SNIA

Education

# The Benefits of Solid State in Enterprise Storage Systems

Roberto Salucci, HDS

# 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.

# Agenda

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



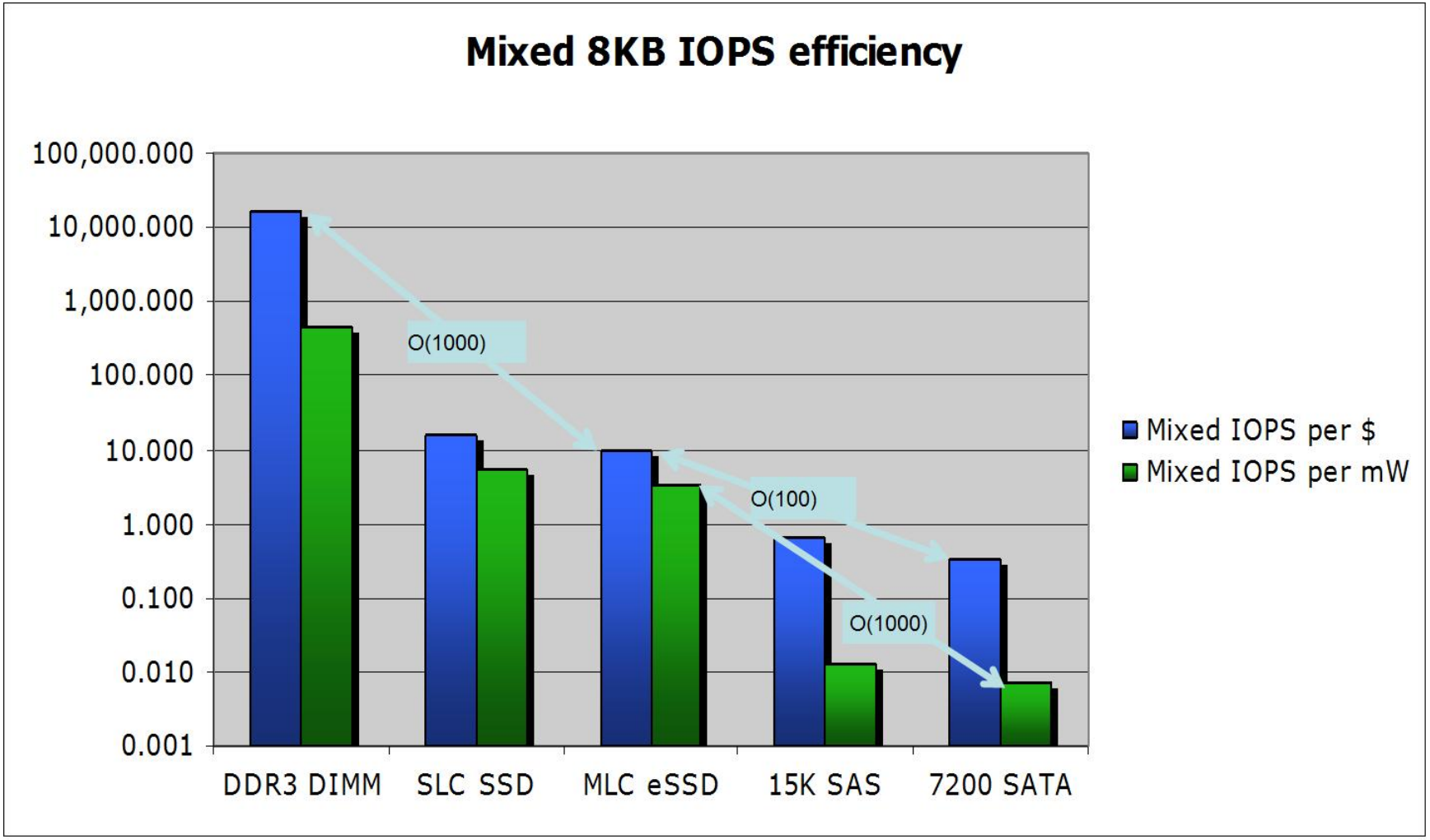

3

# Why Flash in the Data Center Now?

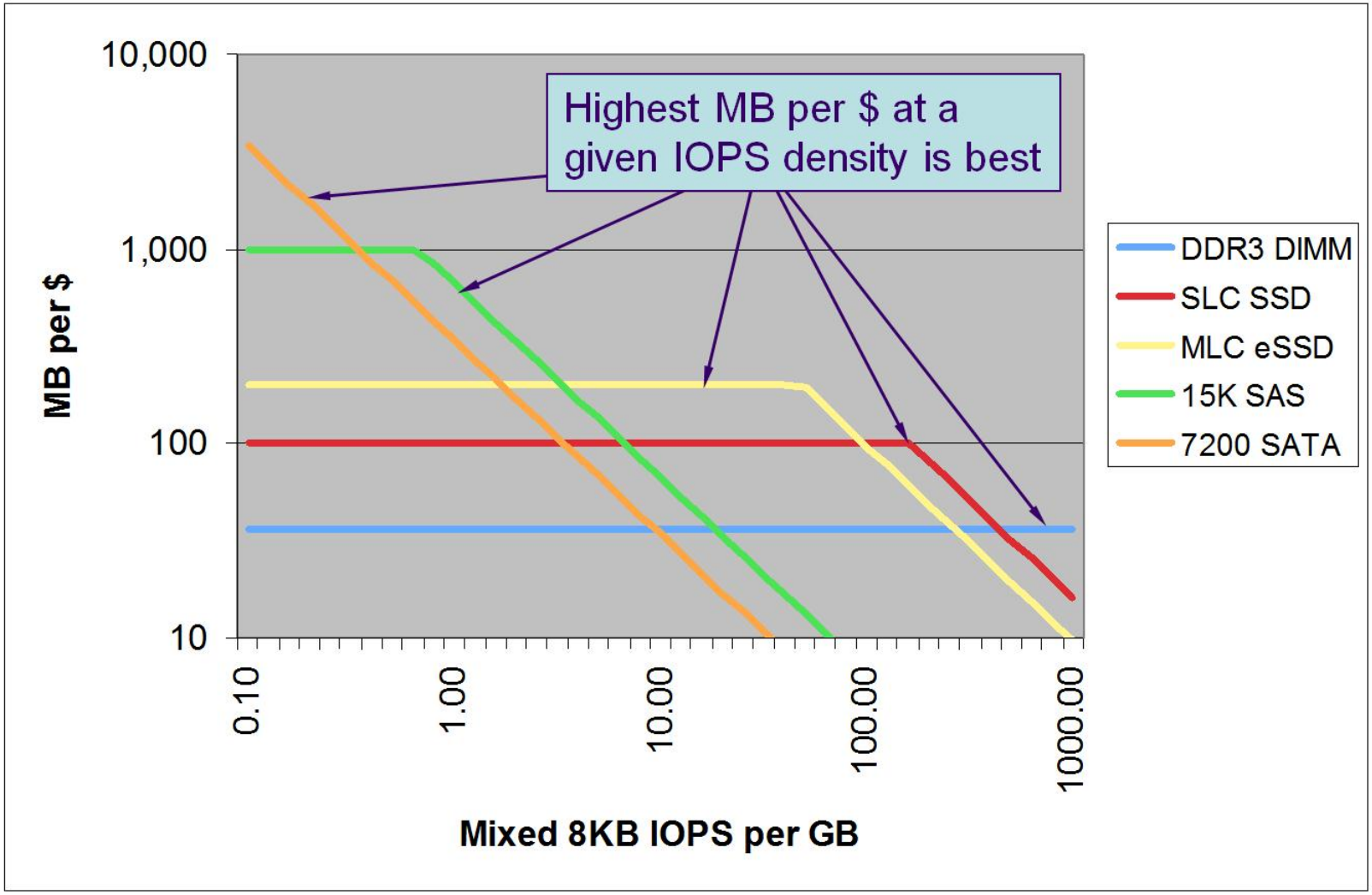
Education SNIA

- 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

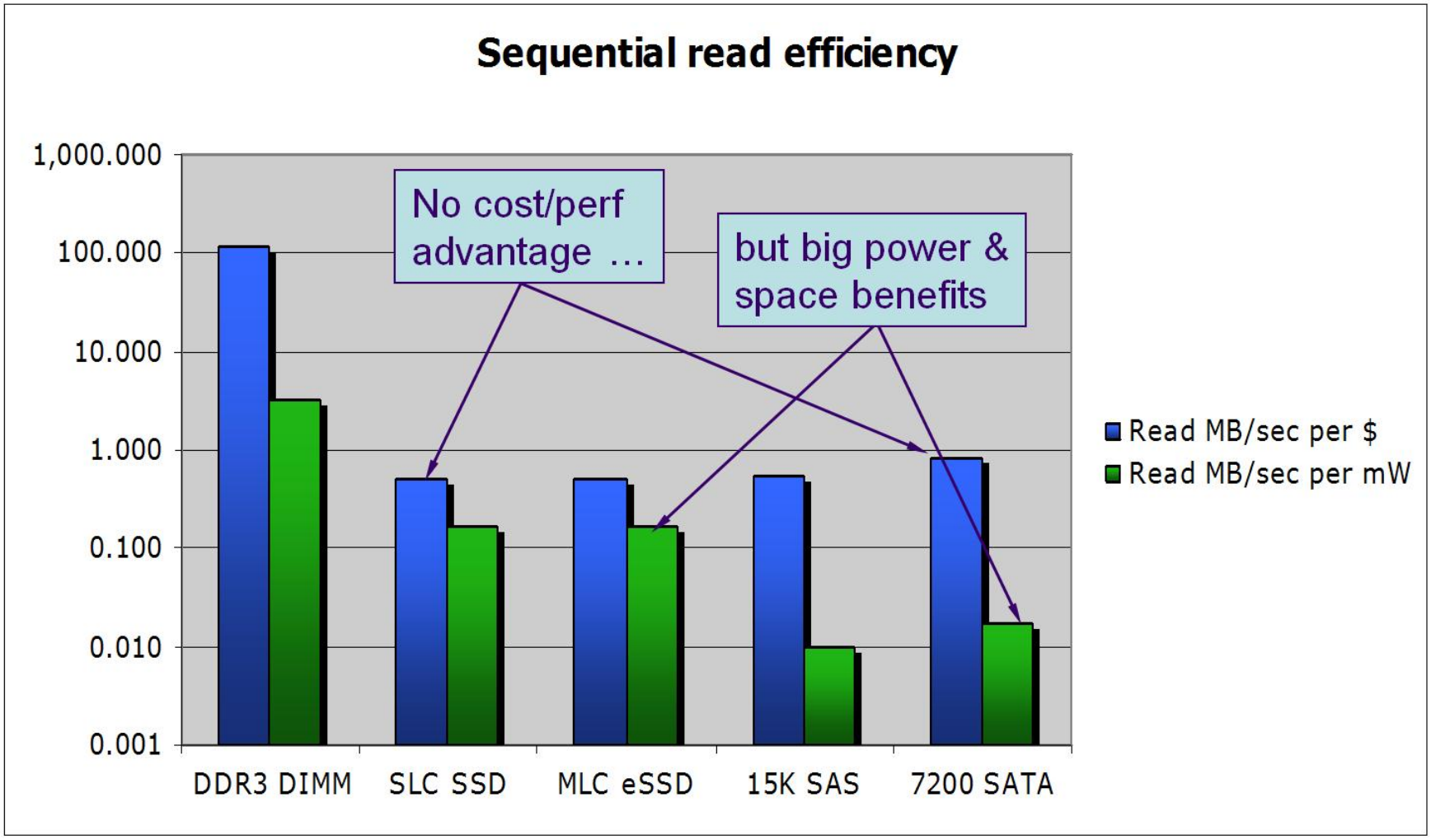
