SNIA

Education

The Benefits of Solid State in Enterprise Storage Systems

David Dale, NetApp

SNIA Legal Notice

- The material contained in this tutorial is copyrighted by the SNIA unless otherwise noted.
- Member companies and individual members may use this material in presentations and literature under the following conditions:
  - Any slide or slides used must be reproduced in their entirety without modification
  - The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.
- This presentation is a project of the SNIA Education Committee.
- Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.
- The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

  NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.

© 2011 Storage Networking Industry Association. All Rights Reserved.

Abstract

- Solid State in Enterprise Storage Systems
  - Targeted primarily at an IT audience, this session presents a brief overview of the solid state technologies which are being integrated into Enterprise Storage Systems today, including technologies, benefits, and price/performance.
  - It then goes on to describe where they fit into typical Enterprise Storage architectures today, with descriptions of specific use cases.
  - Finally the presentation speculates briefly on what the future will bring.

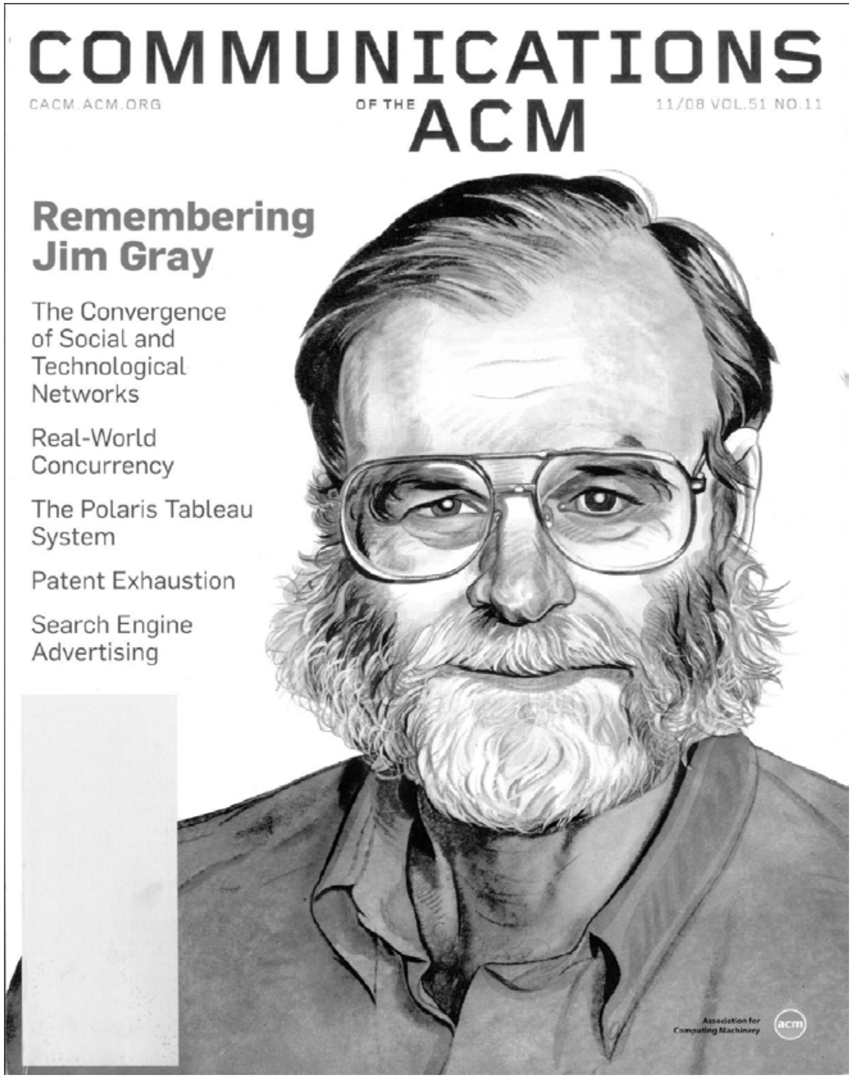
© 2011 Storage Networking Industry Association. All Rights Reserved.

Agenda

- Why flash in the datacenter? Why now?
- Memory, cache and storage
- Flash in enterprise storage today
  - SSD storage tier
  - Storage controller-based cache
  - Host-based flash
  - Network cache
- What's next
- Conclusion

© 2011 Storage Networking Industry Association. All Rights Reserved.

Remembering Jim Gray



Database and systems design pioneer, and co-creator of the Five Minute Rule (1987)

"Flash is a better disk …, and disk is a better tape" ~2006

Lost at sea January 2007

© 2011 Storage Networking Industry Association. All Rights Reserved.

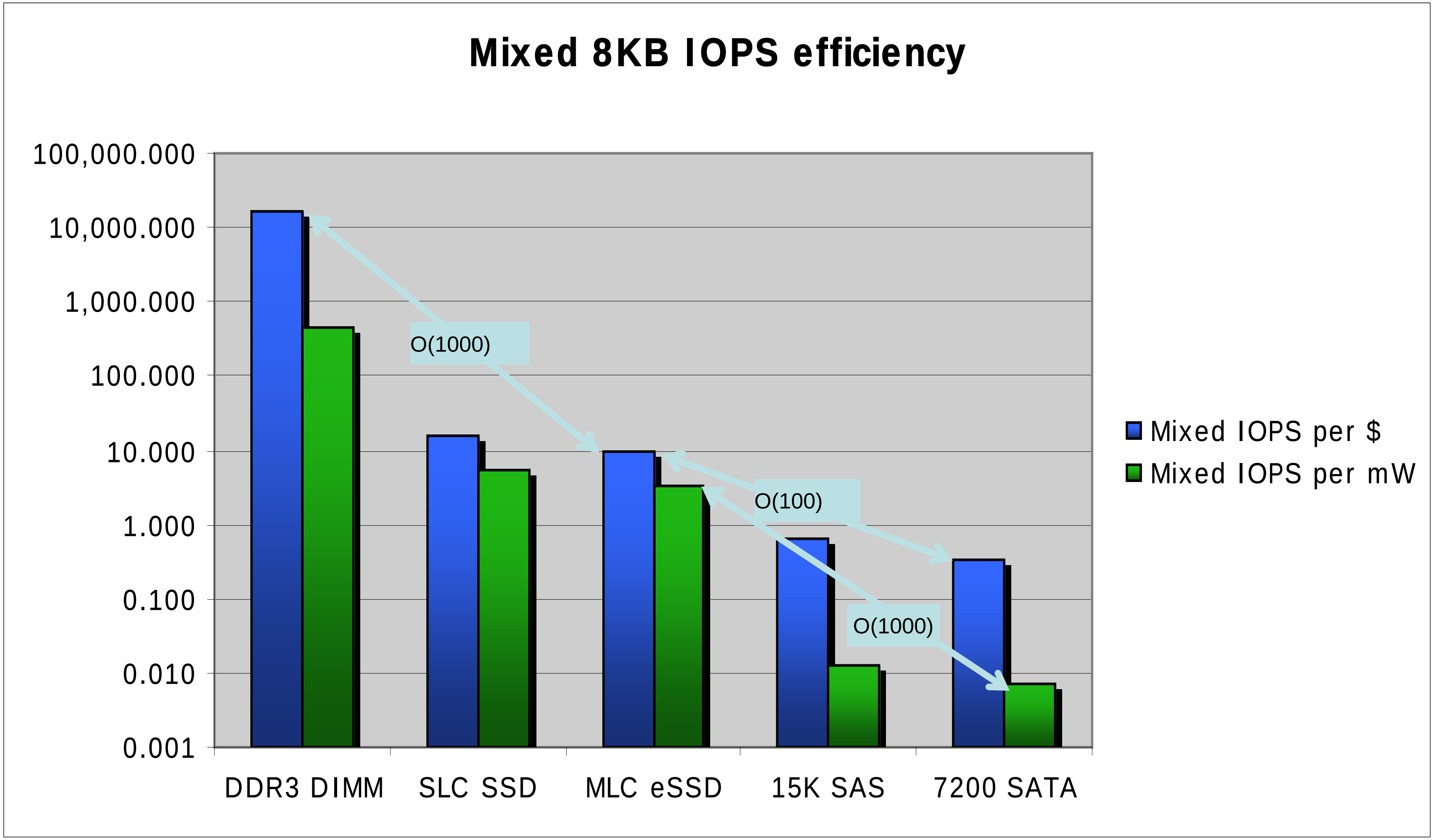
Why Flash in the Data Center Now?

