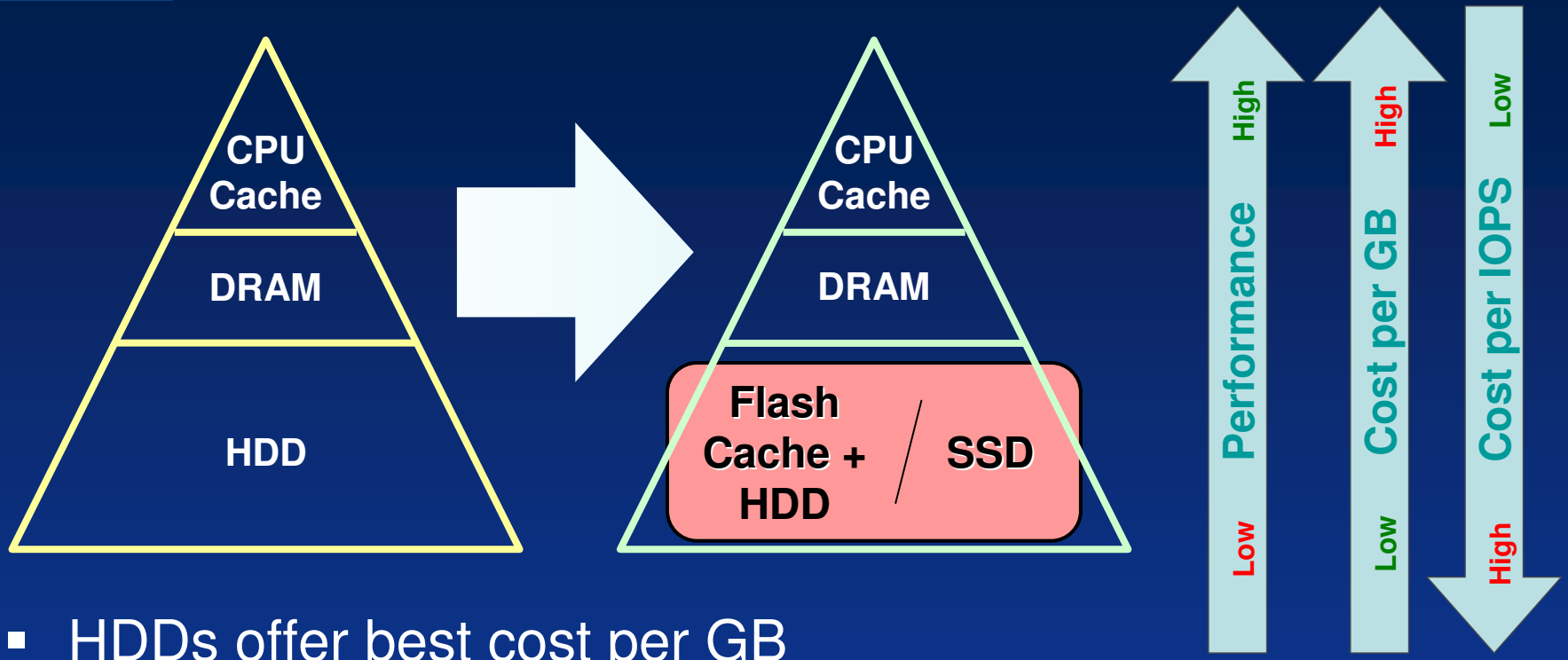
#1 driver for Flash adoption in Enterprise is Performance

End User Applications

- Improve OLTP performance
  - Higher percentage of IO operations
- Accelerate real-time financial data processing
  - Every µsec matters $$
- Reduce 3D rendering time
  - Enable real time rendering for CAD/CAM
- Accelerate Database/Data mining performance
  - Faster data mining in customer relationship management applications