# Why Flash? IOPS Efficiency



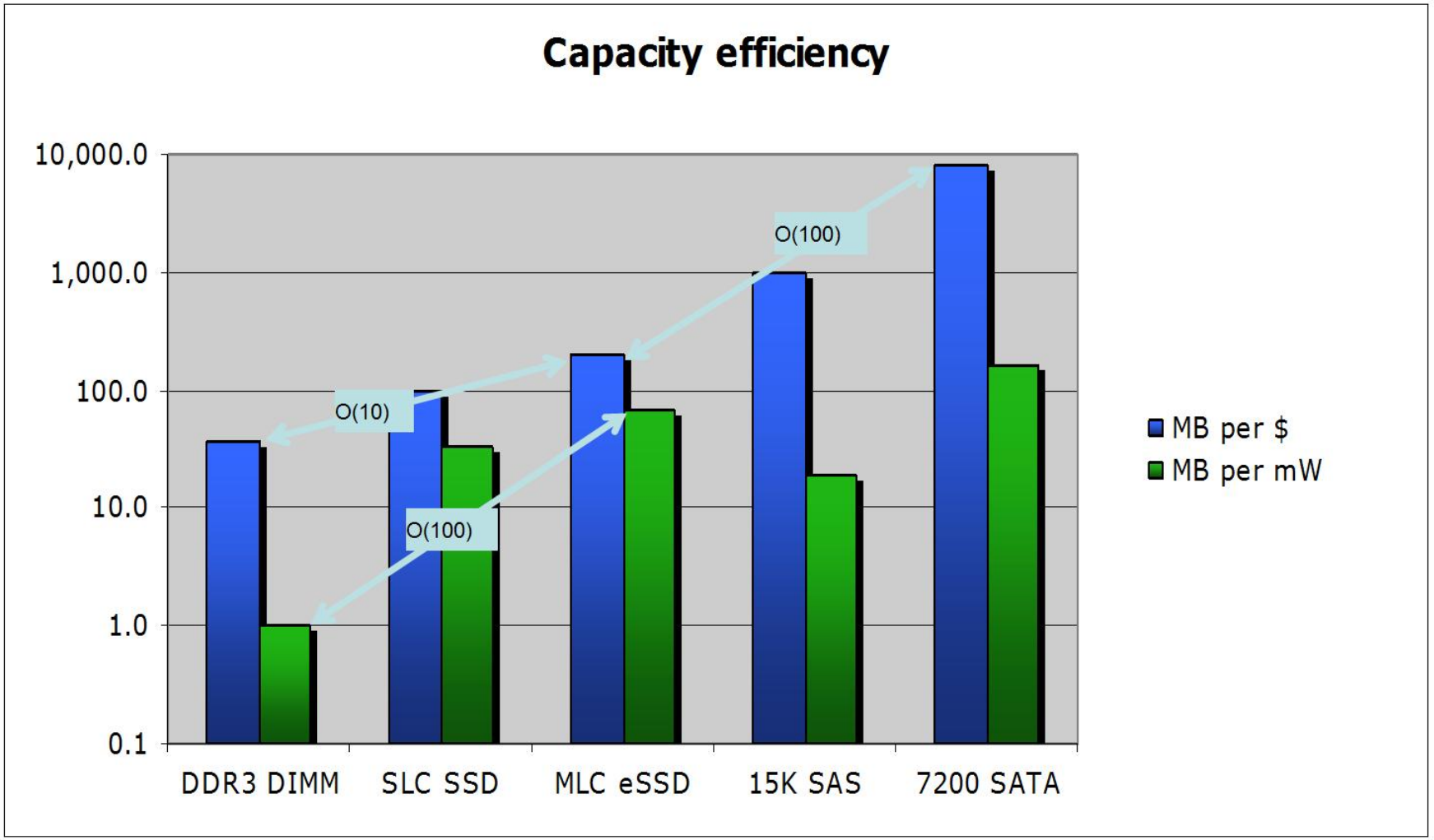
# Why Flash? An IOPS Density View



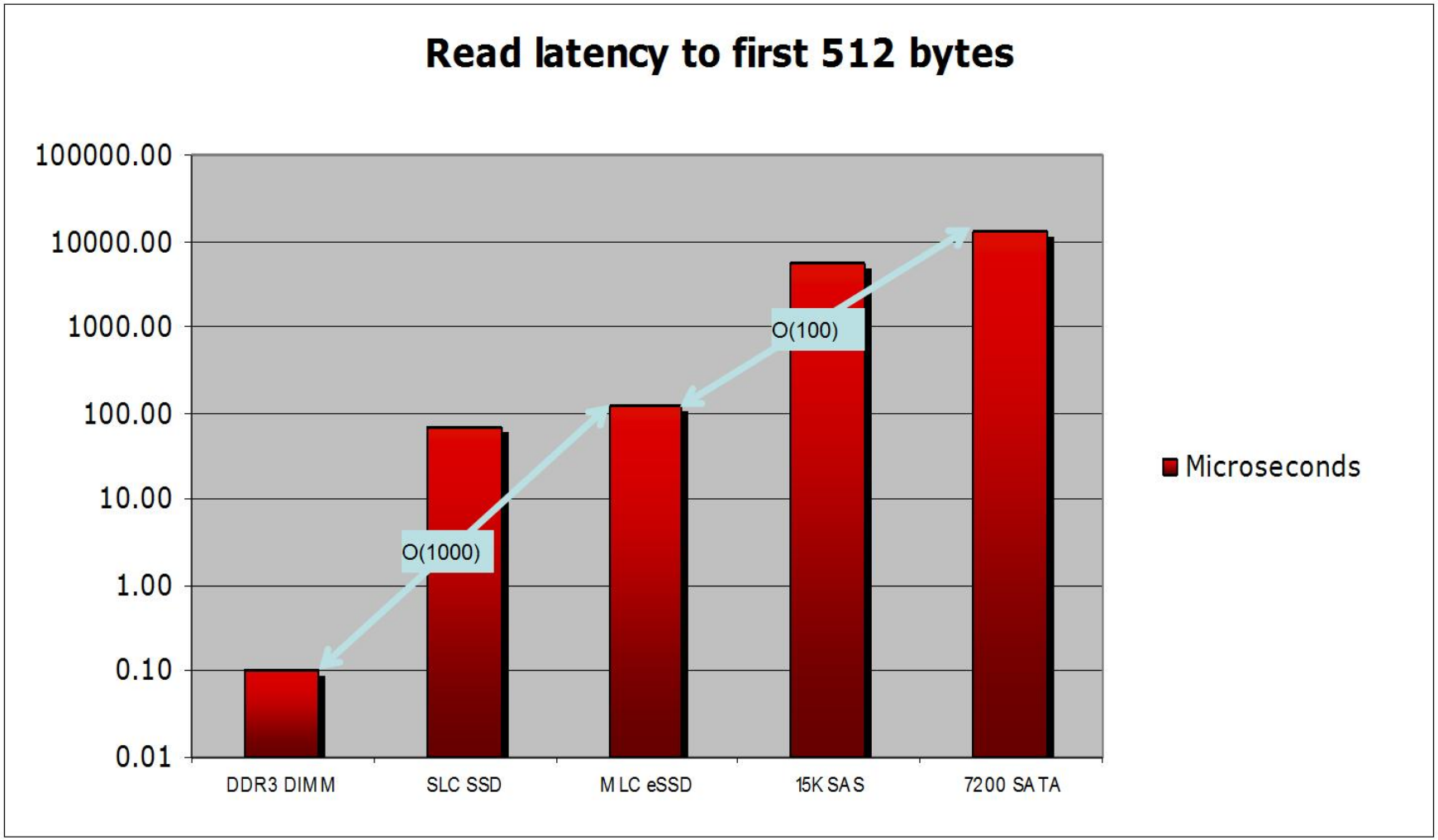
# Why Flash? Read Throughput per Watt



# Why Flash? Capacity Efficiency



# Why Flash? Read Latency



# Application Opportunities

- Intense random reads, e.g. OLTP, metadata
- Sequential read after random write
  - Log-oriented writes convert this to random read after sequential write (e.g. FTL)
- Low read latency (~100x better than HDD)