- Why flash?
  - Capacity efficiency versus DRAM
    - ~5x better $ per GB
    - ~40x better power per GB
  - IOPS efficiency versus HDDs
    - ~40x better $ per IOPS
    - ~600x better power per IOPS
- Why now?
  - Period of rapid density advancements led to HDD-like bit density at lower $/GB than DRAM
  - Innovations in SSD and tiering technology

© 2011 Storage Networking Industry Association. All Rights Reserved.

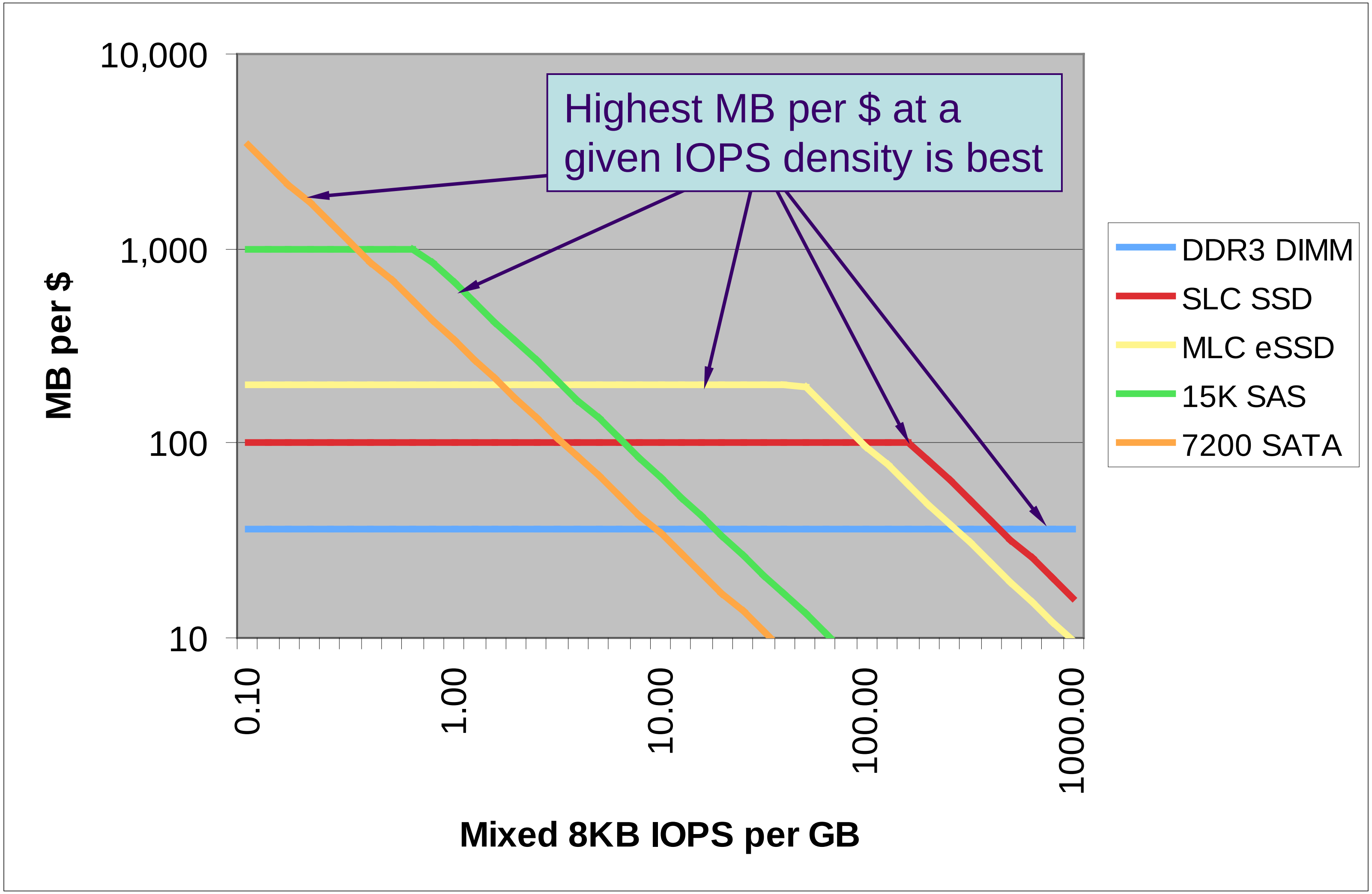
6

Why Flash? IOPS Efficiency



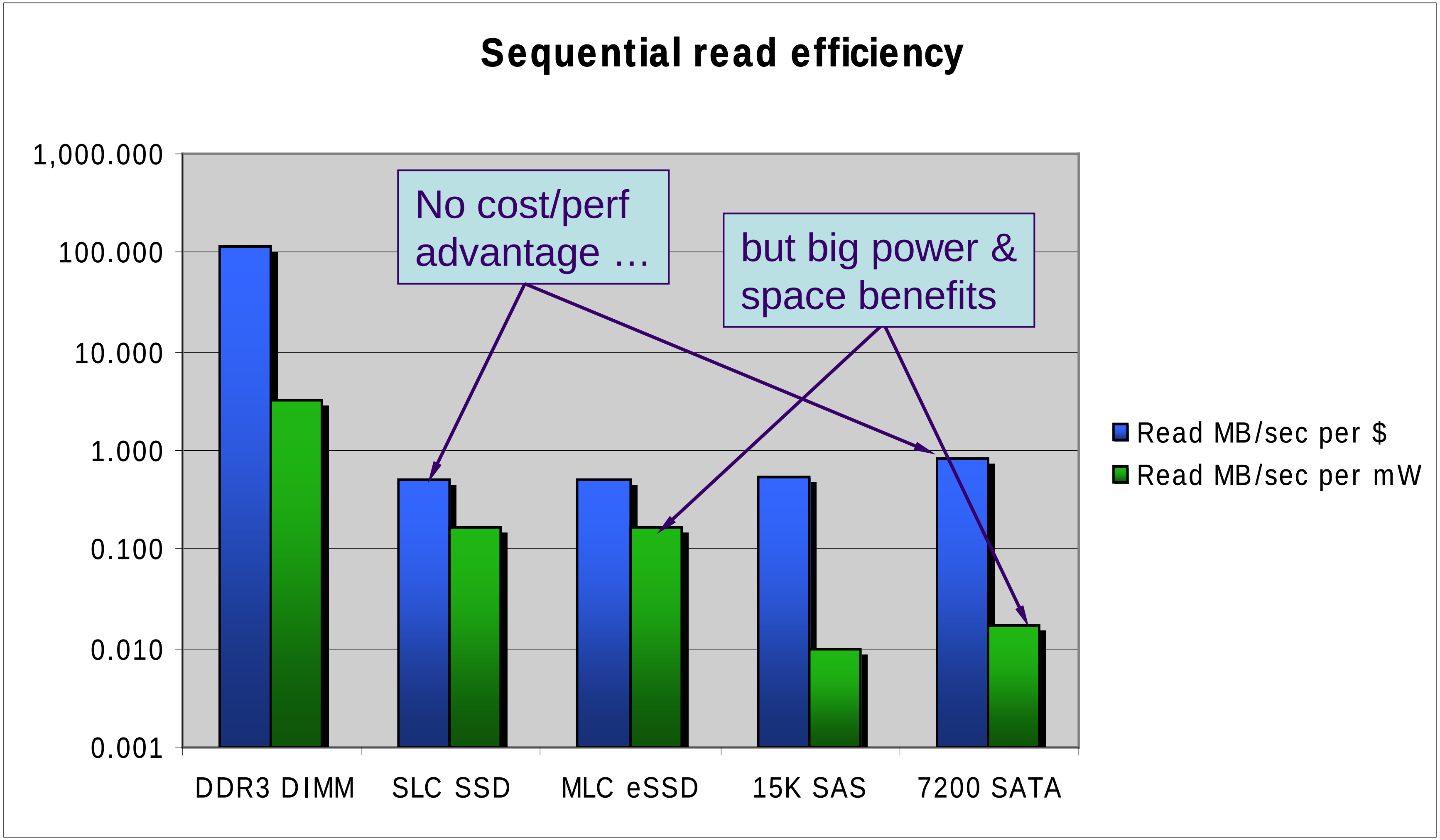
© 2011 Storage Networking Industry Association. All Rights Reserved.