Business Critical Applications Limited by Storage Performance

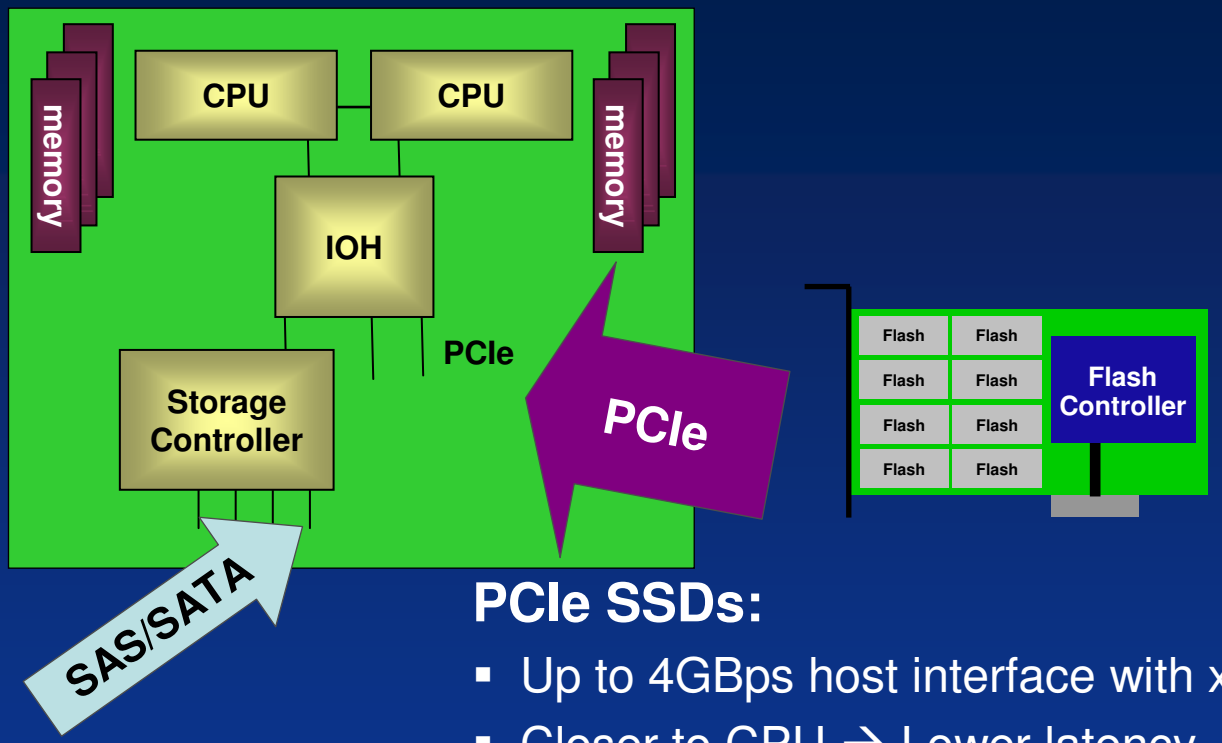
Enterprise Memory/Storage Hierarchy Paradigm Shift



- HDDs offer best cost per GB
- SSDs improve cost per IOPS by >10x vs. HDDs
- New Flash-based layer for application-managed caching

PCIe SSDs' high performance and low latency is ideal for Cache Tier and high performance drive replacement

Server SSD Attach Points



SAS/SATA SSDs:

- 600MBps host interface
- Storage Controller adds latency

PCIe SSDs:

