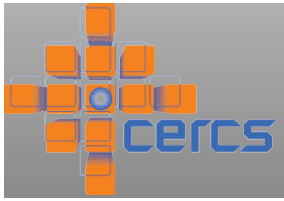
Impact of Flash Memory on Video-on-Demand Storage

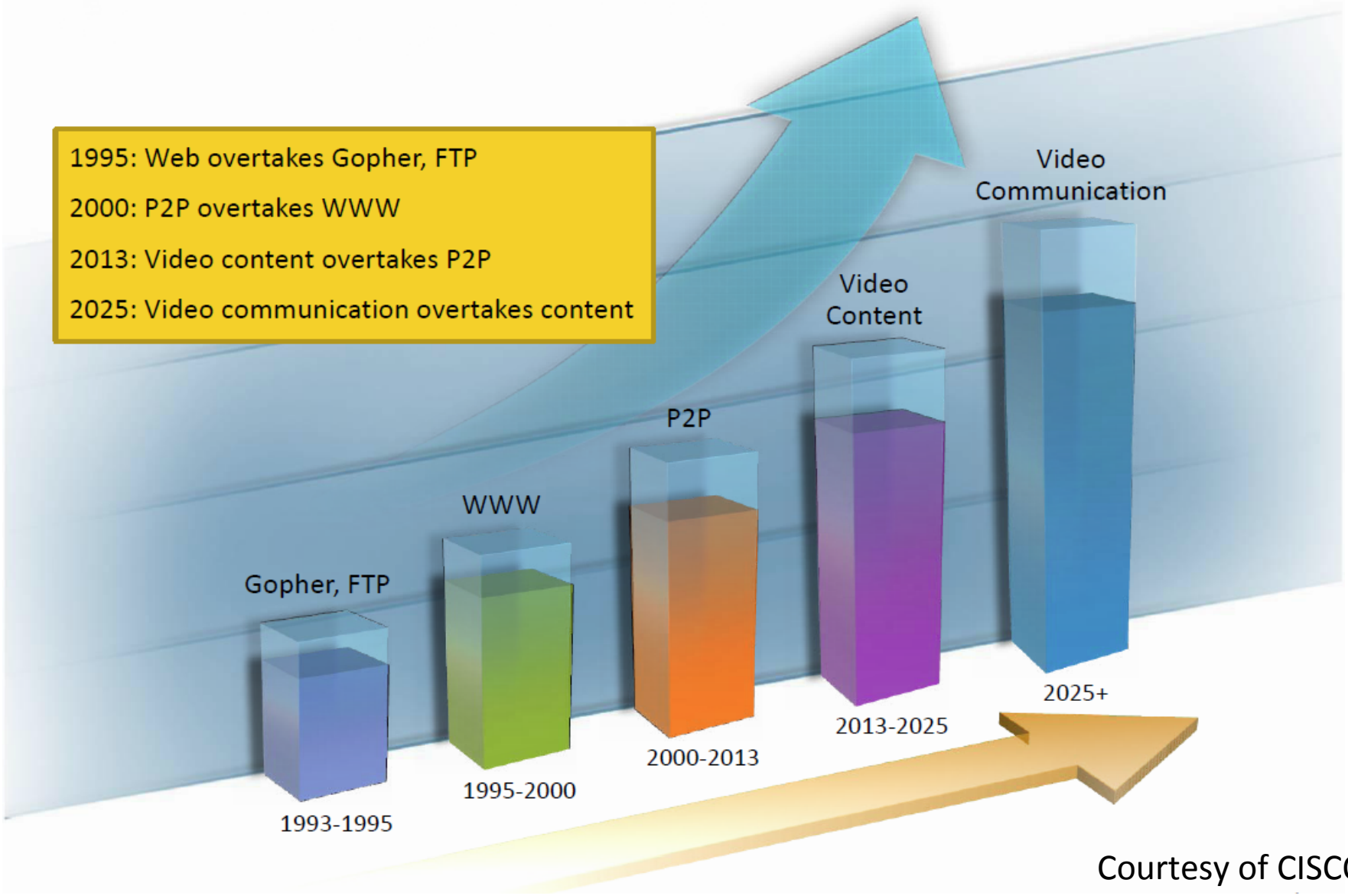
Moonkyung Ryu



Contents

- Motivation
  - What is the problem?
- Background
  - Interval Caching
  - Flash Memory
- How to utilize Flash SSD?
- What kinds of Flash SSD is appropriate?
- What is the best strategy?
- Conclusion

Internet Video Rocks!

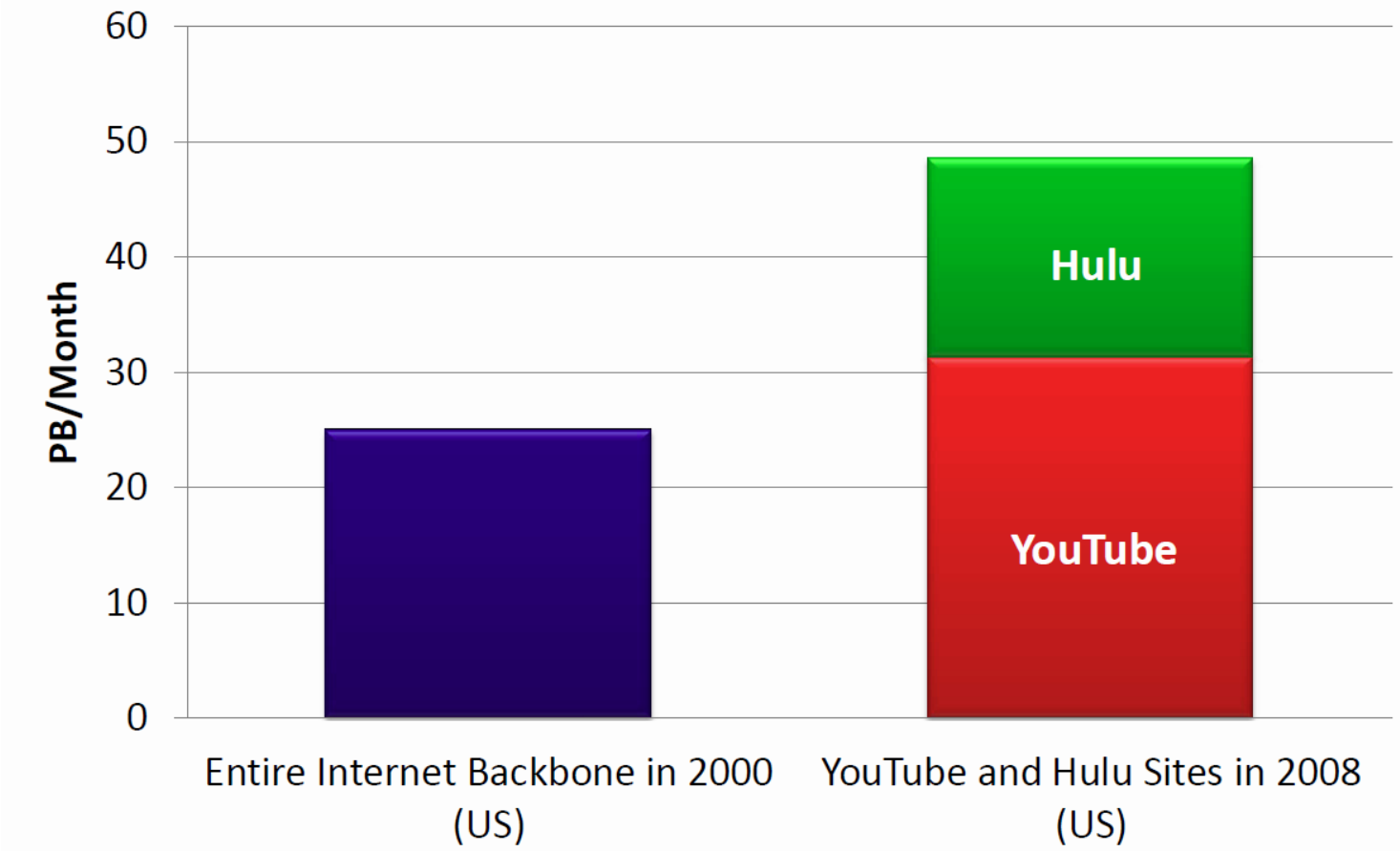


Courtesy of CISCO

Internet Video Rocks!