Why Flash? An IOPS Density View



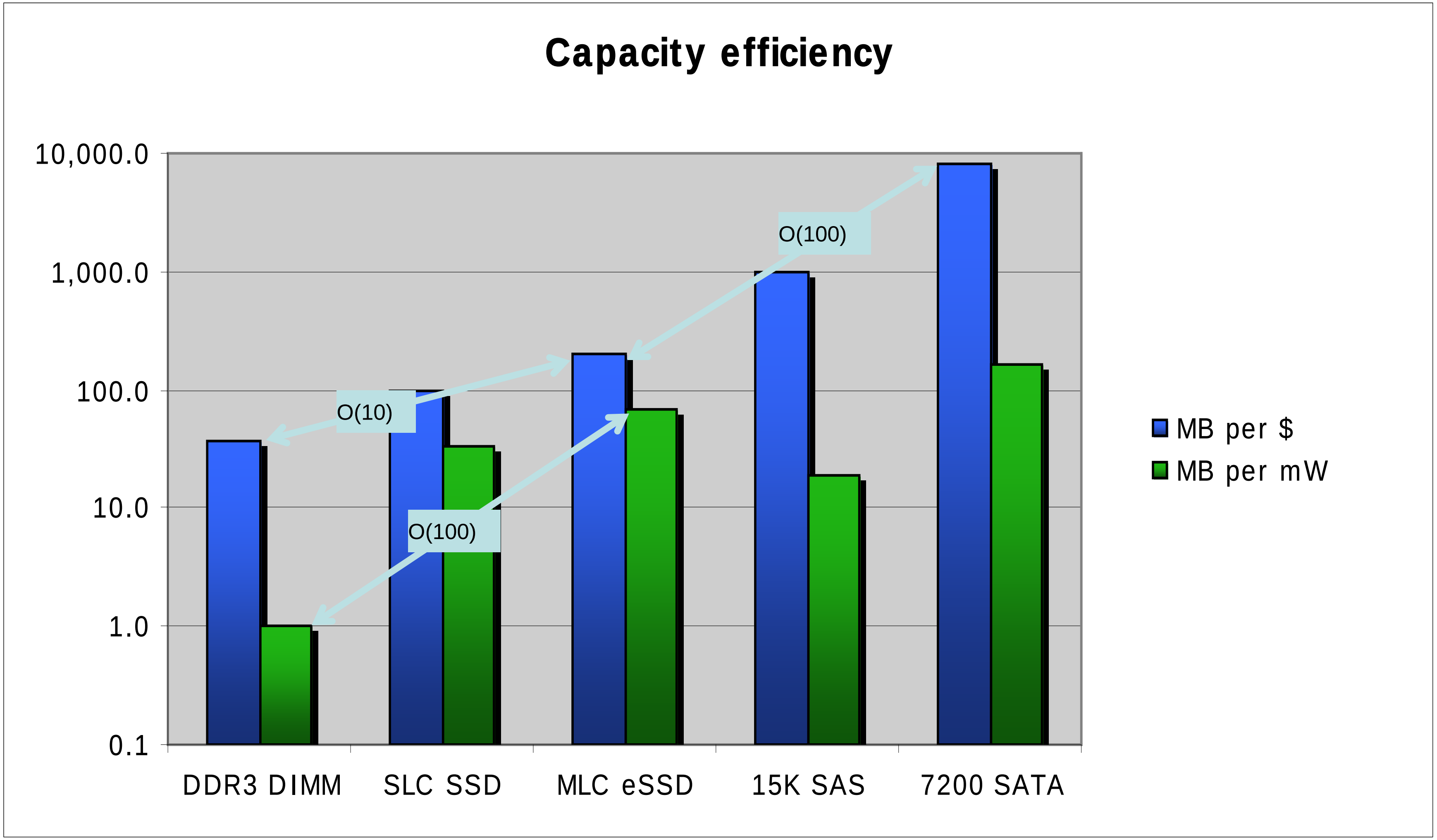
© 2011 Storage Networking Industry Association. All Rights Reserved.