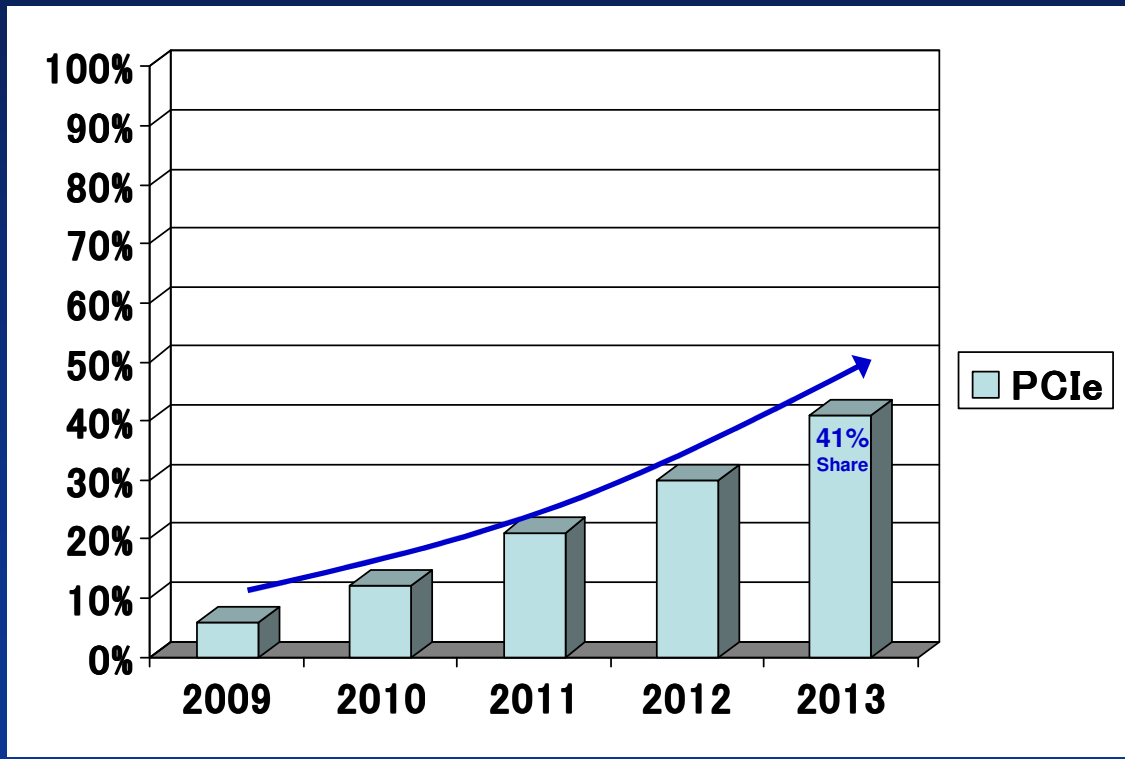
- Up to 4GBps host interface with x8G2
- Closer to CPU → Lower latency
- Plenty of native PCIe connectivity

PCIe SSDs have >5x the throughput of SAS/SATA SSDs

PCIe-based Enterprise SSD Market

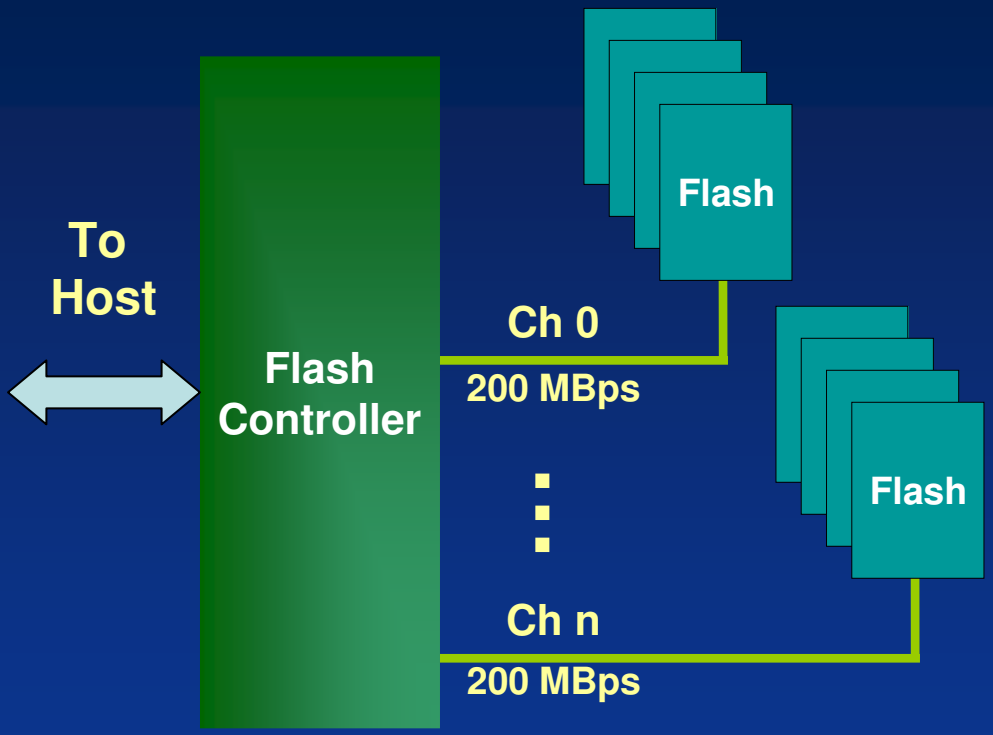
- Enterprise-class SSDs expected to grow to over 5M units in 2013
- Mostly FC or SATA-based today, with SAS and PCIe expecting to dominate the enterprise SSD market by 2013
- 41% of enterprise SSDs expected to use PCIe host interface in 2013
- A standard driver and consistent feature set will expedite PCIe SSD adoption