YouTube and Hulu Traffic

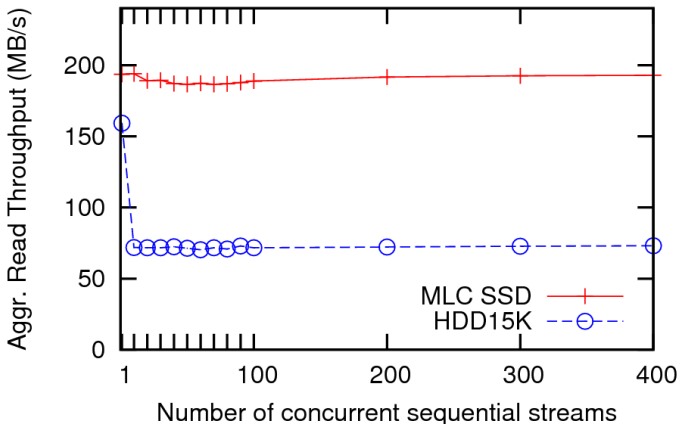
Courtesy of CISCO



VoD Storage Requirements

- Video is continuous media data
- Real-time retrieval
- Large Bandwidth & Large Capacity

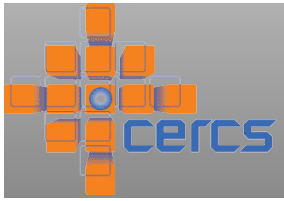
Device	Random Read Throughput (I/O Request Size 1MB)	Cost per GB	Cost per MB/s
DRAM	> 20GB/s	$23.0	$0.001
MLC SSD	> 155MB/s	$1.88	$0.774
HDD15K	70MB/s	$1.23	$2.571



VoD Storage Strategies

Device	Random Read Throughput (I/O Request Size 1MB)	Cost per GB	Cost per MB/s
DRAM	> 20GB/s	$23.0	$0.001
MLC SSD	> 155MB/s	$1.88	$0.774
HDD15K	70MB/s	$1.23	$2.571

- HDD only
- HDD + DRAM
- HDD + Flash SSD
- Questions
  - How to utilize Flash SSD?
  - What kinds of Flash SSD is appropriate?
  - What is the best strategy?



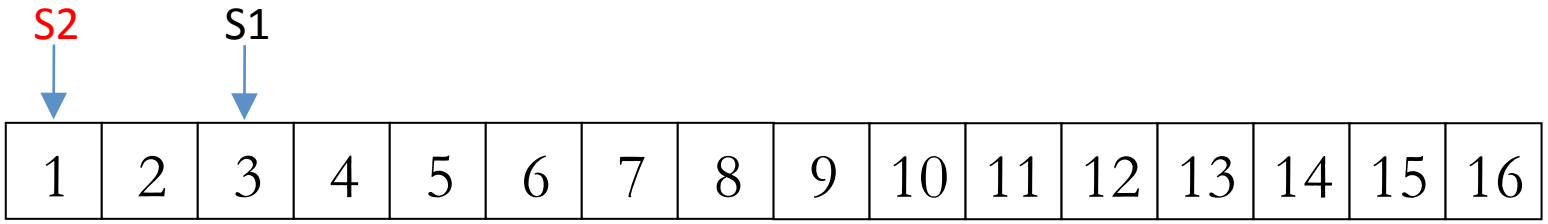
Contents

- Motivation
  - What is the problem?
- Background
  - Interval Caching
  - Flash Memory
- How to utilize Flash SSD?
- What kinds of Flash SSD is appropriate?
- What is the best strategy?
- Conclusion

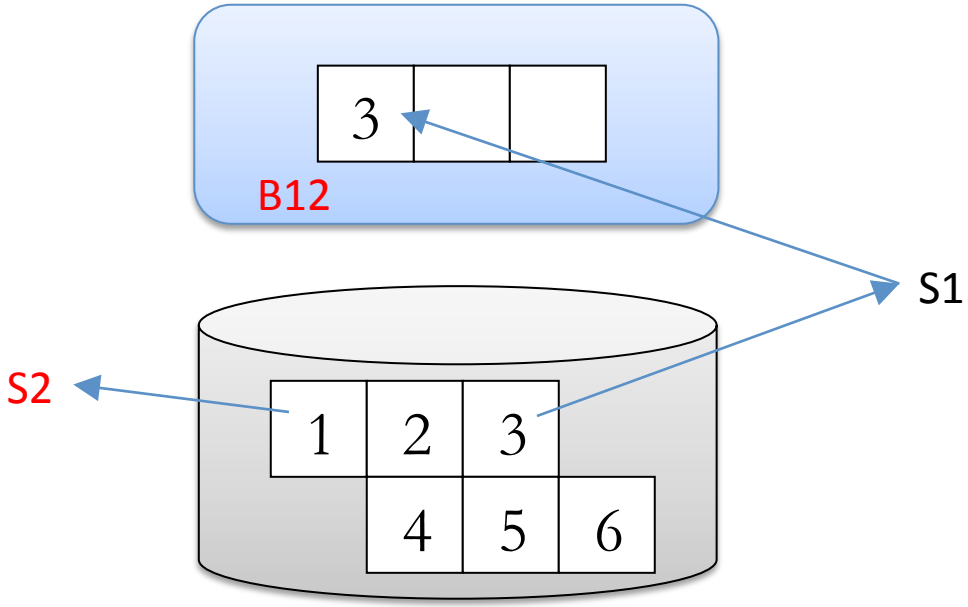
HDD + DRAM

- File server or Web server
  - Small file size
  - No real-time support
  - Block-level or File-level LRU (Least Recently Used)
- Video server
  - Large file size
  - Real-time support
    - via resource reservation
  - Interval Caching
    - Sequential video access
    - Caching the interval between two streams accessing th e same video

