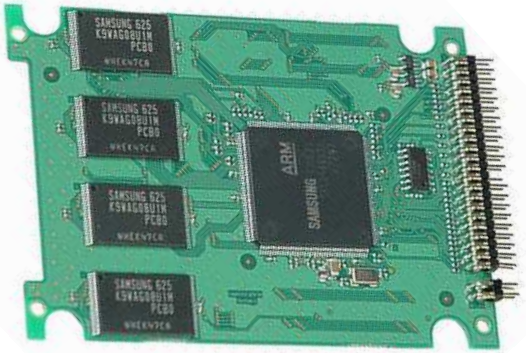
Hybrid Solid State Disks: Combining Heterogeneous NAND Flash in Large SSDs

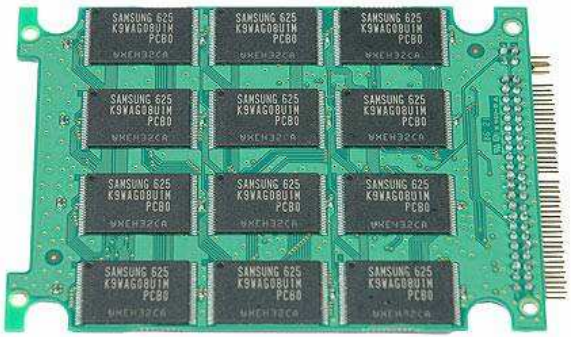
Prof. Li-Pin Chang　張立平
ESSLab @ NCTU
National Chiao-Tung University, Taiwan
ASPDAC 2008

Introduction

- SSDs aim at replacing hard drives in embedded computers
  - The host can not distinguish a hard drive from an SSD
- An SSD comprises an MCU, an optional volatile memory, and an array of NAND flash



Front side



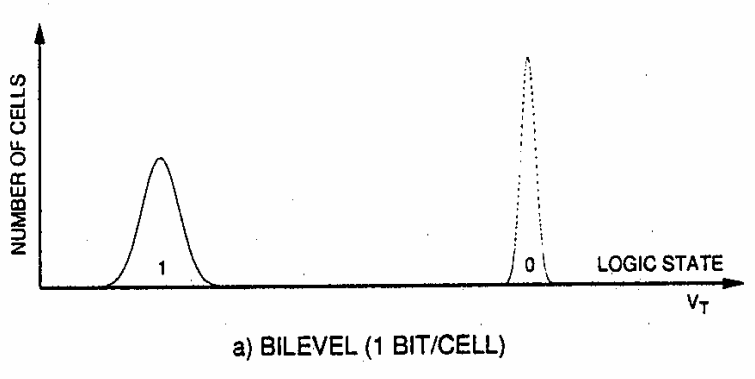
Back side

Introduction

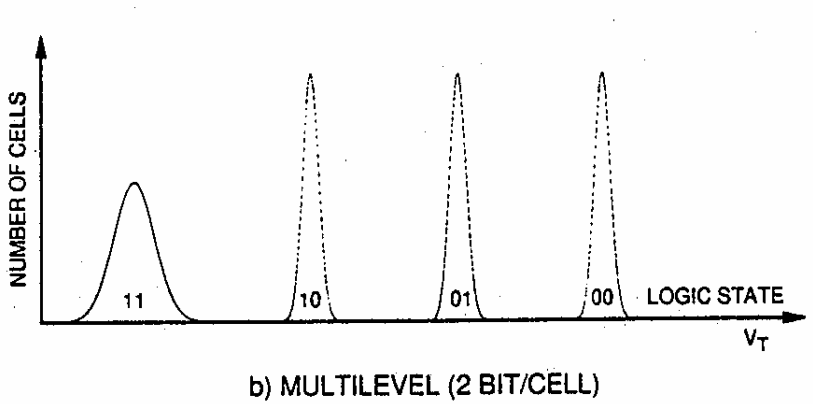
- Competitors
  - Micro-drives vs. SSDs
- It turns out…
  - Recently Hitachi has announced that it will phase out the production of 1.0, 1.5, and 1.8-inch drives
- However,
  - NAND flash is still expensive!

Introduction

- An SLC cell stores a binary value
- A 2BC MLC cell stores information of 2 bits by using 4 voltage thresholds
  - Information density is doubled!