PCIe Enterprise SSDs



Source: Gartner 2010

Enterprise SSDs Saturating Host Interface

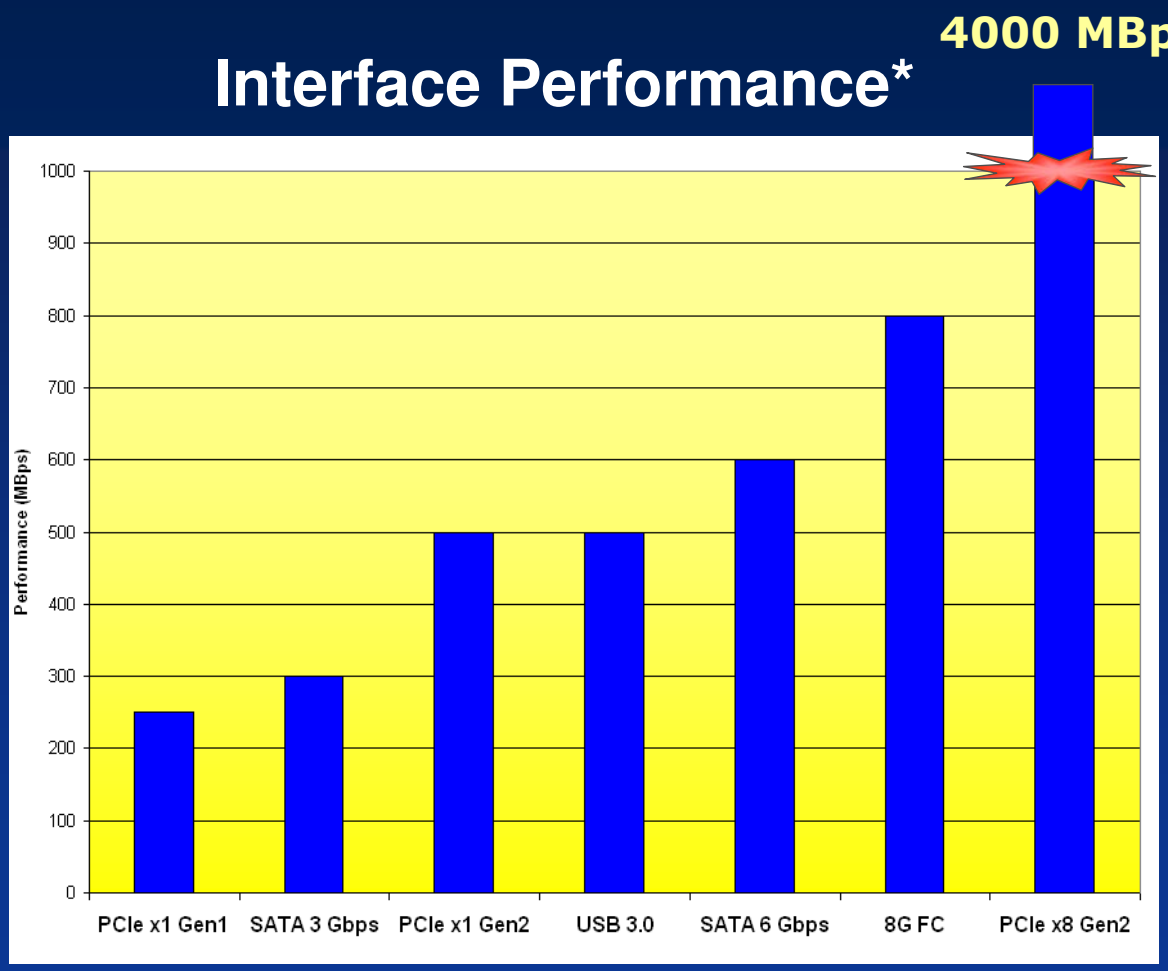


- Flash more efficient and higher performance
- Page size increasing from 512B to 8KB
- Channel performance increasing from 40MT/s to 133-200MT/s (DDR)
- 4 flash channels with 8 targets on each will saturate SATA 6G host interface

Performance bottleneck shifting to Host Interface

What is the right Host Interface?