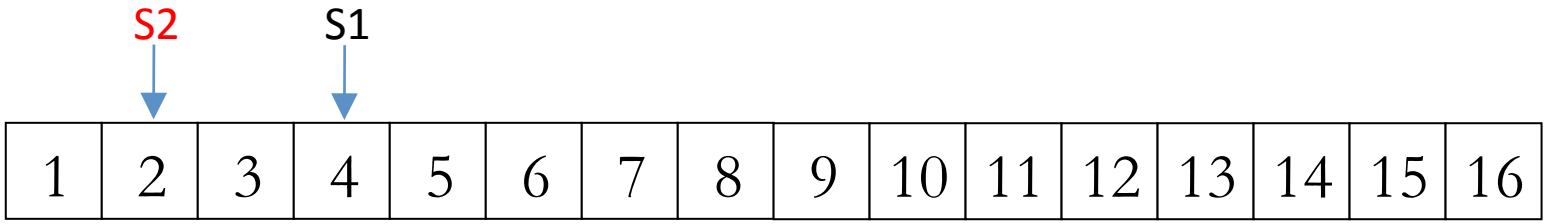
Interval Caching



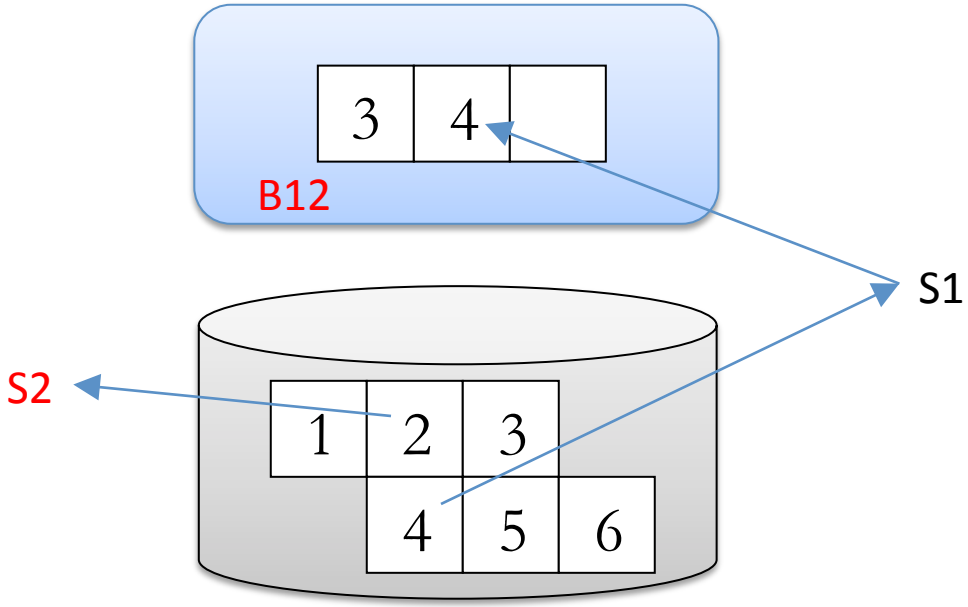
Video block sequences



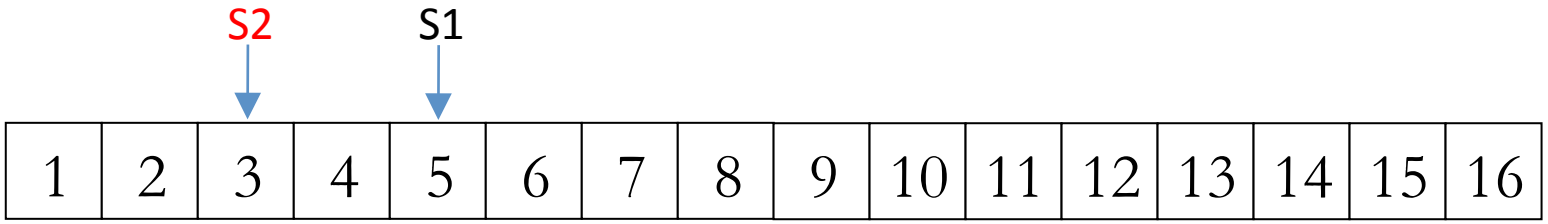
Interval Caching



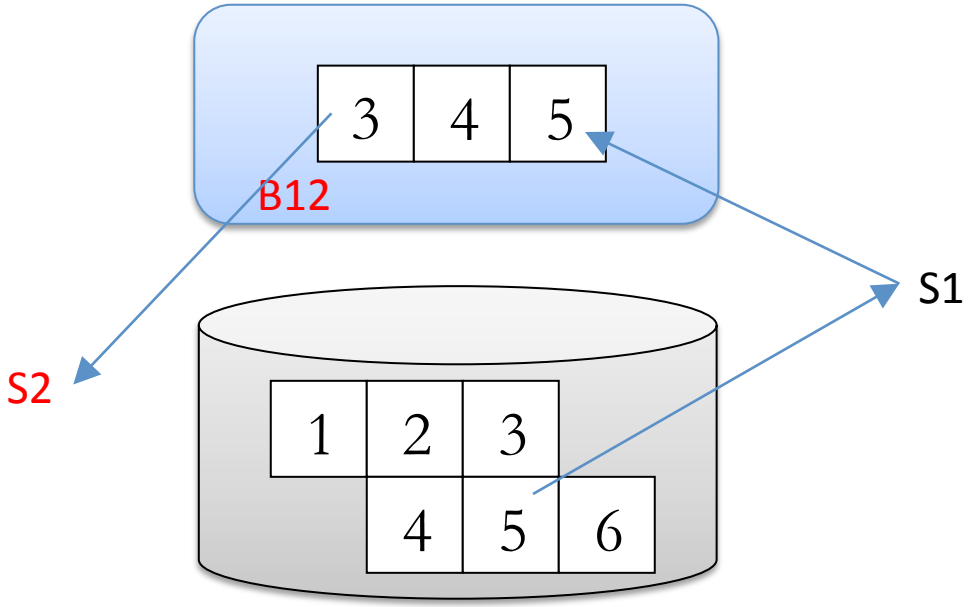
Video block sequences



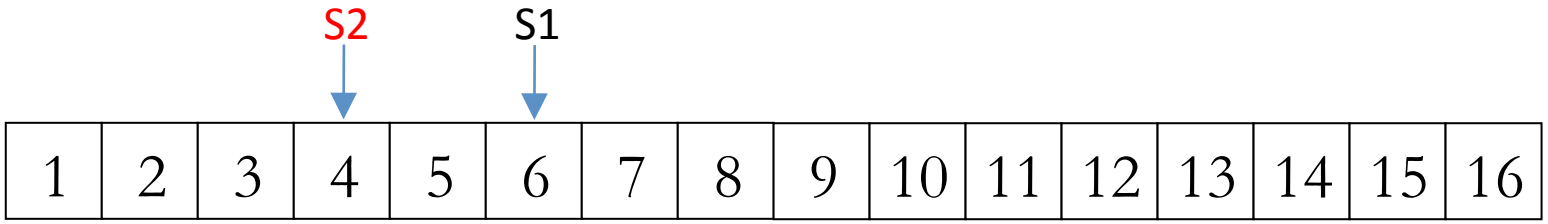
Interval Caching



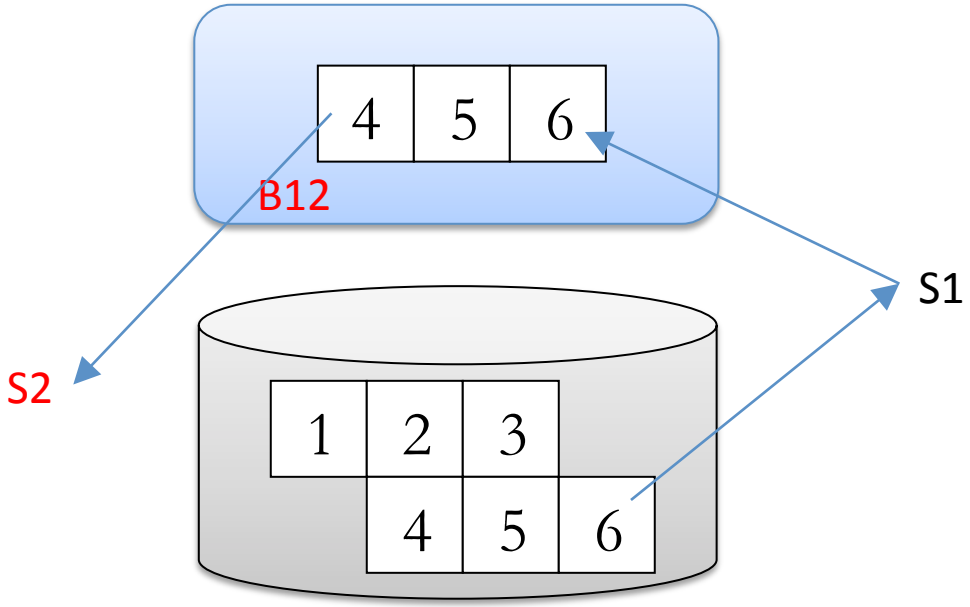
Video block sequences



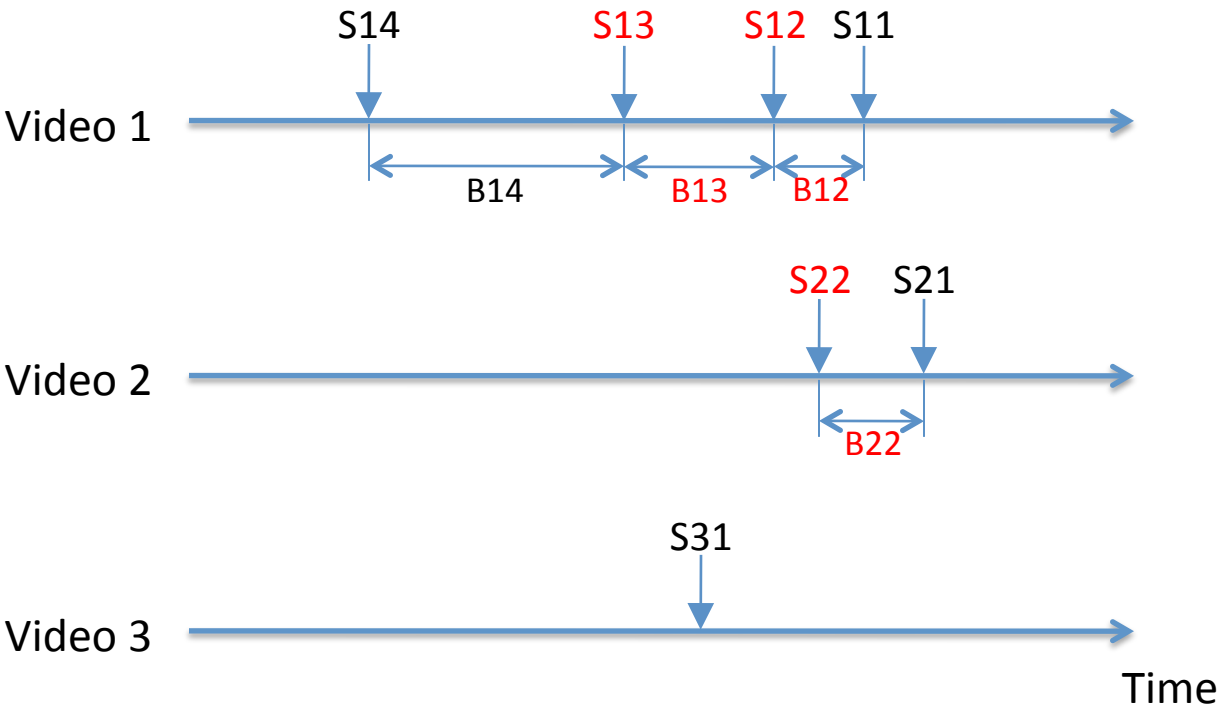
Interval Caching



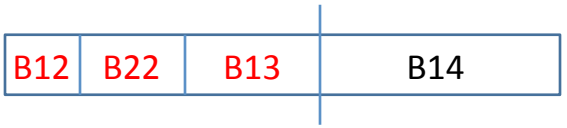
Video block sequences



Interval Caching



Ordered buffer requirement list:

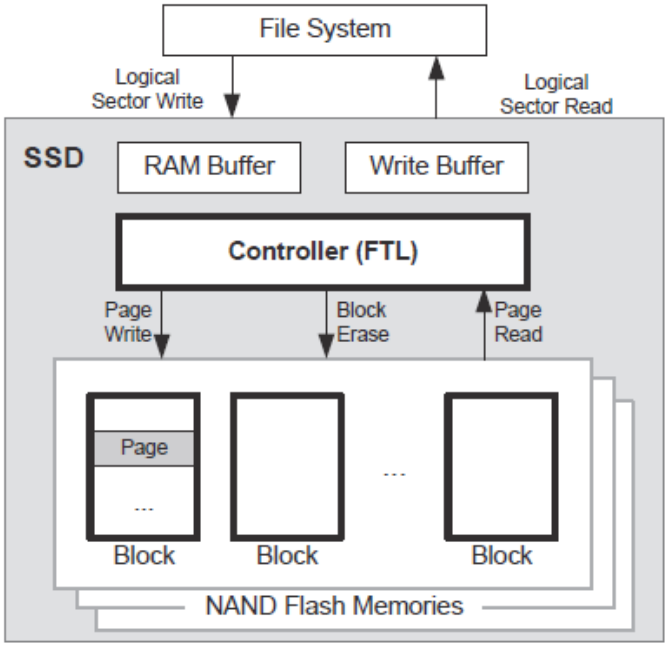


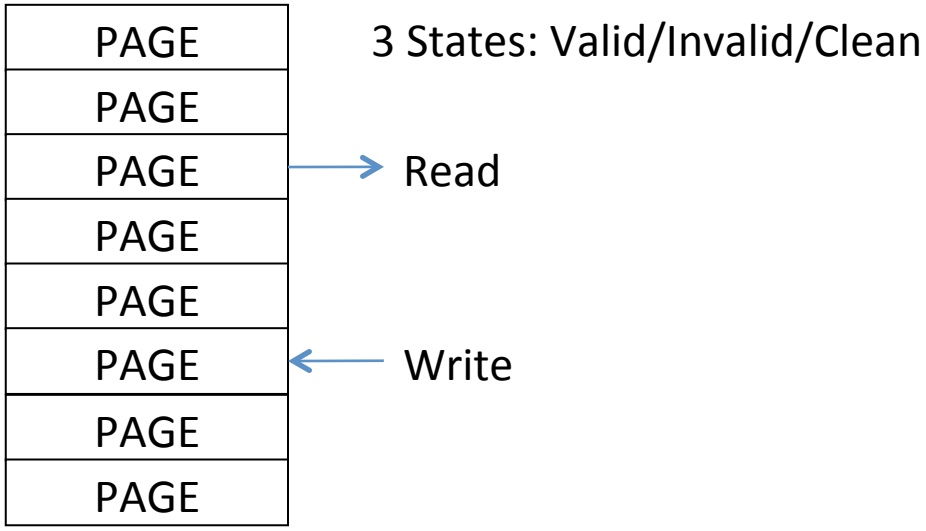
Buffered streams: S12, S22, S13

Interval Caching

- Beauty
  - 100% cache hit for buffered streams
    - RAM space is reserved
    - Real-time support
- Pitfalls
  - No support for fast-rewind or fast-forward
    - Assume the same data rate for streams accessi ng the same video

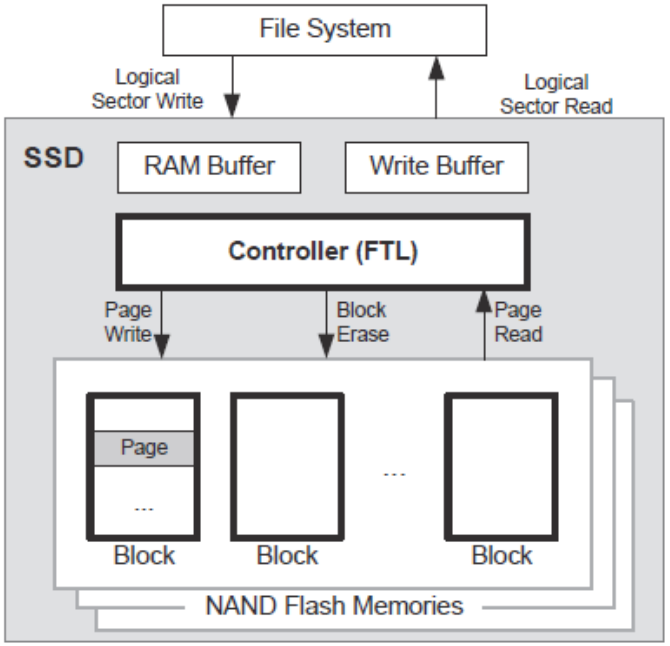
Flash Memory

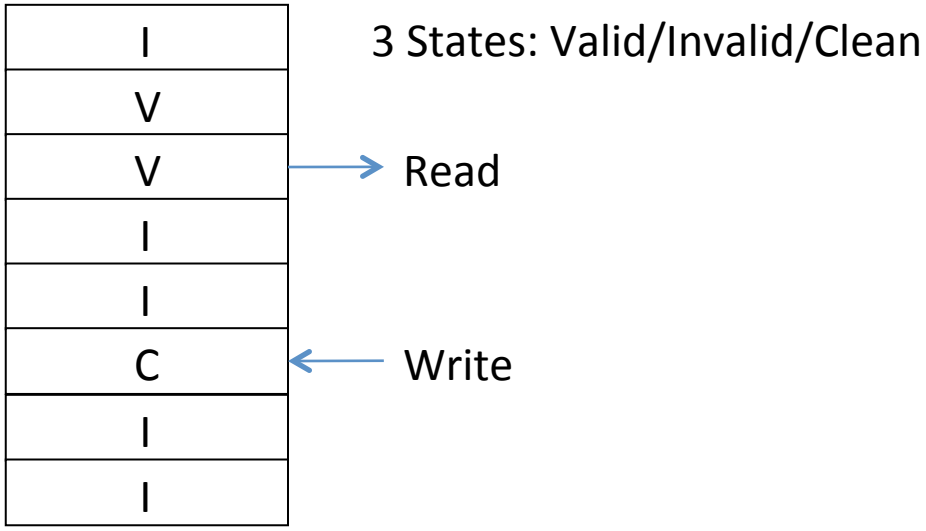




Flash Type	Data Unit Size			Access Time		
	Page (Bytes)		Block (Bytes)	Page READ (us)	Page WRITE (us)	Block ERASE (ms)
	Data	OOB				
Small Block	512	16	(16K+512)	41.75	226.75	2
Large Block	2048	64	(128K+4K)	130.9	405.9	2

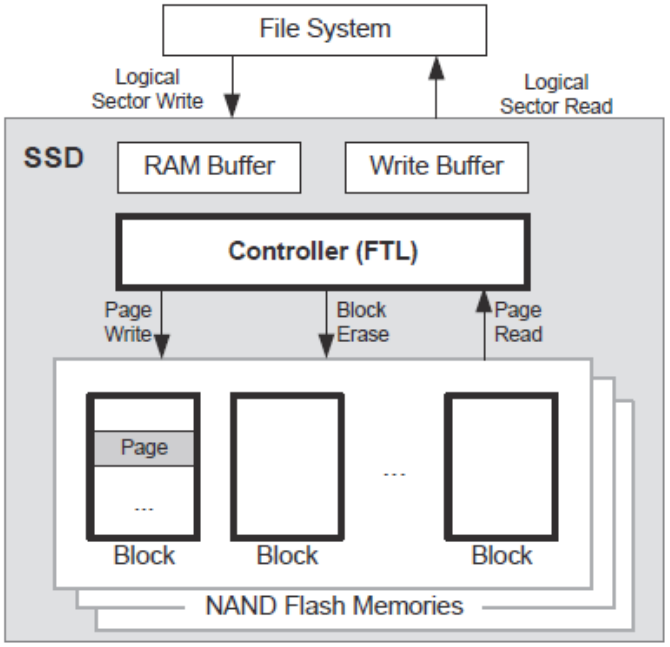
Flash Memory

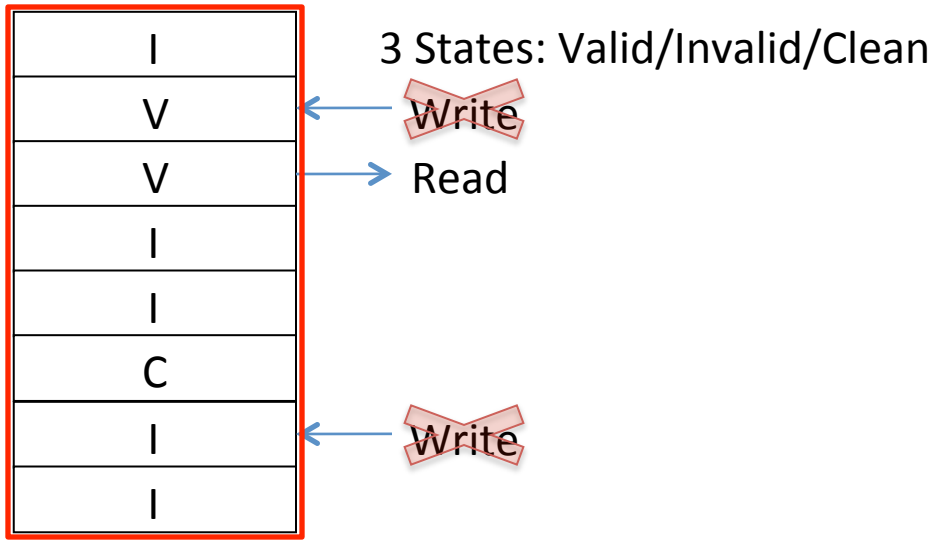




Flash Type	Data Unit Size			Access Time		
	Page (Bytes)		Block (Bytes)	Page READ (us)	Page WRITE (us)	Block ERASE (ms)
	Data	OOB				
Small Block	512	16	(16K+512)	41.75	226.75	2
Large Block	2048	64	(128K+4K)	130.9	405.9	2