Why Flash? Read Throughput per Watt



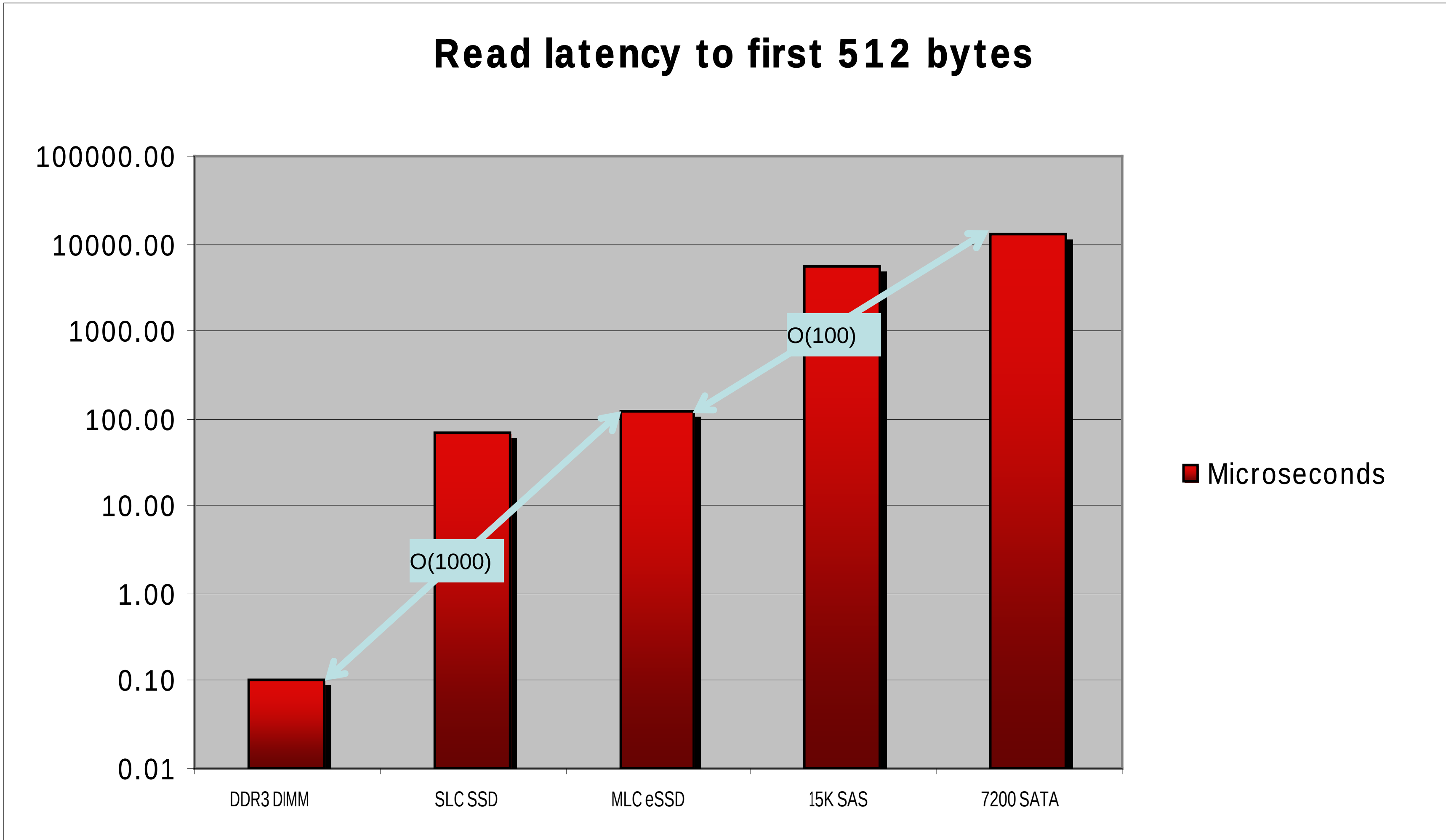
© 2011 Storage Networking Industry Association. All Rights Reserved.

Why Flash? Capacity Efficiency



© 2011 Storage Networking Industry Association. All Rights Reserved.

Why Flash? Read Latency



© 2011 Storage Networking Industry Association. All Rights Reserved.

Five Minute Rule, 1987

- Assuming that the cost of a cache is dominated by its capacity, and the cost of a backing store is dominated by its access cost (cost per IOPS), then the breakeven interval for accessing a page of data in cache is given by:

  Break-Even-Interval =
  Backing-Store-Cost-Per-IOPS /
  Cache-Cost-Per-Page

- 1987: Disk $2,000 / IOPS; RAM $5 / KB → 1 KB breakeven = 400 seconds ~= 5 minutes

© 2011 Storage Networking Industry Association. All Rights Reserved.

Five Minute Rule, 2010: DRAM & HDD

Education SNIA

- Disk $1 / IOPS (2,000x reduction)
- DRAM $25 / GB (200,000x reduction)

➔ 100 KB breakeven ~= 5 minutes
➔ 8 KB breakeven ~= 1 *hour*
➔ 1 KB breakeven ~= 10 hours *as Gray predicted*

- 200,000x / 2,000x = 100-fold decrease in breakeven access rate for a DRAM cache page backed by disk ➔ much bigger DRAM caches

© 2011 Storage Networking Industry Association. All Rights Reserved.

13

Five Minute Rule, 2010: Flash & HDD

Education SNIA

- Disk $1 / IOPS

- MLC eSSD ~$5 / GB

→ SSD 100 KB breakeven ~= 30 minutes
→ SSD 8 KB breakeven ~= 7 hours (5x DRAM)

Flash economically caches working sets with 5x longer access intervals than DRAM.

© 2011 Storage Networking Industry Association. All Rights Reserved.

Five Minute Rule, 2010: DRAM & Flash

- MLC eSSD ~$0.10 / mixed 8 KB IOPS
- DRAM $25 / GB

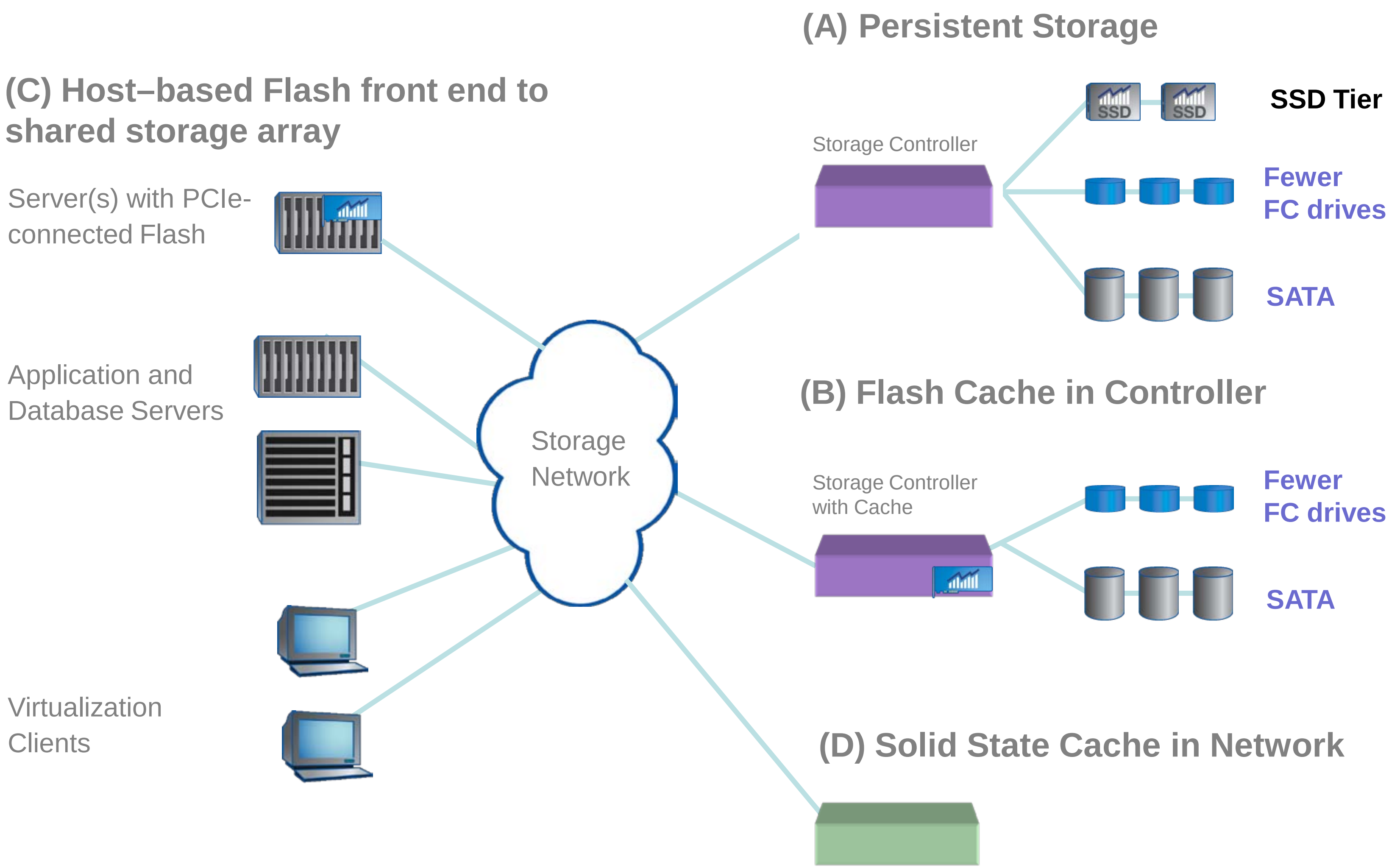
→ 8 KB breakeven ~= 8 minutes (1/10th DRAM)

Adding flash between DRAM and HDD reduces the breakeven access interval for DRAM by 10x, indicating that DRAM capacity could be reduced to hold working sets for data accessed 1/10th as often

© 2011 Storage Networking Industry Association. All Rights Reserved.