Interface Performance*



- PCIe improves overall system performance by reducing latency and increasing throughput
- CPU complex provides plenty of PCIe ports, simplifying server integration

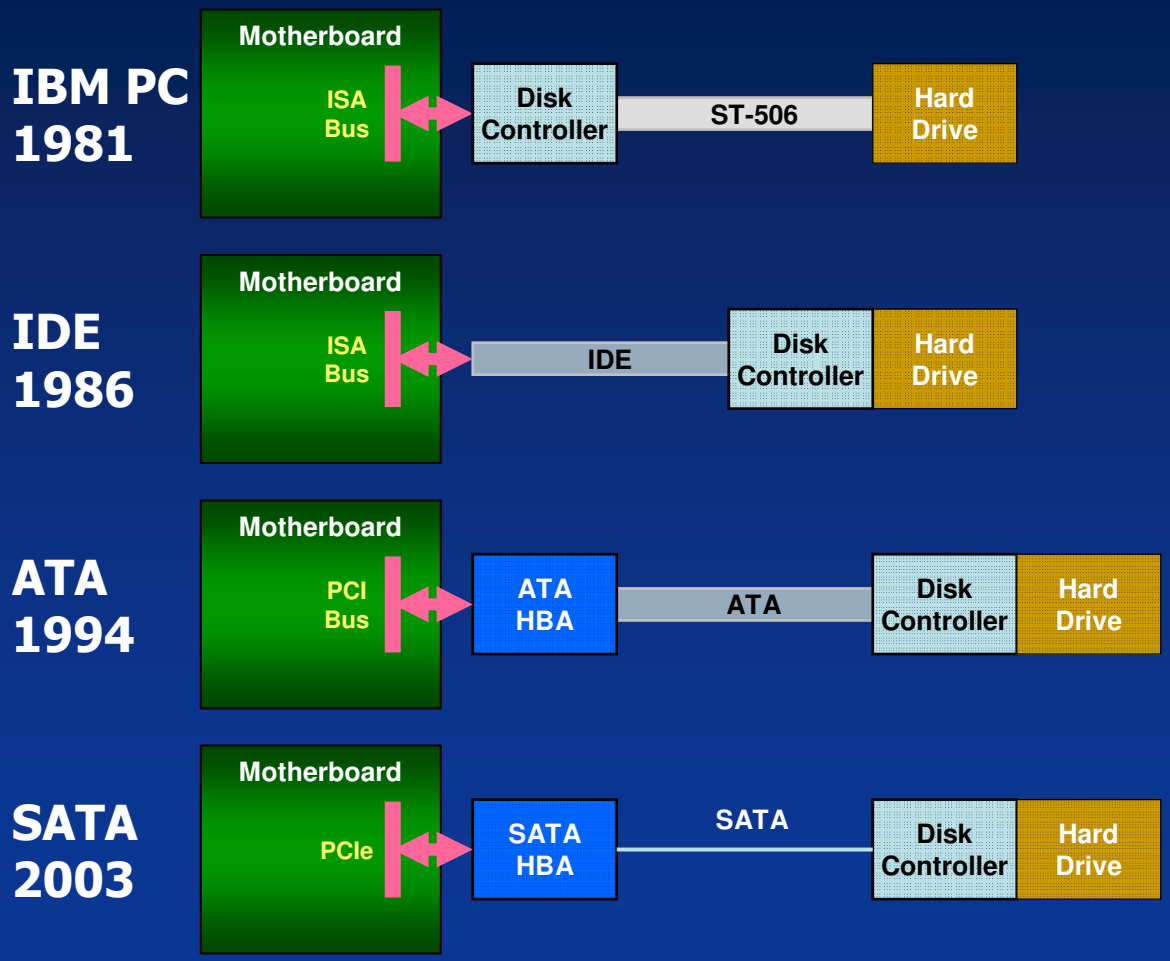
* **All performance numbers assume zero protocol overhead**

Why PCIe SSD? Performance.