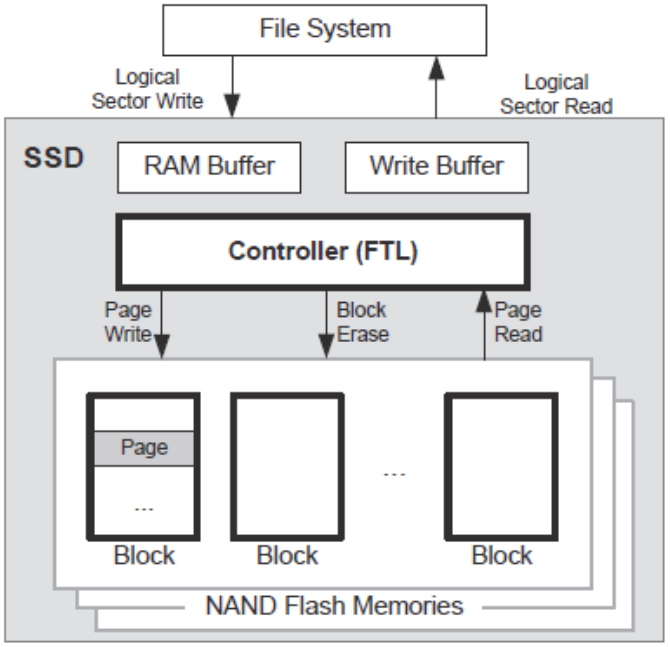
Flash Memory

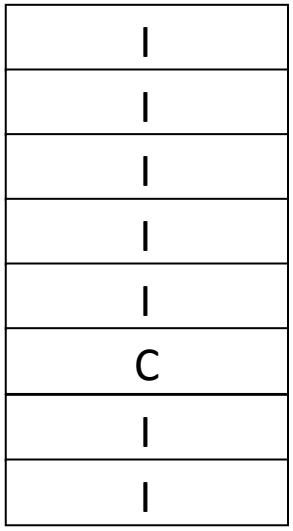




Flash Type	Data Unit Size			Access Time		
	Page (Bytes)		Block (Bytes)	Page READ (us)	Page WRITE (us)	Block ERASE (ms)
	Data	OOB				
Small Block	512	16	(16K+512)	41.75	226.75	2
Large Block	2048	64	(128K+4K)	130.9	405.9	2

Flash Memory

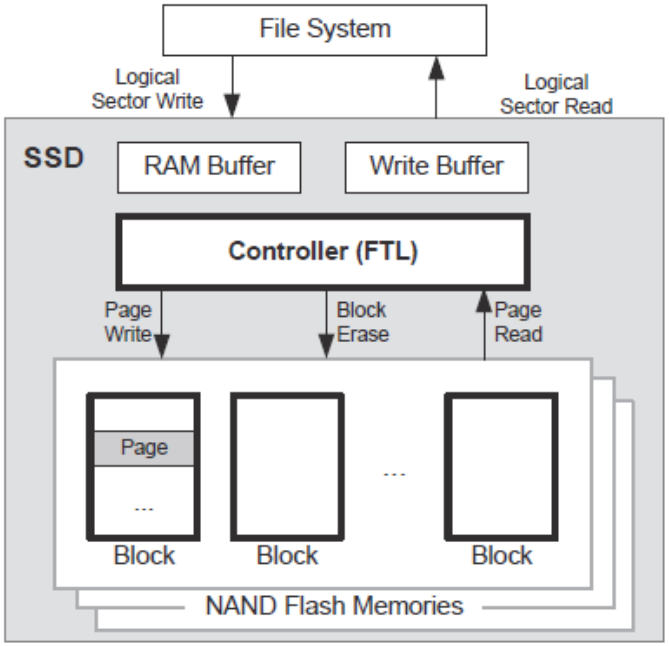




3 States: Valid/Invalid/Clean

Flash Type	Data Unit Size			Access Time		
	Page (Bytes)		Block (Bytes)	Page READ (us)	Page WRITE (us)	Block ERASE (ms)
	Data	OOB				
Small Block	512	16	(16K+512)	41.75	226.75	2
Large Block	2048	64	(128K+4K)	130.9	405.9	2

Flash Memory

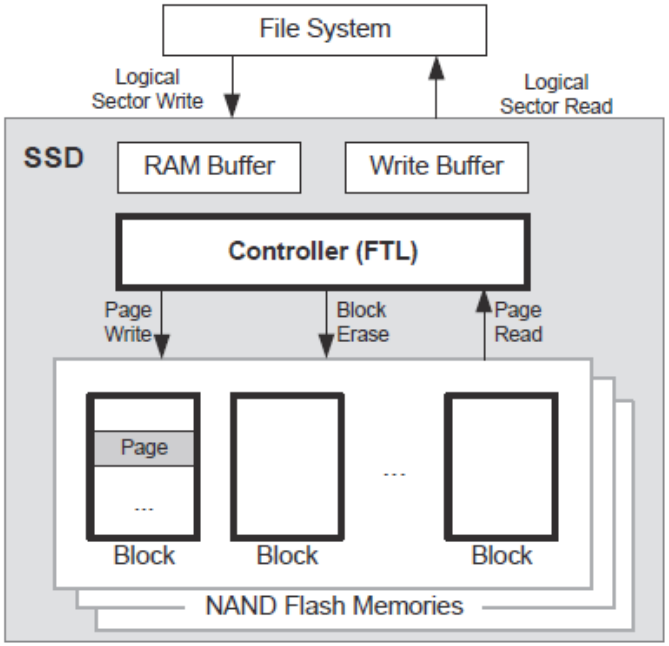


C
C
C
C
C
C
C
C

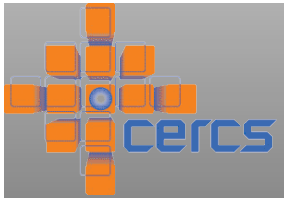
3 States: Valid/Invalid/Clean

Flash Type	Data Unit Size			Access Time		
	Page (Bytes)		Block (Bytes)	Page READ (us)	Page WRITE (us)	Block ERASE (ms)
	Data	OOB				
Small Block	512	16	(16K+512)	41.75	226.75	2
Large Block	2048	64	(128K+4K)	130.9	405.9	2

Flash Memory



- Constraints
  - Read/Write in a PAGE unit
  - Erase in a BLOCK unit
  - No in-place update
    - No write on a page unless it's clean
  - Limited erase count
    - SLC (100,000), MLC (10,000)
    - Wear-leveling
- Flash Translation Layer
  - Mapping table on SRAM
    - Virtual address (exposed to upper level) → physical address (on flash chips)
  - Garbage Collection
    - Log structured mechanism
    - Hide erase operation to upper level

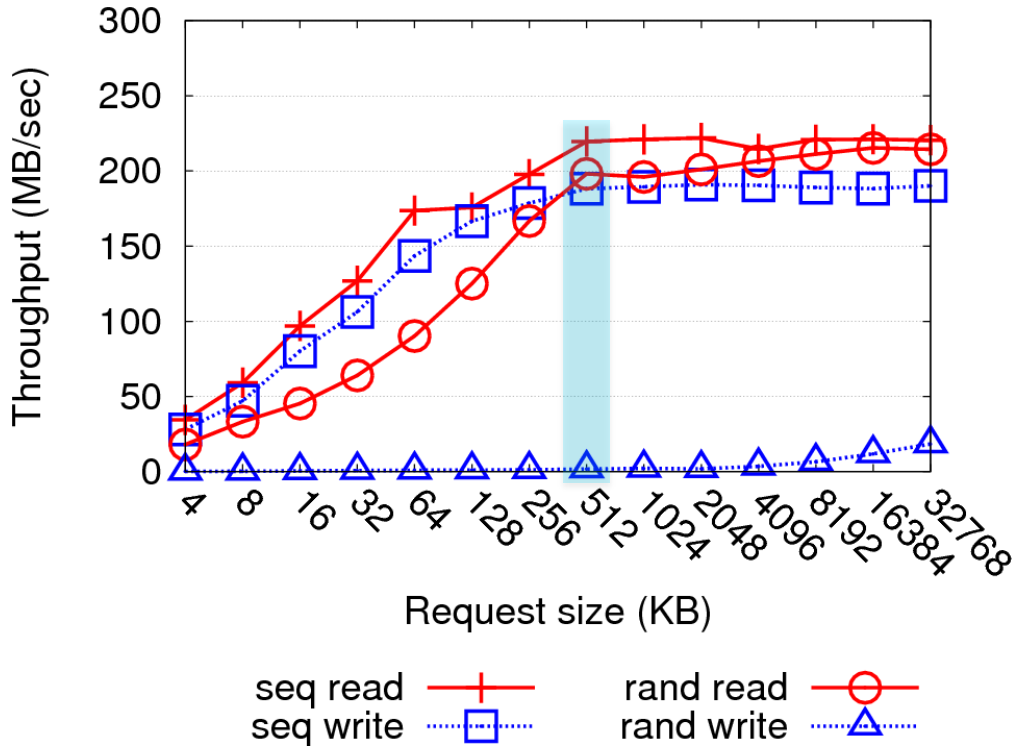


Contents

- Motivation
  - What is the problem?
- Background
  - Interval Caching
  - Flash Memory
- How to utilize Flash SSD?
- What kinds of Flash SSD is appropriate?
- What is the best strategy?
- Conclusion