An SLC-flash cell



An MLC-flash cell

Introduction

- SLC flash : MLC flash
  - Two times faster write
  - Lower power consumption
  - 10~20 times longer endurance
  - Finer geometry (page size, block size)
  - Two times more expensive!

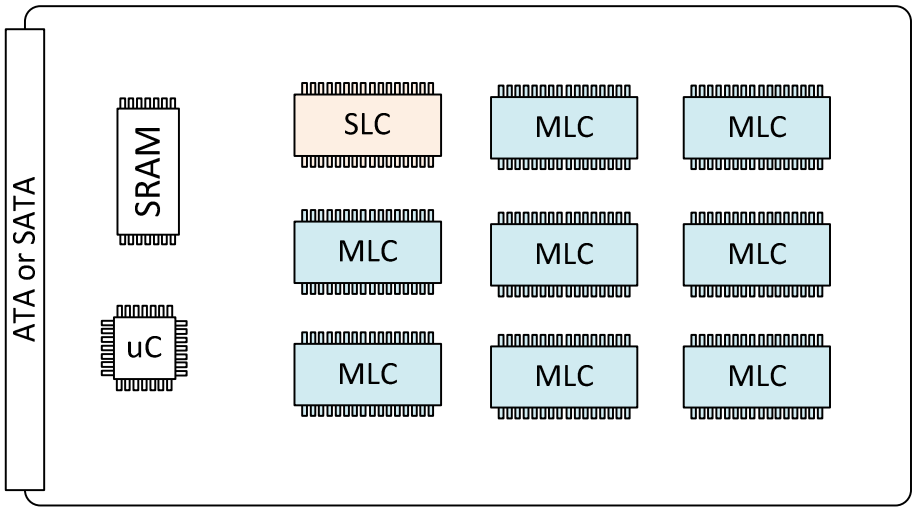
Introduction

- Observations
  - SLC flash is fast and durable
  - MLC flash is large and cheap
  - SLC flash and MLC flash are pin-compatible

- Idea
  - SLC flash and MLC flash are complements to each other!

Hybrid Solid-State Disks

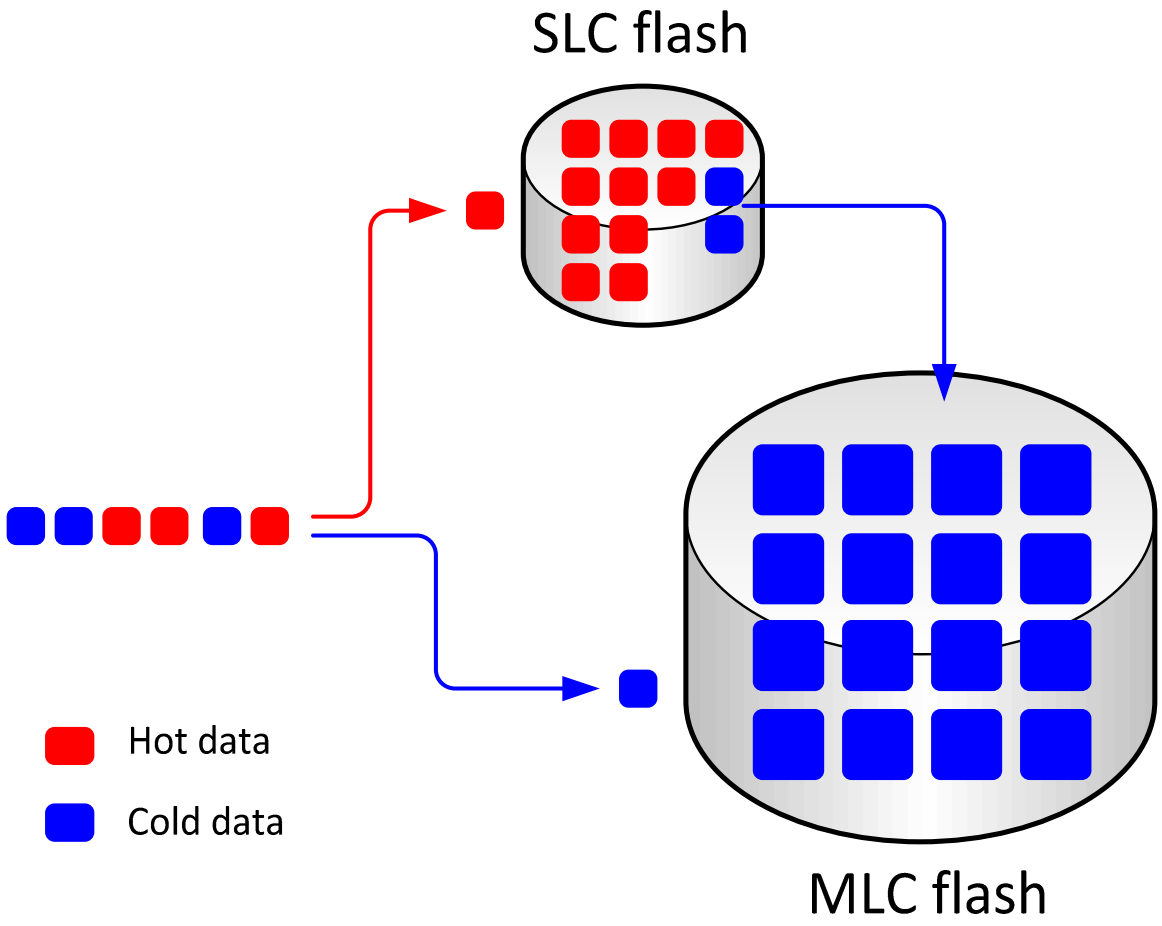
- A hybrid solid-state disk
  - Using SLC flash and MLC flash together
  - SLC flash and MLC flash are pin-compatible