  - Facilitates DRAM extension by allowing high read throughput with limited read concurrency
  - Paging datacenter apps can be practical again
  - Memory capacity to consolidate more servers with underutilized CPU
- Enabling memory-resident datasets, e.g.
  - OLTP
  - Data warehouses (*viz* TPC-H results)
  - Large metadata




10

# End-to-End Flash Categories

C

## Host-side Flash Software

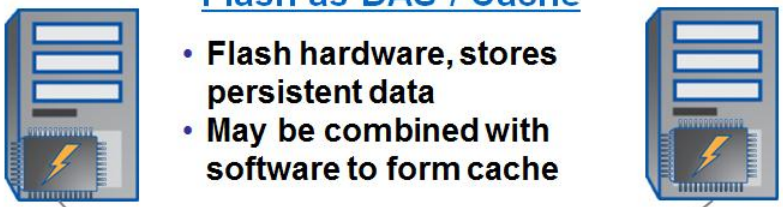
- Software only, may be tied to particular flash hardware

## Flash as DAS / Cache

- Flash hardware, stores persistent data
- May be combined with software to form cache

## Flash-based Virtual Storage Appliance

- Software



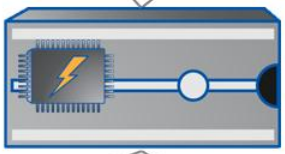
## Network-based Flash

- Flash Hardware and software
- "Bump in the wire"

B

## Flash in Storage Controller

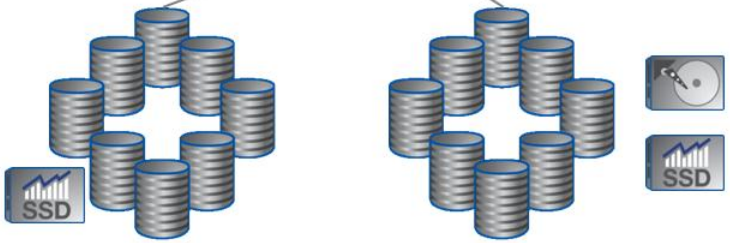
- Flash hardware and software
- "Behind wire"
- E.g. Flash Cache