Flash Memory

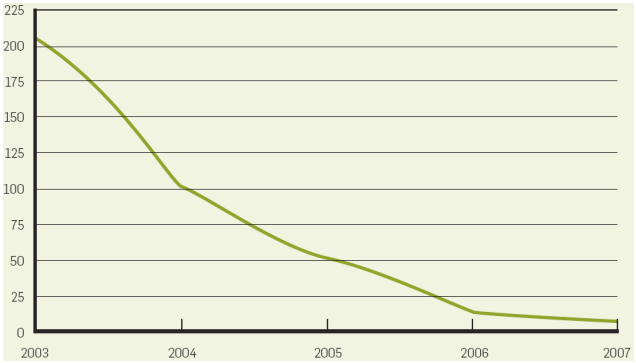
- Fast Random Read and Sequential Write



SAMSUNG MMDOE56G5MXP-0VB MLC 256GB

Flash Memory

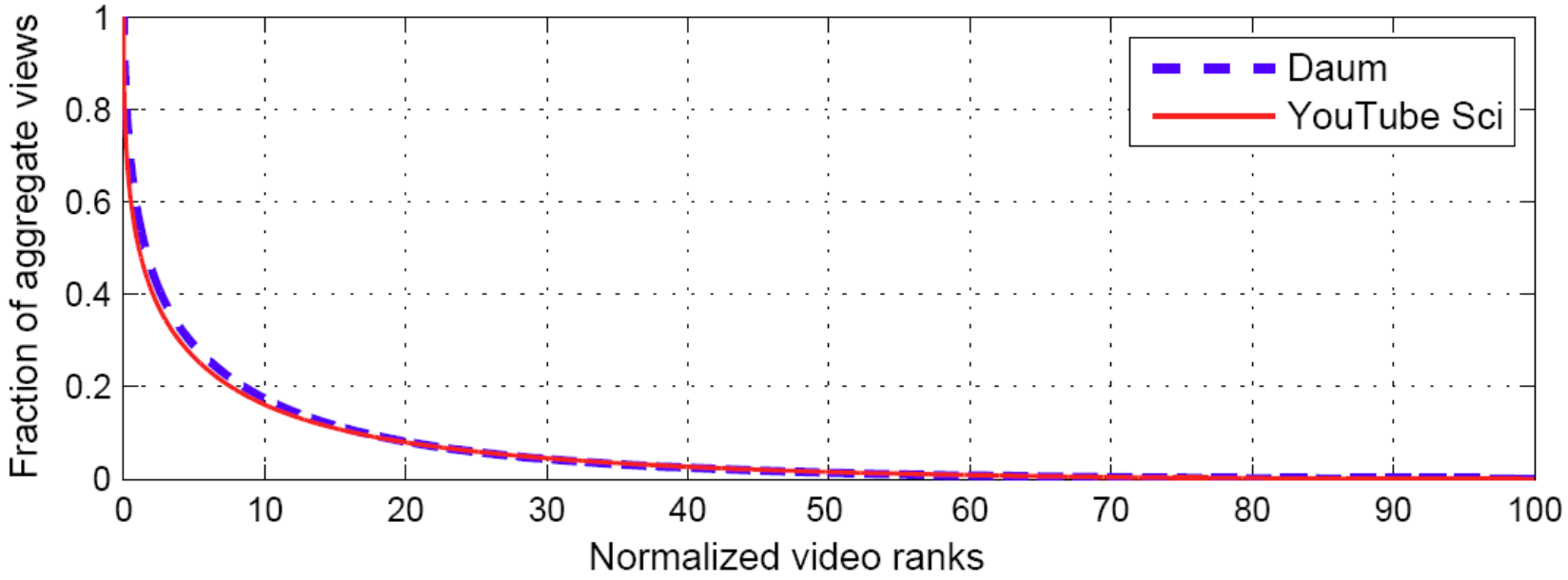
- Cheap MLC Flash Memory



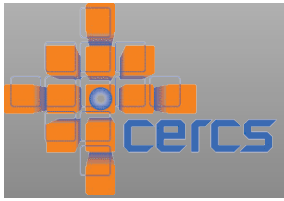
Device	Seq. Read (MB/s)	Rand. Read (MB/s) (I/O Request Size 1MB)	$ / GB
DRAM	> 20480	> 20480	23.00
HDD15K	160	70	1.23
MLC SSD	> 155	> 155	1.88

Flash Memory

- Highly skewed video access pattern



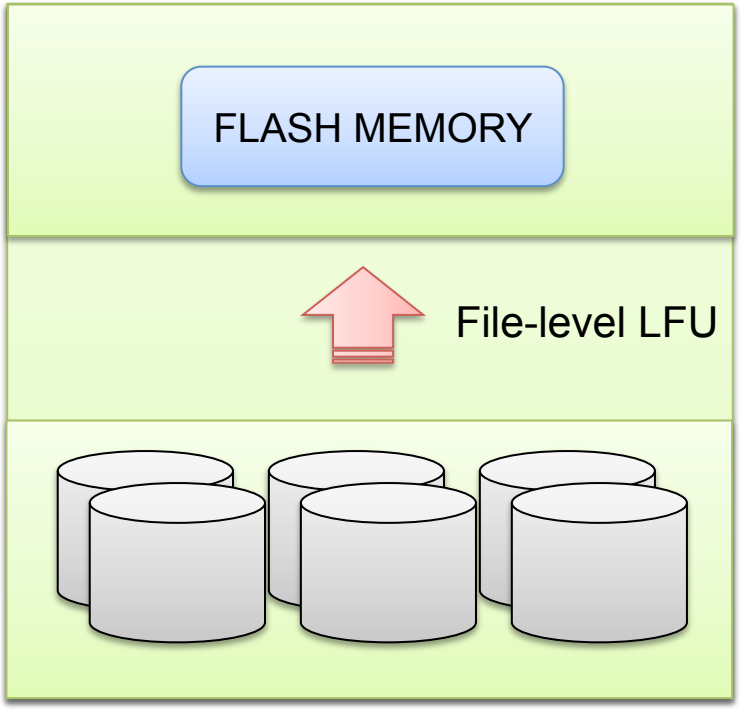
Reference: Cha et al., "I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System", In Proc. of IMC, 2007



Flash Memory

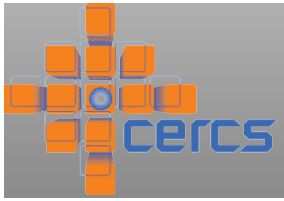
- Reasons to go for Flash Memory
  - Fast Random Read
  - Fast Sequential Write
  - Cheap MLC Flash Memory
  - Skewed video access pattern
  - Very Low Power
- Problems
  - Higher cost/GB than HDDs
  - Poor small random write
  - Unpredictable garbage collection
  - Limited life time

How to utilize Flash Memory?



- Guarantee cache hit
- Easy to guarantee continuous delivery

- Flash Memory for B/W and Capacity
- HDD for Capacity
- Power reduction

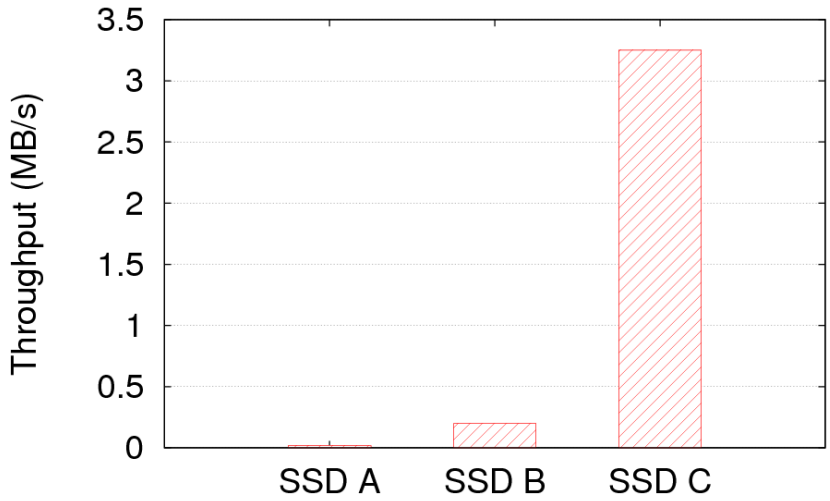


Contents

- Motivation
  - What is the problem?
- Background
  - Interval Caching
  - Flash Memory
- How to utilize Flash SSD?
- What kinds of Flash SSD is appropriate?
- What is the best strategy?
- Conclusion

Which Flash Memory SSD?