Hybrid Solid-State Disks

- Technical issues
  - Data placement
    - Data of what attribute should be write to SLC?
  - Data-attribute identification
    - How to do on-line data attribute identification?
  - Data migration
    - How to phase out data from SLC?
  - Wearing balance
    - How to balance the wearing of SLC and MLC?

Hybrid Solid-State Disks

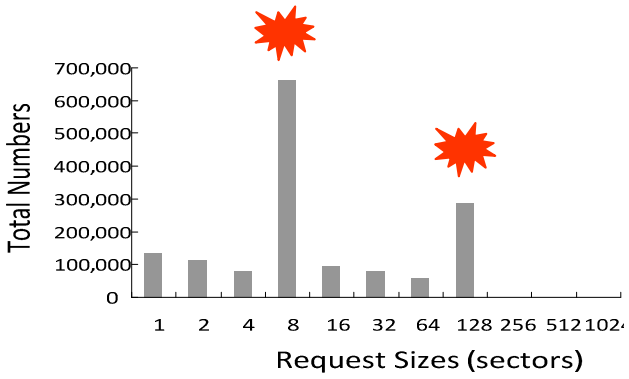


Hybrid Solid-State Disks

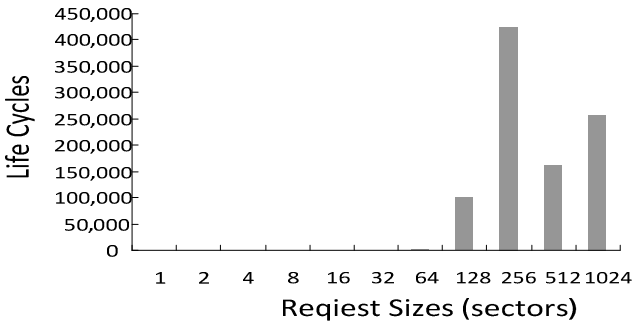
- Data placement
  - SLC flash
    - is expensive
    - has small pages/blocks
    - is fast on write
    - endures more erasure operations
  - Small writes to hot data can be serviced by SLC flash

Hybrid Solid-State Disks

- Data-attribute identification
  - By analyzing a one-month disk traces collected from a UMPC, we found that
    - Many small writes goes to hot data
    - Many large writes goes to cold data
    - The distribution of write sizes is bimodal