15

Storage Networking with Flash



© 2011 Storage Networking Industry Association. All Rights Reserved.

Available Solutions Compared

	Pros	Cons
Solid State Drives	❒ Assured performance levels ❒ Low cost per IOPS ❒ Admin has direct control over data stored in SSD tier	❒ High cost per gigabyte ❒ Requires partitioning of hot data ❒ Limited practical applications at current SSD costs
Controller Cache	❒ Hot data automatically flows into cache – no admin required ❒ Deployment non-disruptive ❒ Viable for common enterprise applications	❒ Cache must be populated before it becomes effective ❒ Less predictable performance than static placement
Host-based Cache/tier	❒ IOPS tier of storage on host; capacity tier on shared storage ❒ Flash on host will become common ❒ Cache and AST implementations	❒ Big variations in implementation capabilities – block granularity ; AST software maturity; degree on admin required
Network Cache	❒ Hot data automatically flows into the caching tier ❒ Deployment relatively non-disruptive ❒ Scalable solution for HPC apps	❒ Cache must be populated before it becomes effective ❒ In-network placement may constrain protocols or use cases

© 2011 Storage Networking Industry Association. All Rights Reserved.

(A) Solid State Disk Tier

Advantages:

- Fast random I/O for small blocks
- Low read and write latency time
- Low power consumption
- Low noise
- Better mechanical reliability

Disadvantages:

- Very high price, typically 10-30X comparable fast HDDs
- Relatively limited capacities
- Less advantaged random write speeds (erase of blocks)
- Less advantaged sequential write (HDD strength)

© 2011 Storage Networking Industry Association. All Rights Reserved.

Practical Use Cases

- **Database acceleration solution**
  - Entire database on SSD tier, or
  - Hot random access files on SSD and rest of database on standard disk
    - Indexes and temp space
- **Large scale virtual machine environments**
  - Solves "boot storm" problem for large numbers of virtual machines
  - Deduplication of VM data, e.g. virtual desktops
    - Reduces capacity requirements, increasing IOPS density, potentially making SSD economical

© 2011 Storage Networking Industry Association. All Rights Reserved.

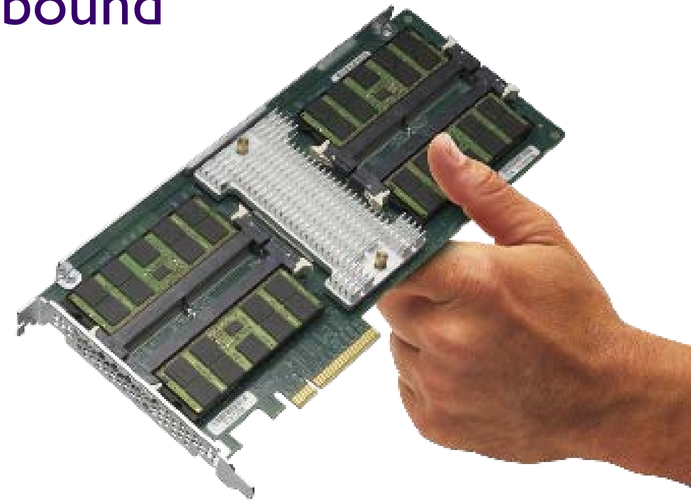
Automated Storage Tiering (AST)

- Utilizing both SSD and HDD for any workload will be the most cost-efficient use of SSDs in over the next few years
- Area of intense innovation among enterprise storage vendors
  - Media-spanning LUNs
  - Hybrid aggregates of volumes
  - Automated storage tiering software
- Issue is to automate data placement and movement
  - Automated tiering
  - Policy-based; no administrator overhead imposed
- Alternatively you can eliminate data movement
  - Caching implementations
  - Some vendors refer to this as tier-less storage

© 2011 Storage Networking Industry Association. All Rights Reserved.