A

## Pure Flash in Array

- All flash




## Hybrid Flash / HDD Array

- Mixed flash / HDD
- E.g. Flash Pool

# (A) Solid State Disk Tier

Education SNIA

## 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 10-30 X comparable FC drives
- Limited capacities
- Slow random write speeds, e.g. erase of blocks
- Slow sequential write throughput

# 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



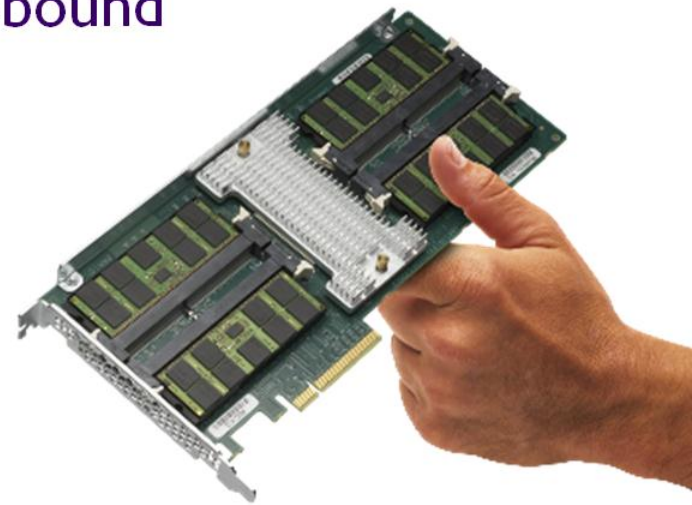

13

# Automated Tiering or Tier-less

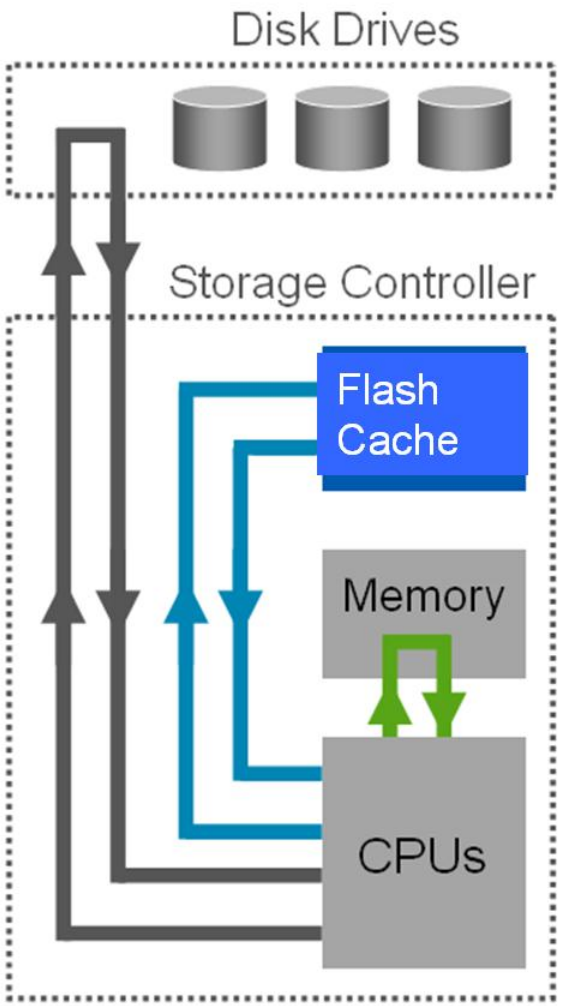
- Mixing SSD and HDD for a particular workload will probably be the most cost-efficient use of SSDs in over the next few years
- Area of intense innovation among enterprise storage vendors
- Issue is to automate data placement and movement
  - Automated tiering
  - Policy-based
  - No administrator overhead imposed
  - Some vendors refer to this as tier-less storage
- As SSD prices fall this will become increasingly important