Storage Type	Performance	
	Latency	Max. Throughput
HDD	2000µs	130 MBps
SATA 3G SSD	50µs + Protocol Controller	280 MBps
SATA 6G SSD	50µs + Protocol Controller	400 MBps
PCIe x8 G2 SSD	**50µs**	**3000 MBps**

PCIe SSDs offer lower latency and >5x the performance of SAS/SATA SSDs

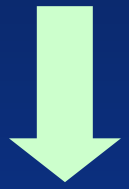
Myth: PCIe is an exotic storage interface (SATA history)

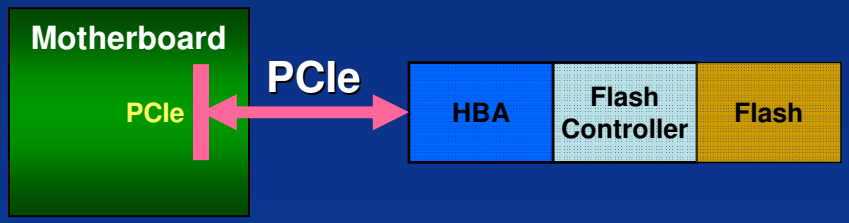


Disk controller attached directly to the Host Bus

Keep old HDD architecture, improve interface

Myth: PCIe is an exotic storage interface





Transition from HDD to SSD creates opportunity to reconsider host interface

Running host bus to disk is not a radical idea; it was done two decades ago!

Benefits of PCIe as an SSD Interconnect (Summary)

- Higher performance
  - >5x the throughput of SAS/SATA SSDs
- Lower latency
  - Closer to CPU
- Lower power
  - Eliminates SAS/SATA SerDes power
- Lower cost
  - Eliminates Storage HBA
- Less board real estate
  - More room for Flash devices

Some say… *Any "SSD" needs to be drop-in disk replacement, using existing software infrastructure*

We say… *It's time to bypasses legacy HDD technology for a direct host connection → get* ***blistering Performance at low Latency!***

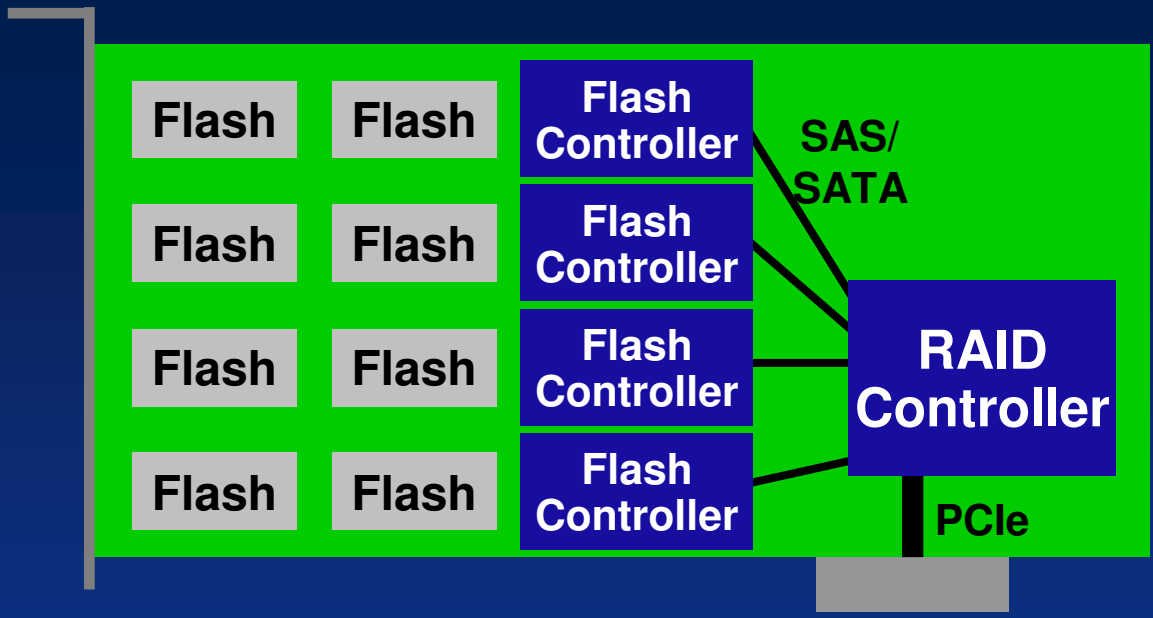
PCIe SSD Implementation Options

How would you build a PCIe SSD card?