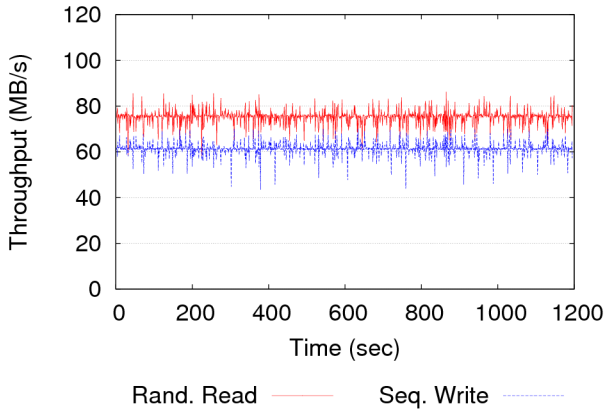
	A	B	C	D
MODEL	RiData NSSD-S25-64-C06MPN	SAMSUNG MMDOE56G5MXP-0VB	INTEL X25-M G1	Fusion-io ioDRIVE
CLASS	LOW	LOW	HIGH	ENTERPRISE
TYPE	MLC	MLC	MLC	SLC
CAPACITY (GB)	64	256	80	80
SEQ. READ B/W (MB/s)	155	220	250	750
$/GB	1.88	2.34	2.50	61.95



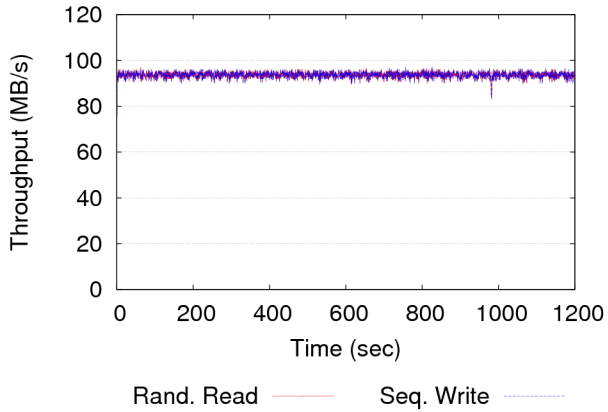
Random Write throughput when request size is 4KB.

Which Flash Memory SSD?

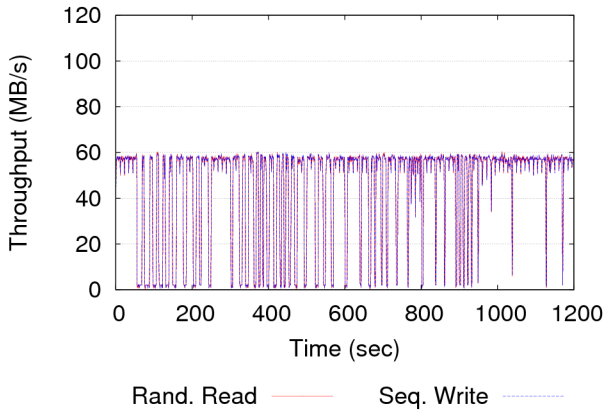
Is High-end SSD good for the architecture? NO!



SSD A

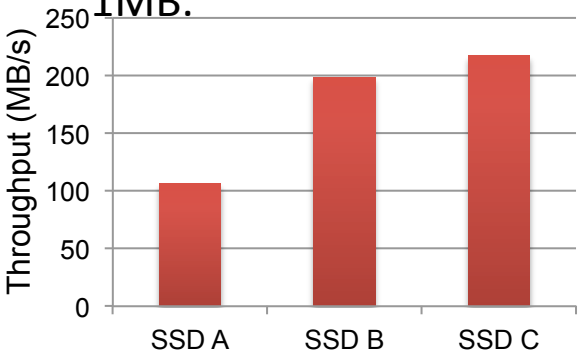


SSD B

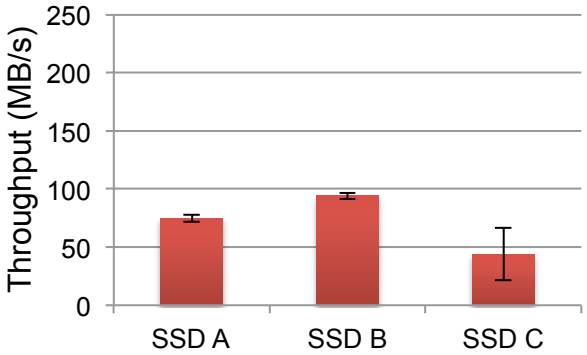


SSD C

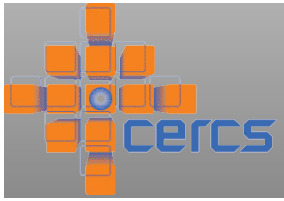
Random Read and Sequential Write operations are mixed by 1:1. Request size is 1MB.



Average Random Read throughput without write operations



Average and Standard deviation of Random Read throughput with write operations

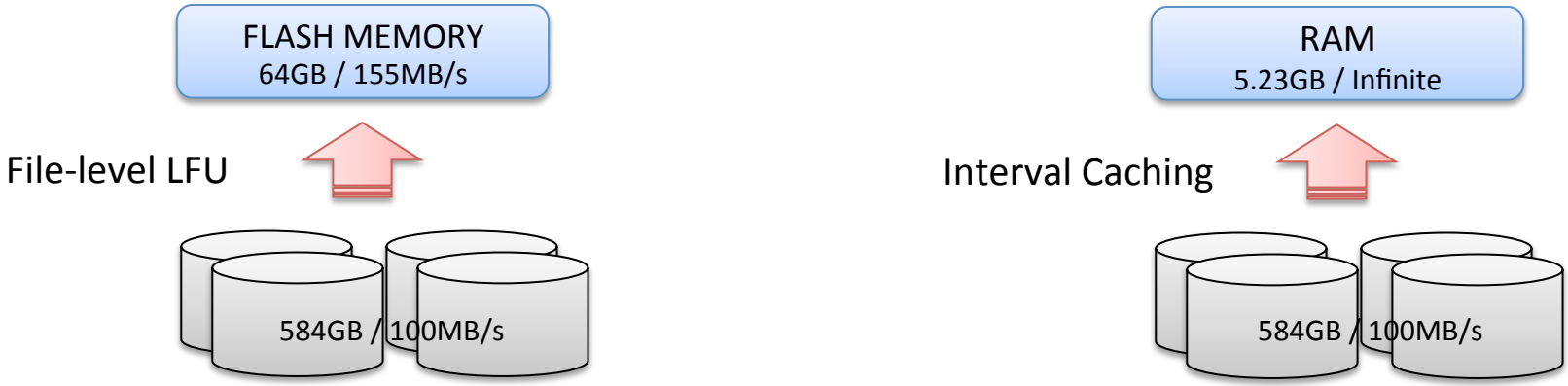


Contents

- Motivation
  - What is the problem?
- Background
  - Interval Caching
  - Flash Memory
- How to utilize Flash SSD?
- What kinds of Flash SSD is appropriate?
- What is the best strategy?
- Conclusion