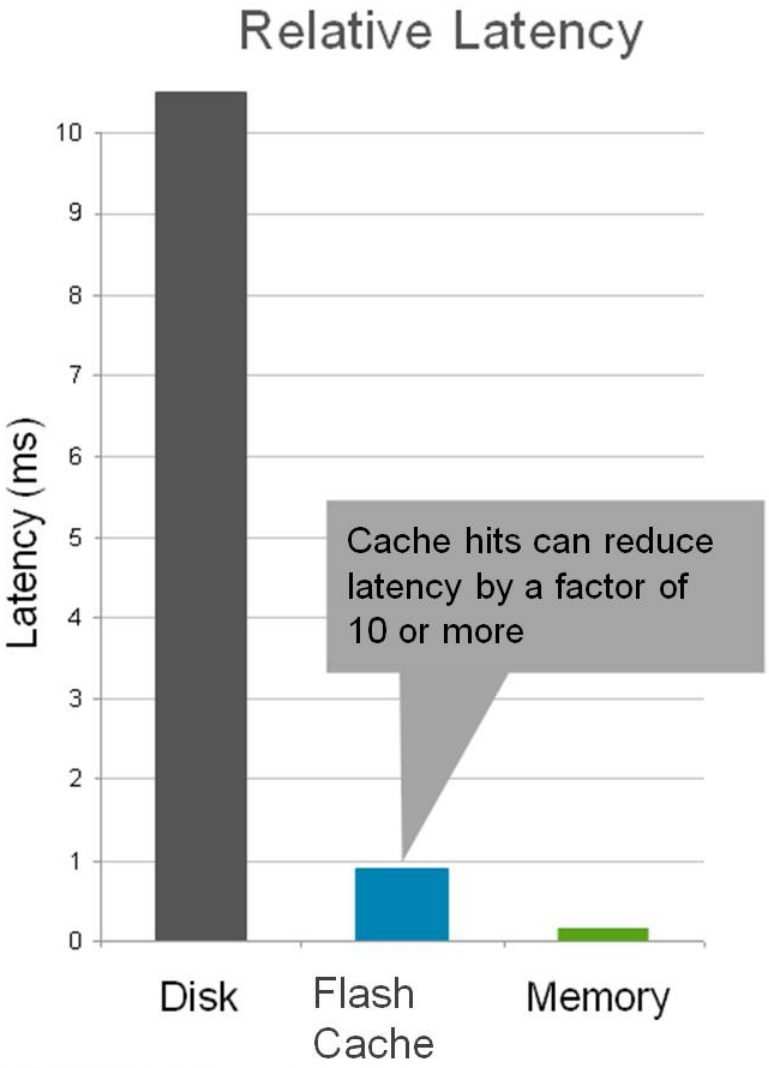
# (B) Controller-based Flash Cache

- 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

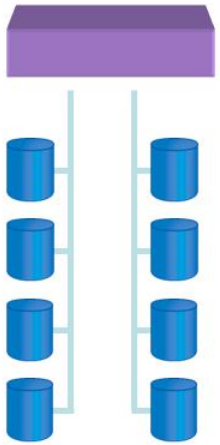


# Reduce Latency with Flash Cache



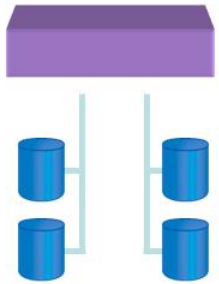



# Use case: Scale Performance of Disk-bound Systems



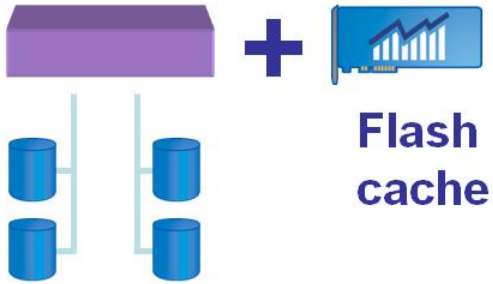
## Add Spindles

- 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

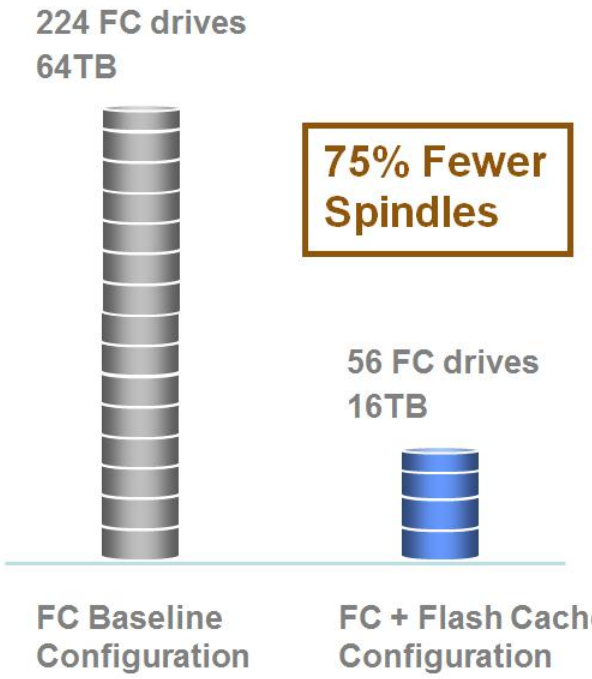



## Add Flash Cache

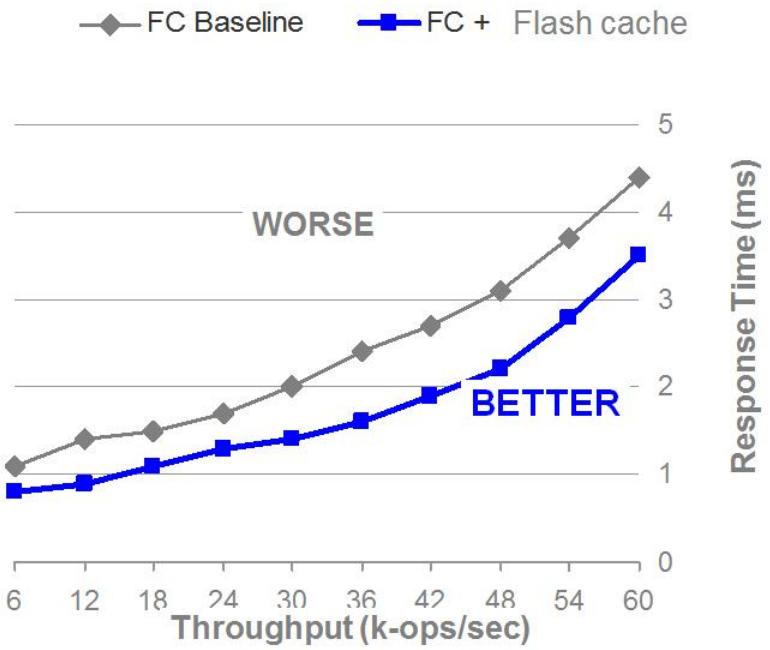
- Use cache to provide more IOPs
- Improves response times
- Uses storage efficiently
- Achieves cost savings for storage, power, and space