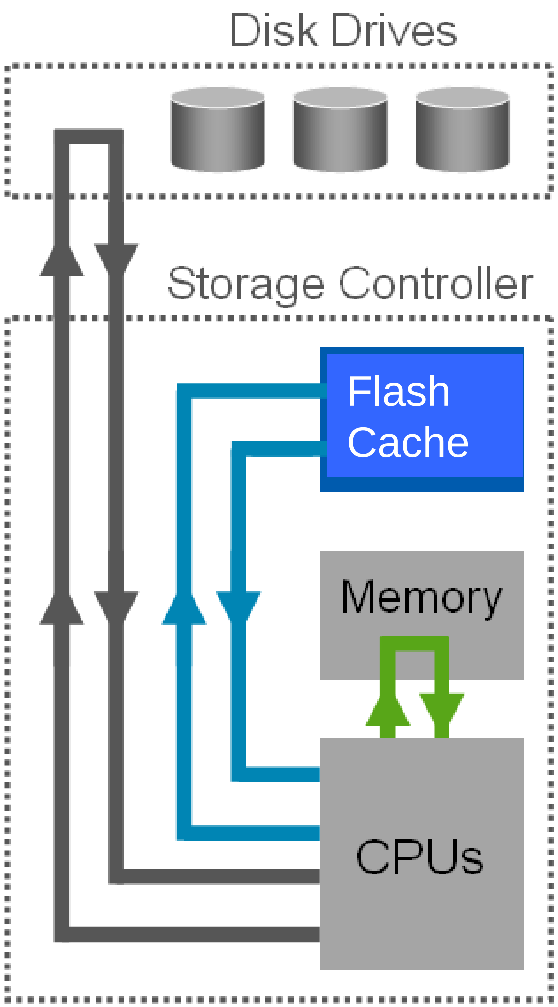
(B) Controller-based Flash Cache

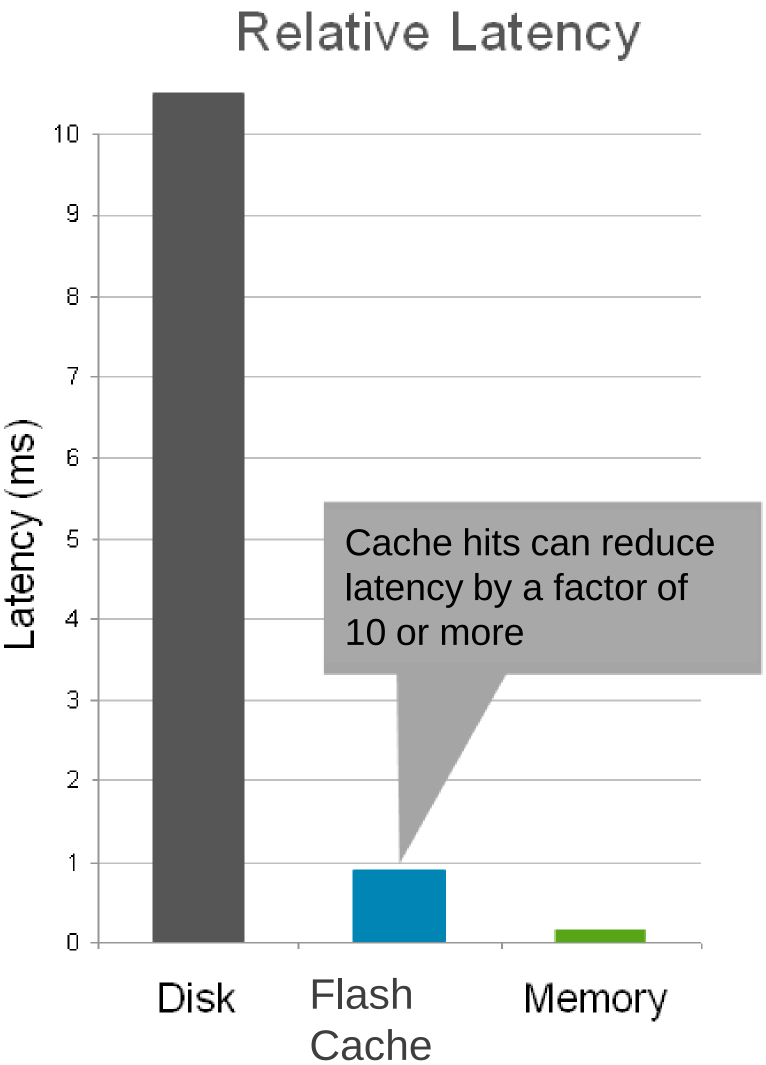
- Functions as an intelligent read cache for data and metadata
- Automatically places active data where access can be fast
- Provides more I/O throughput without adding high-performance disk drives to a disk-bound storage system
- Effective for file services, OLTP databases, messaging, and virtual infrastructure



© 2011 Storage Networking Industry Association. All Rights Reserved.

Reduce Latency with Flash Cache

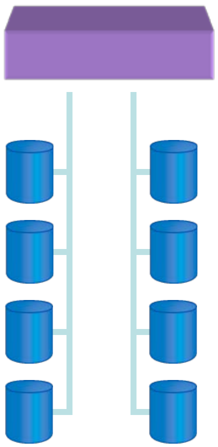




© 2011 Storage Networking Industry Association. All Rights Reserved.

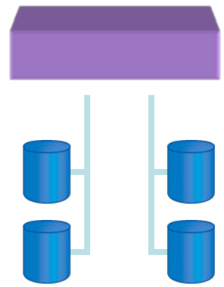
Use case: Scale Performance of Disk-bound Systems

Education SNIA



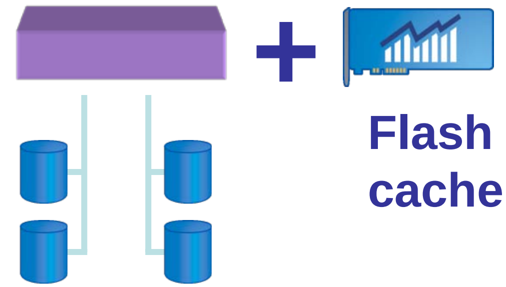
Add Spindles

- Use more disks to provide more IOPs
- May waste storage capacity
- Consumes more power and space



Starting Point:
Need More IOPs

- Performance is disk-bound
- Have enough storage capacity
- Random read intensive workload



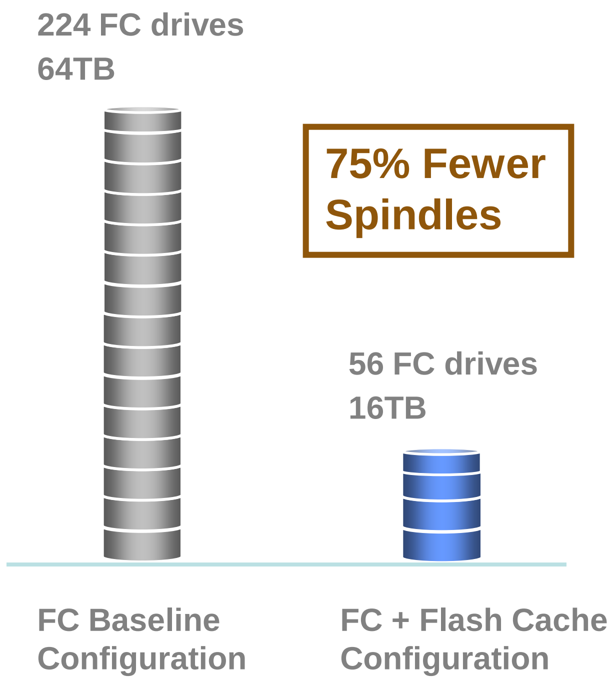
Add Flash Cache

- Use cache to provide more IOPs
- Improves response times
- Uses storage efficiently
- Achieves cost savings for storage, power, and space

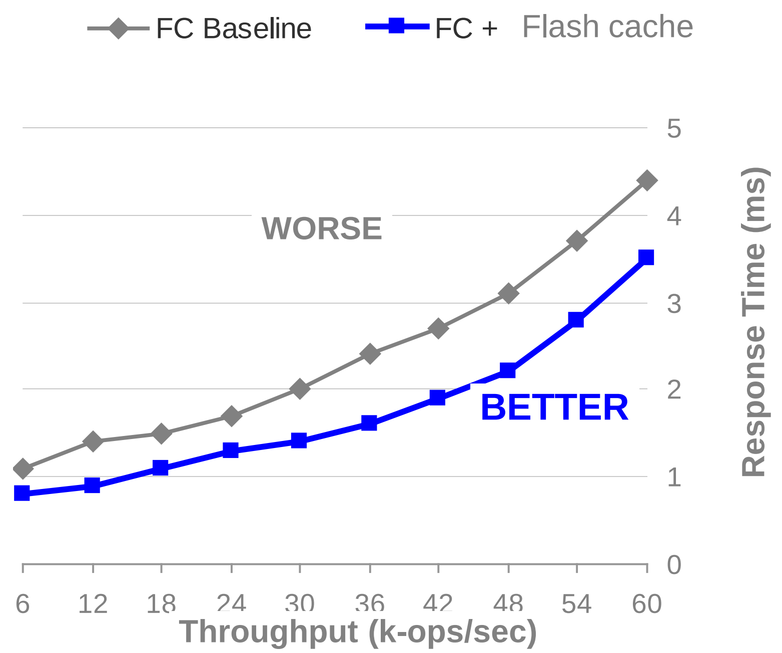
© 2011 Storage Networking Industry Association. All Rights Reserved.

FC HDD plus Flash Cache Example

Benchmarked Configurations



SPECsfs2008 Performance



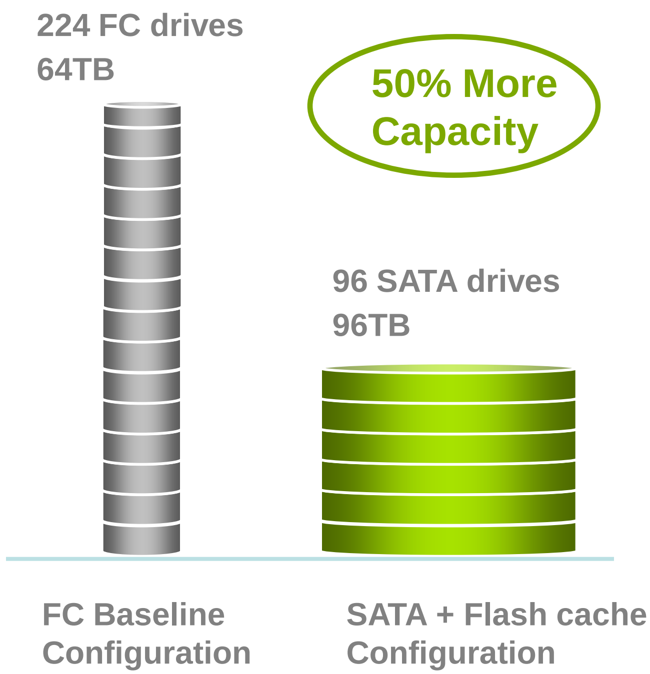
- Purchase price is **50% lower** for FC + Flash cache compared to Fibre Channel baseline
- FC + Flash cache yields **67% power savings** and **67% space savings**

For more information, visit http://spec.org/sfs2008/results/sfs2008nfs.html.
SPEC® and SPECsfs2008® are trademarks of the Standard Performance Evaluation Corp.