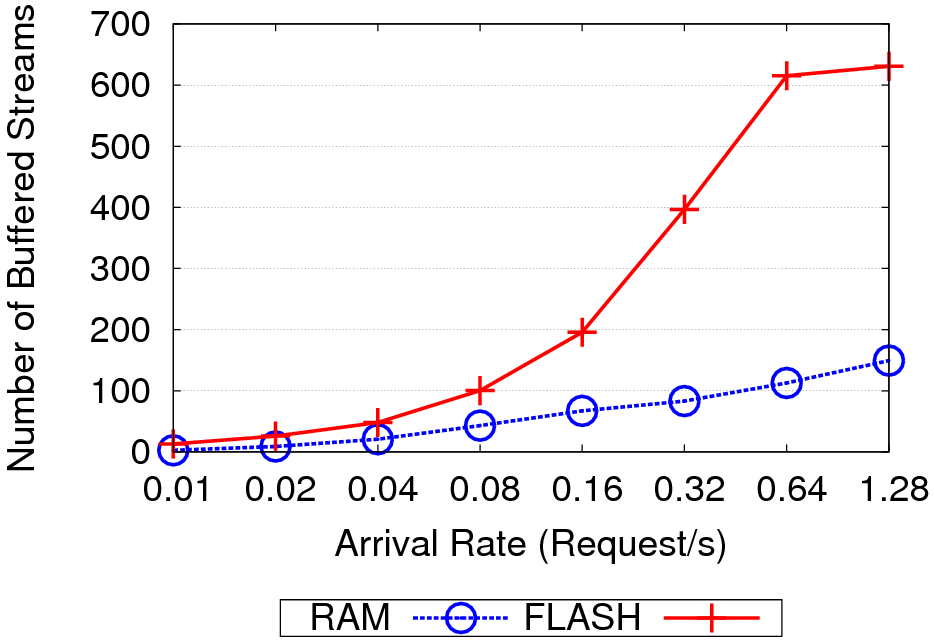
File-level LFU v.s. Interval Caching

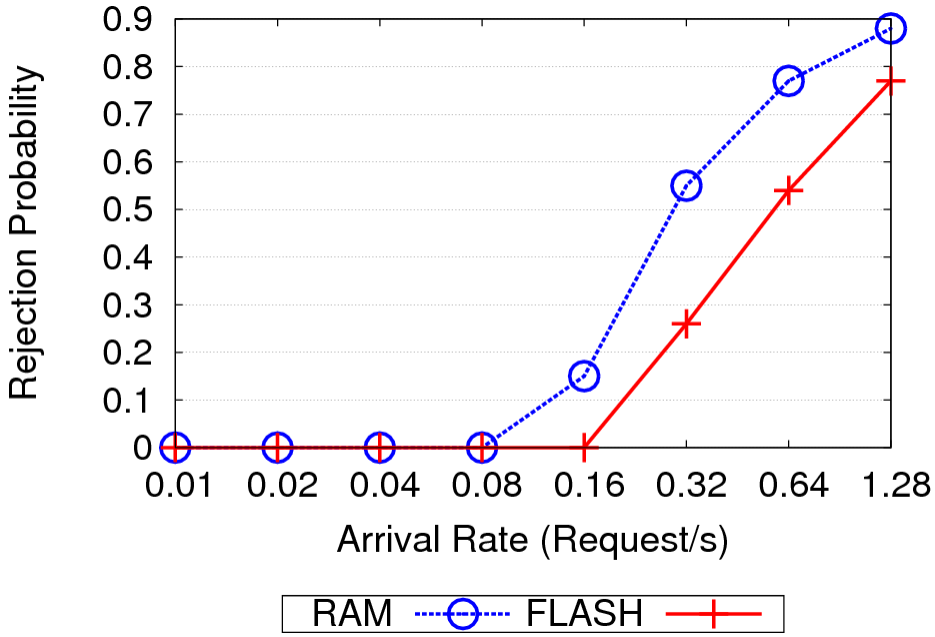
- Assumptions
  - 300 videos in total
  - Homogeneous constant bitrate 245KB/s
  - Homogeneous video length
  - Popular videos are cached in advance
- Control Parameters
  - Arrival rate (req/sec)
  - Video popularity (Zipf distribution)
  - Video length
- Performance metrics
  - Number of buffered streams
  - Hit ratio
  - Rejection probability



File-level LFU v.s. Interval Caching

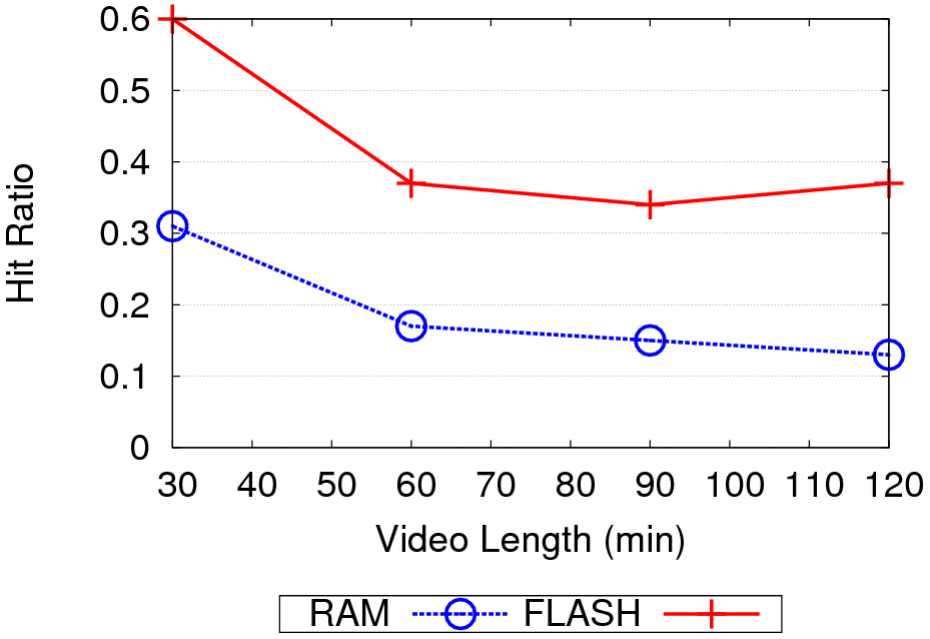
Zipf 0.271, Video Length 60mins

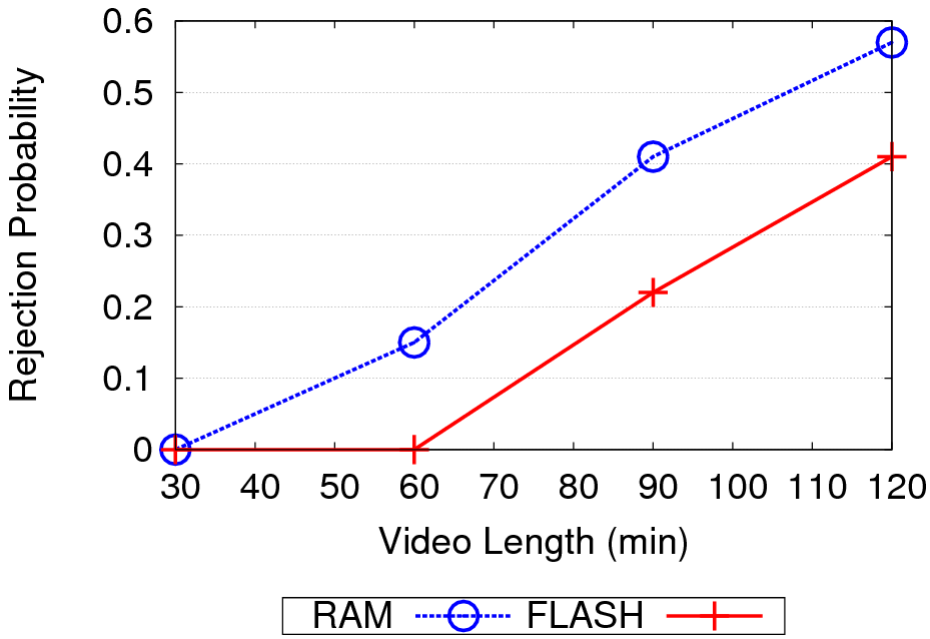




File-level LFU v.s. Interval Caching

Zipf 0.271, Arrival Rate 0.16





What is the best strategy?

- Strategies
  - HDD only
  - HDD + DRAM (Interval Caching)
  - HDD + MLC Flash SSD (File-level LFU)
- Total cost of ownership =
  total cost of acquisition +
  total cost of operation

What is the best strategy?

- Simulation Parameters

Video Library	300
Zipf (Video Popularity)	0.271
Video Bitrate	245 KB/s
Video Length	60 min
Arrival Rate	1 req / sec
Rejection Probability	0 %

3600 concurrent streams

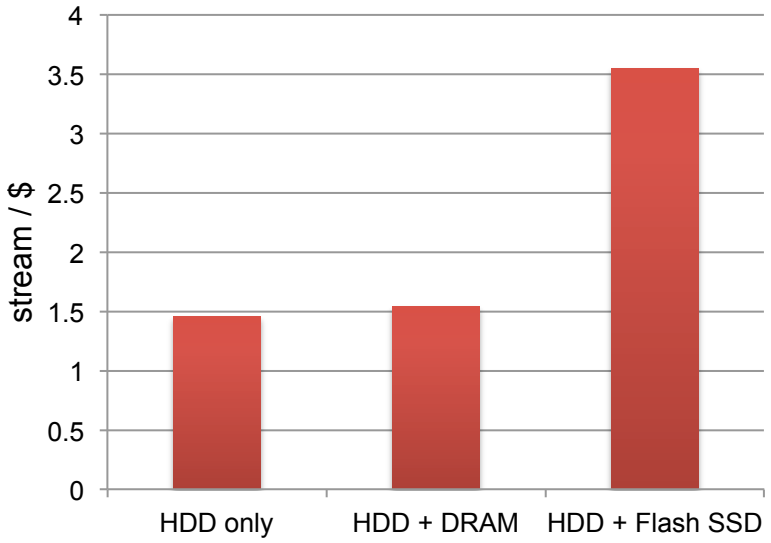
- Device profile

Device	Capacity (GB)	Rand. Read (MB/s) (I/O Request Size 1MB)	$
DRAM	1	∞	23
HDD15K	146	70	180
MLC SSD	64	155	120

What is the best strategy?

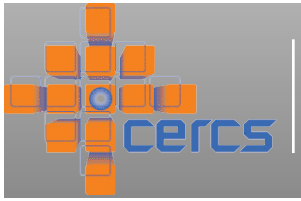
- Minimum cost configuration

Strategy	Configuration	Streams / $
HDD only	13	1.46
HDD + DRAM (Interval Caching)	2 + 12	1.54
HDD + Flash SSD (File-level LFU)	5 + 2	3.55



Conclusion

- NAND Flash Memory is promising
- Questions
  - How to utilize Flash SSD?
    - Avoid random write operations
    - File-level LFU
  - What kinds of Flash SSD is appropriate?
    - Low-end MLC Flash SSD is better
  - What is the best strategy?
    - HDD+MLC Flash SSD



Thank You!