# FC HDD plus Flash Cache Example

## Benchmarked Configurations




## SPECsfs2008 Performance

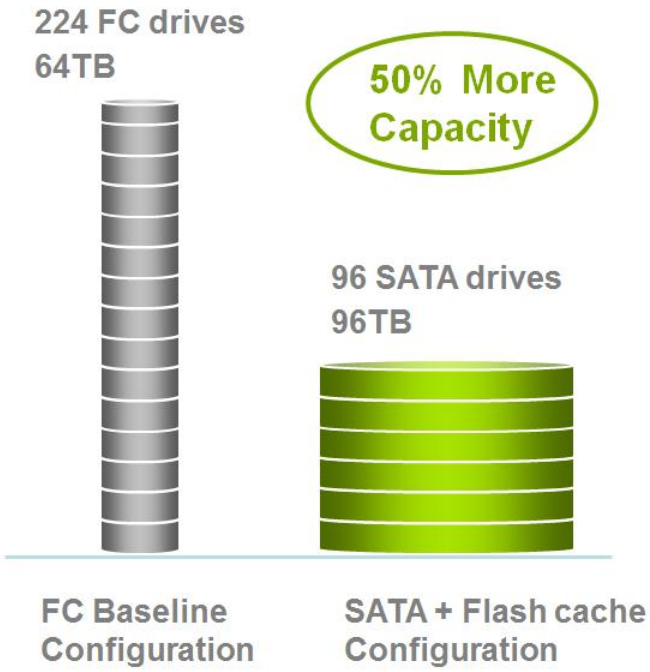



- Purchase price is **50% lower** for FC + Flash cache compared to Fibre Channel baseline
- FC + Flash cache yields **67% power savings** and **67% space savings**

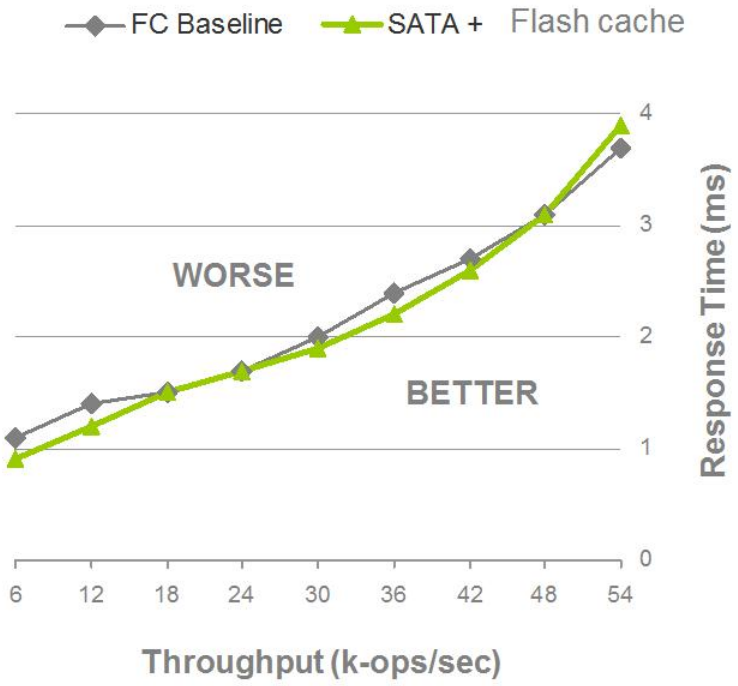
For more information, visit http://spec.org/sfs2008/results/sfs2008nfs.html.
SPEC® and SPECsfs2008® are trademarks of the Standard Performance Evaluation Corp.