Let's review a few different approaches…

- RAID-based using SATA/SAS Flash Controllers
- Native PCIe Flash Controller with Host Managed Flash
- Native PCIe Flash Controller on SSD

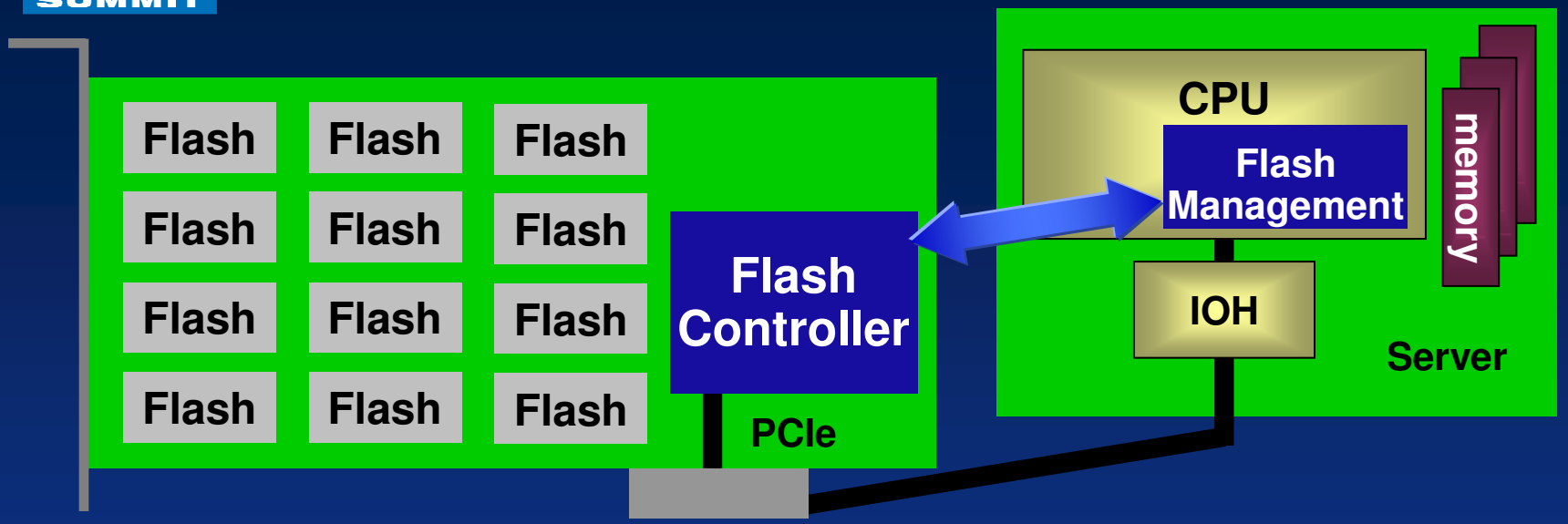
Option A – *RAID-based PCIe SSD*



- **Pro's**
  - Time to Market – All building blocks exist today
- **Con's**
  - Two-level architecture increases latency and reduces performance
  - High device count results in higher cost, uses up more board space
  - Higher power because of additional SerDes links and high device count

RAID-based PCIe SSD is readily available today, but not optimized for performance/power