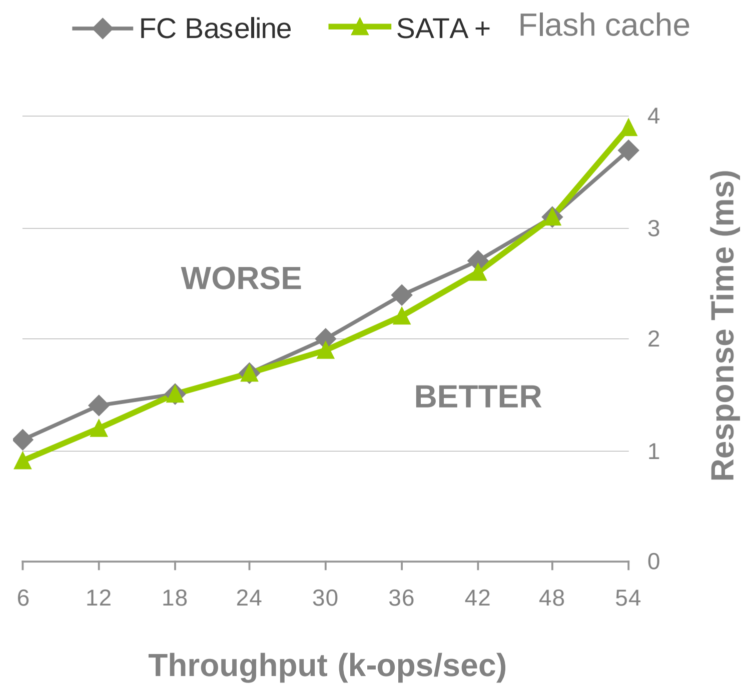
© 2011 Storage Networking Industry Association. All Rights Reserved.

SATA HDD plus Flash Cache Example

Benchmarked Configurations



SPECsfs2008 Performance



- Purchase price is **39% lower** for SATA + Flash cache compared to FC baseline
- SATA + Flash cache yields **66% power savings** and **59% space savings**

For more information, visit http://spec.org/sfs2008/results/sfs2008nfs.html.
SPEC® and SPECsfs2008® are trademarks of the Standard Performance Evaluation Corp.

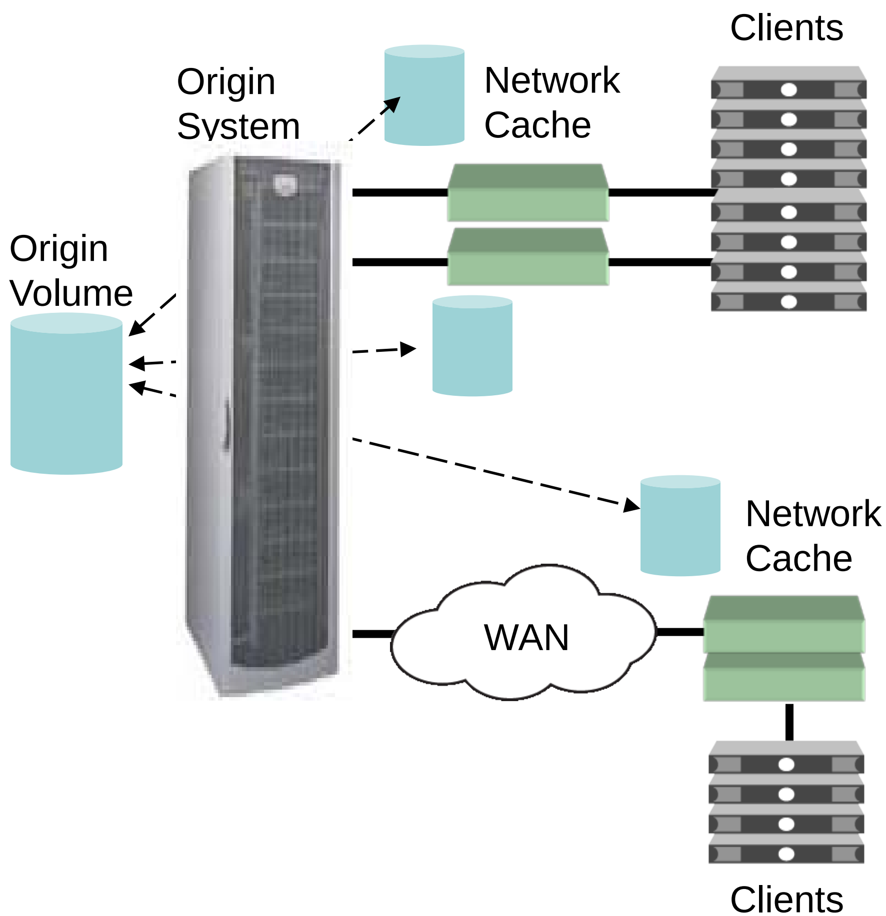
© 2011 Storage Networking Industry Association. All Rights Reserved.

(C) Host-based Flash

- Flash card on PCI bus in host system
- Can support storage semantics (SCSI) or device driver model
- Acts as Tier 0 storage (IOPS tier) in front of networked storage (capacity tier)
- Multiple implementations in development:
  - Requiring no data movement (caching)
  - Requiring data movement (AST)
- Area of intense industry and standards activity

© 2011 Storage Networking Industry Association. All Rights Reserved.

(D) Network Cache Topology



© 2011 Storage Networking Industry Association. All Rights Reserved.

Practical Use Cases

- **Network cache solutions**
  - All files on HDD in shared storage array
  - Accelerated by SSD-based network cache
  - Self–tuning write-through cache
- **Same pros and cons as SSD tier**
- **Typical applications**
  - Rendering
  - Seismic
  - Financial modeling
  - ASIC design

© 2011 Storage Networking Industry Association. All Rights Reserved.