# 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.

# (C) Host-based Flash

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

# Cost Structure of Memory/Storage Technologies

Education SNIA

**Cost** 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 $4F^2$]

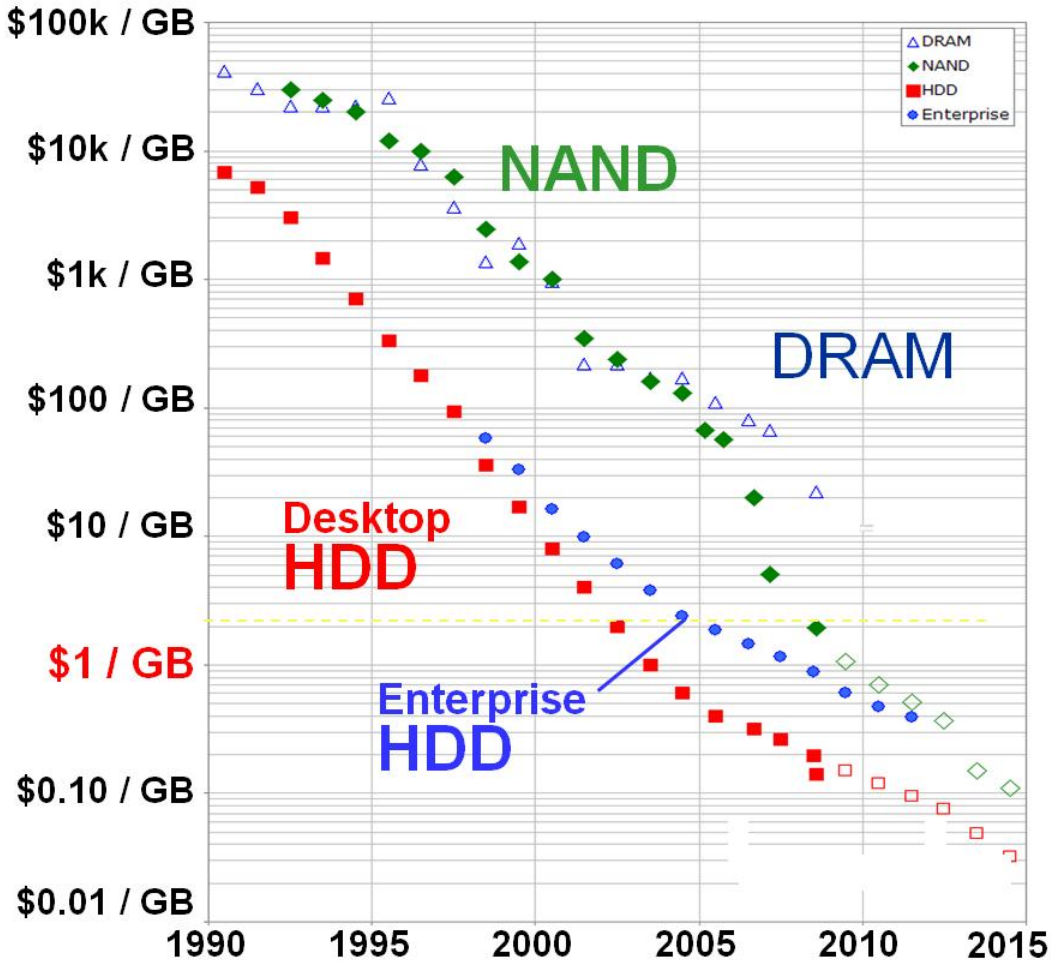
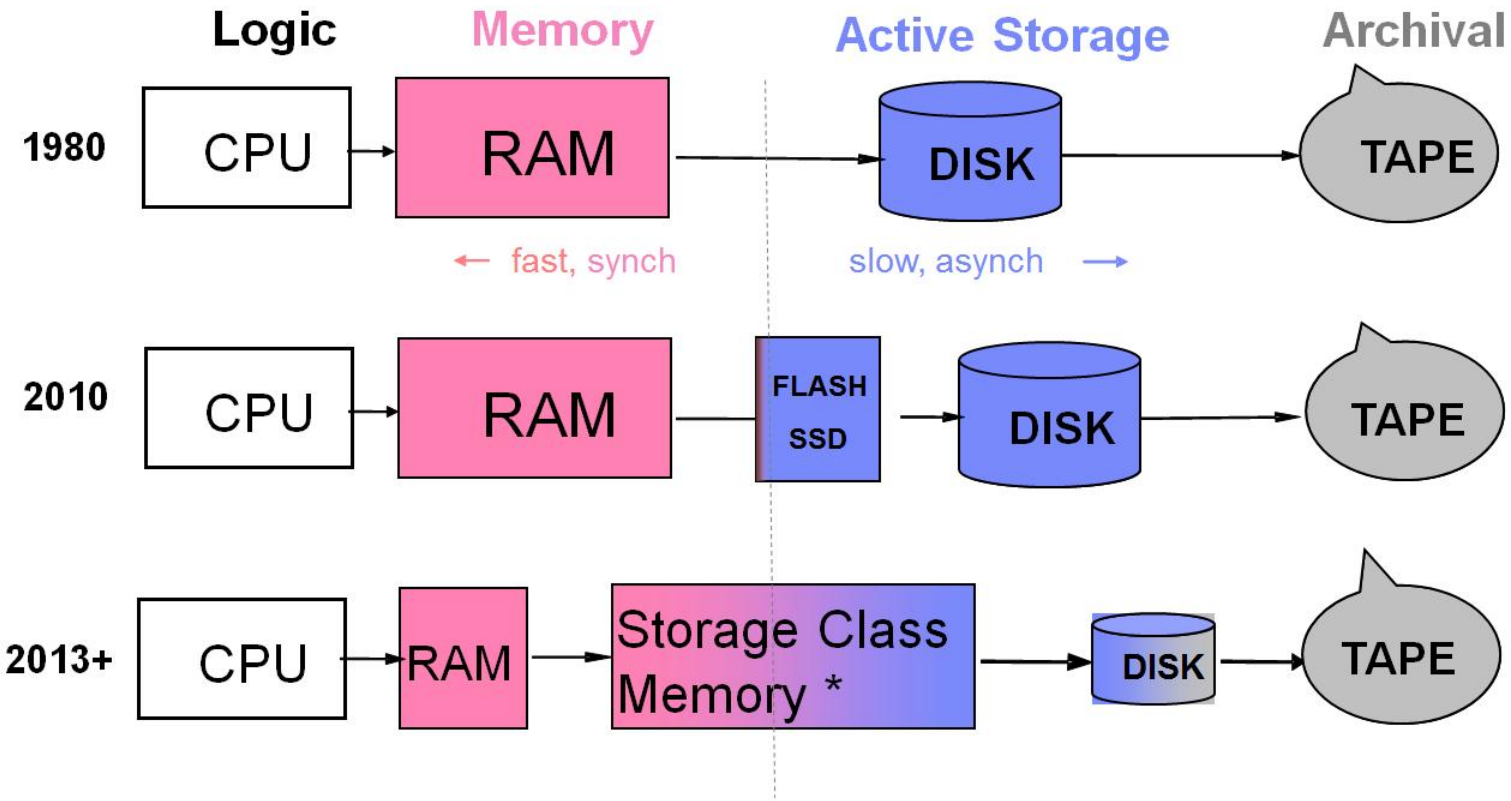



Chart courtesy of Dr. Chung Lam, IBM Research updated version of plot from 2008 *IBM Journal R&D* article

# System Evolution




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

# Summary

- 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 only the beginning…

# Attribution & Feedback

The SNIA Education Committee would like to thank the following individuals for their contributions to this Tutorial.

## Authorship History

**Name/Date of Original Author here:**
David Dale; Spring 2010

**Updates:**
David Dale; Fall 2010
David Dale; Spring 2011
David Dale; Fall 2011
David Dale; Spring 2012

## Additional Contributors

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

*Please send any questions or comments regarding this SNIA Tutorial to*
***tracktutorials@snia.org***