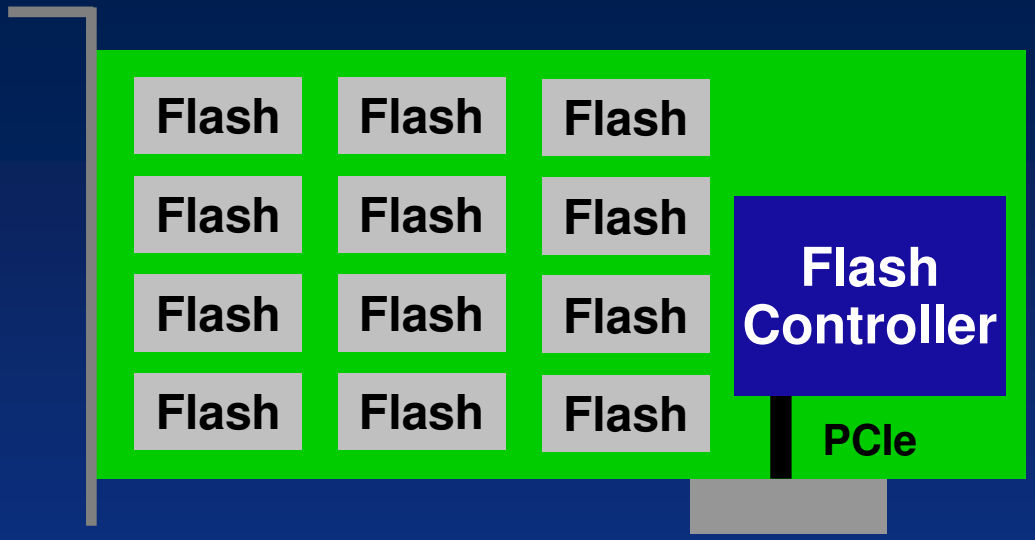
Option B – *PCIe SSD with Host Managed Flash*



- **Pro's**
  - Requires simple Flash controller
- **Con's**
  - Consumes host CPU processing resources
  - Consumes host memory
  - Requires proprietary divers, adding software complexity and extending OEM qualification

Running flash management algorithms on host drains Host CPU/RAM resources

Preferred Option
– *Native PCIe Flash Controller on SSD*



- **Pro's**
  - Improves performance
  - Reduces cost
  - Reduces power
  - Enables standard OS drivers
  - Reduces host CPU/RAM utilization
  - Frees up board real estate for Flash devices
- **Con's**
  - Requires sophisticated Flash controller

Native PCIe Flash Controller improves performance, while reducing cost & complexity

PCIe SSD Software Challenge

What about the software?

- Currently there is no standard OS/driver infrastructure in place for PCIe SSDs
  - No standard host controller register interface
  - Each vendor has to provide drivers for every OS
  - OEMs have to validate each vendor's drivers
- AHCI makes PCIe SSD look like a SATA SSD, but need to extend with proprietary drivers to optimize performance
  - No standard drivers, validation costs, etc.

Solution: **Proprietary drivers today, migrating to standard "Enterprise NVMHCI" in the future**

What is Enterprise NVMHCI?

Enterprise Non-Volatile Memory Host Controller Interface

- Standard host controller interface for Enterprise class PCIe SSDs
- A standard register programming interface for NVM
- Not tied to any specific NVM technology
- All NAND management abstracted out
- Optimized for both cache and SSD usage models
- Driven by Intel, Dell, Microsoft, IDT, and 50+ other companies

Enterprise NVMHCI Objectives

Addressing Enterprise PCIe SSD requirements

- Address Enterprise server scenarios
- SSD vendors focus on building great SSDs
- OS vendors deliver drivers for all PCIe SSDs
- Simplifies OEM qualification
- Features are implemented in a consistent fashion, reducing time to market for PCIe SSDs

Enterprise NVMHCI Architectural Goals

- Increase parallelism, eliminate performance bottlenecks seen in other interfaces
  - Support multiple deep command queues
  - Simplify command decoding and processing
- Provide an efficient & streamlined command set enabling very high IOPS
  - Do not carry forward HDD command set legacy
  - Commands optimized for NVM
- Provides Enterprise features
  - End-to-end data protection
  - Firmware update
  - Encryption
  - Comprehensive statistics
  - Health status reporting
  - Robust error reporting & handling

Enterprise NVMHCI Timeline

- Timeline
  - Aug 2010: 0.7 revision
  - Oct 2010: 0.9 revision (erratum only after this point)
  - Dec 2010: 1.0 release, published
- To get involved in the specification definition, join the NVMHCI Workgroup

IDT Demonstrating Industry's First Enterprise NVMHCI Flash Controller prototype NOW!

Summary

- Need to increase storage I/O throughput and reduce latency for Enterprise applications
- PCIe SSDs deliver the best $$/IOPS and IOPS/W
- Software interface expected to migrate from today's proprietary drivers to standard Enterprise NVMHCI
- Enterprise NVMHCI fosters faster adoption for PCIe SSDs, with standard OS drivers and reduced OEM validation