28

Cost Structure of Memory/Storage Technologies

<u>Cost</u> determined by

- cost per wafer
  - # of dies/wafer
    - memory area per die [sq. µm]
      - memory density [bits per $4F^2$]
        - patterning density [sq. µm per $4F^2$]

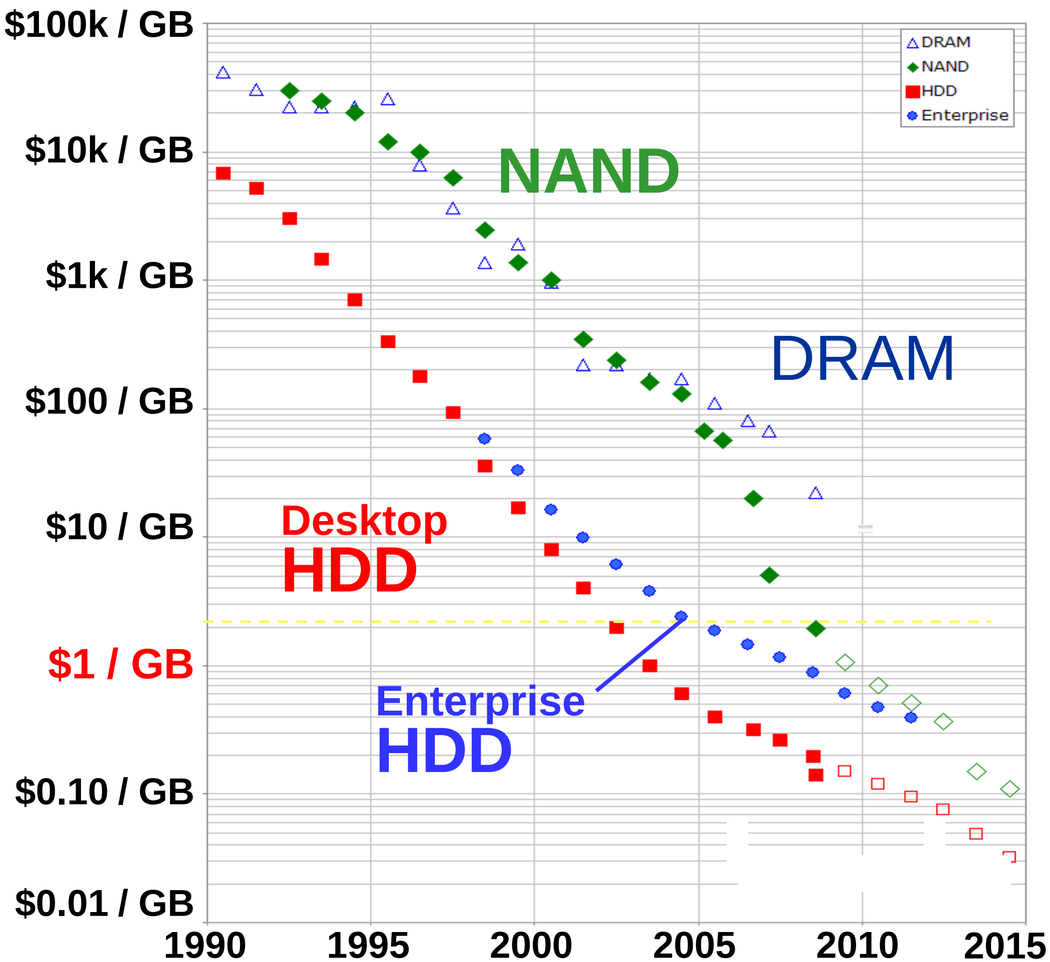
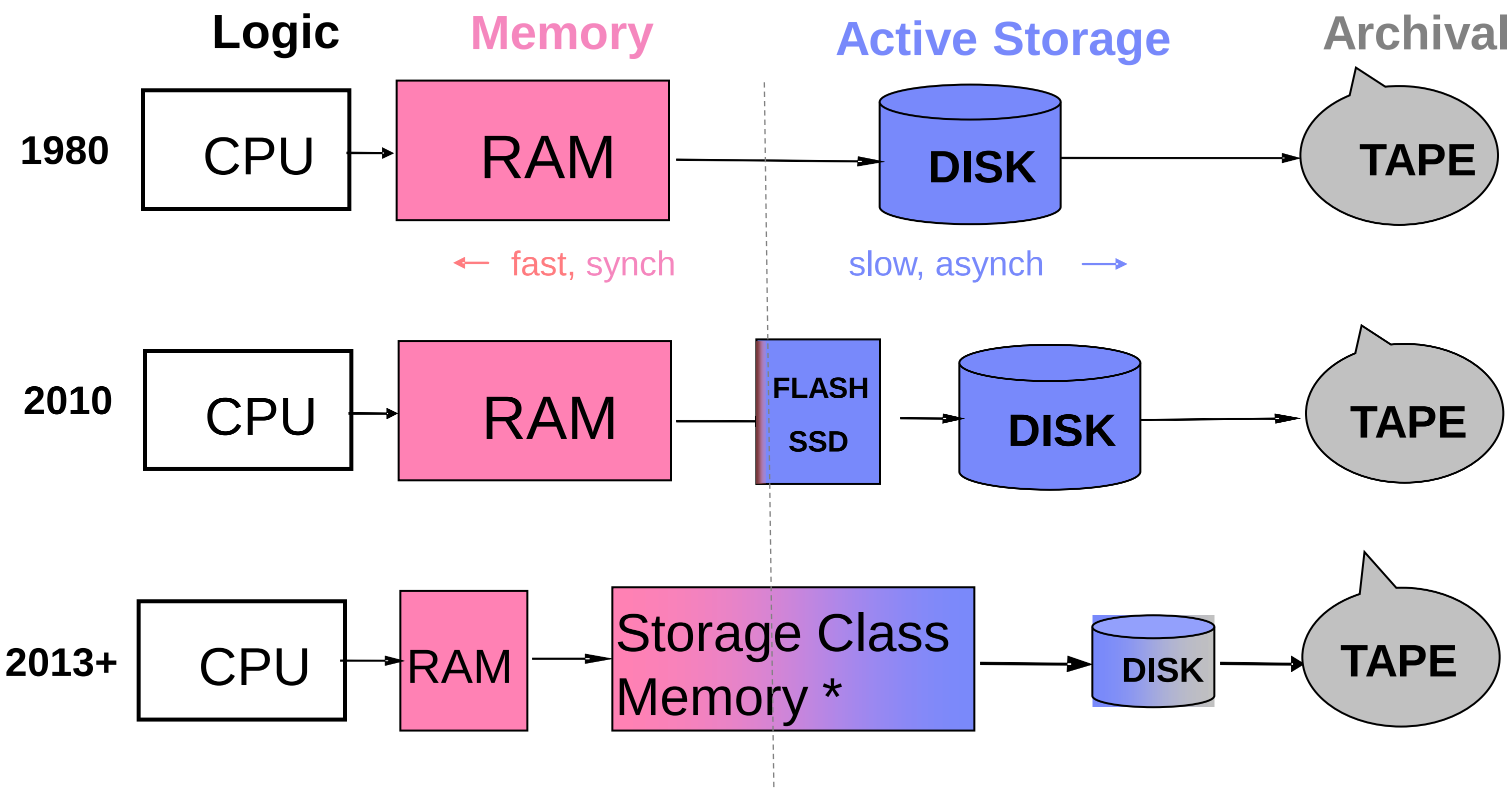


Chart courtesy of Dr. Chung Lam, IBM Research updated version of plot from 2008 *IBM Journal R&D* article

© 2011 Storage Networking Industry Association. All Rights Reserved.

System Evolution



* e.g. Phase change memory
Memristor
Solid Electrolyte
Racetrack memory

© 2011 Storage Networking Industry Association. All Rights Reserved.

Summary

Education SNIA

- Over the next 5 years solid state technologies will have a profound impact on enterprise storage
- It's not just about replacing mechanical media with solid state media
- The architectural balance of memory, cache and persistent storage will change
- Today's solid state implementations in enterprise storage demonstrate these changes
- It's still only the beginning…

© 2011 Storage Networking Industry Association. All Rights Reserved.

Q&A / Feedback

Education SNIA

- Please send any questions or comments on this presentation to SNIA: add your track reflector here

Many thanks to the following individuals for their contributions to this tutorial.

- SNIA Education Committee

David Dale	**Amit Shah**
Jeff Kimmel	**Mark Woods**
Chris Lionetti	**Alan Yoder**
Phil Mills	

© 2011 Storage Networking Industry Association. All Rights Reserved.