(a) Total numbers of writes

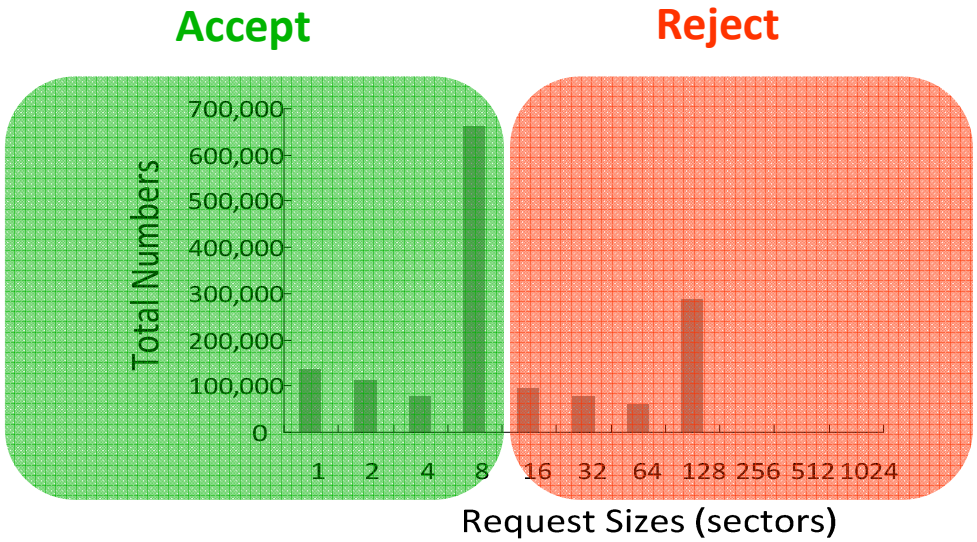


(b) Data life cycles

Hybrid Solid-State Disks

- Hot-data filter
  - Periodically invoking a 2-means clustering algorithm to identify the two peaks
  - Needless to keep track of data attributes actually, the SLC accepts simply small writes and rejects the others

Because hot data are of short life cycles, the SLC flash needs not be large!



Hybrid Solid-State Disks

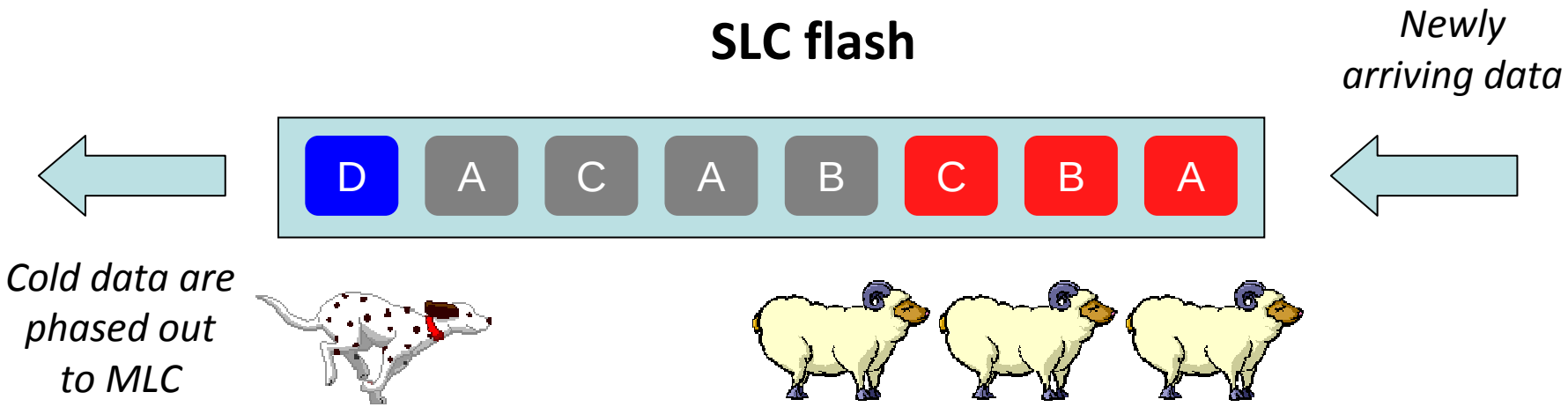
- Data migration
  - Hot data may become cold
  - To leave free space for hot data, cold data should be phased out from SLC flash
    - Copy cold data to MLC flash
  - Challenges
    - Cold data hide in SLC flash must be identified
    - Too costly to examine every piece of data

Hybrid Solid-State Disks

- Data migration
  - SLC flash itself needs garbage collection
    - Garbage collection involves data copy
  - Is it possible to combine the overheads of data migration and garbage collection??
    - Yes!

Hybrid Solid-State Disks

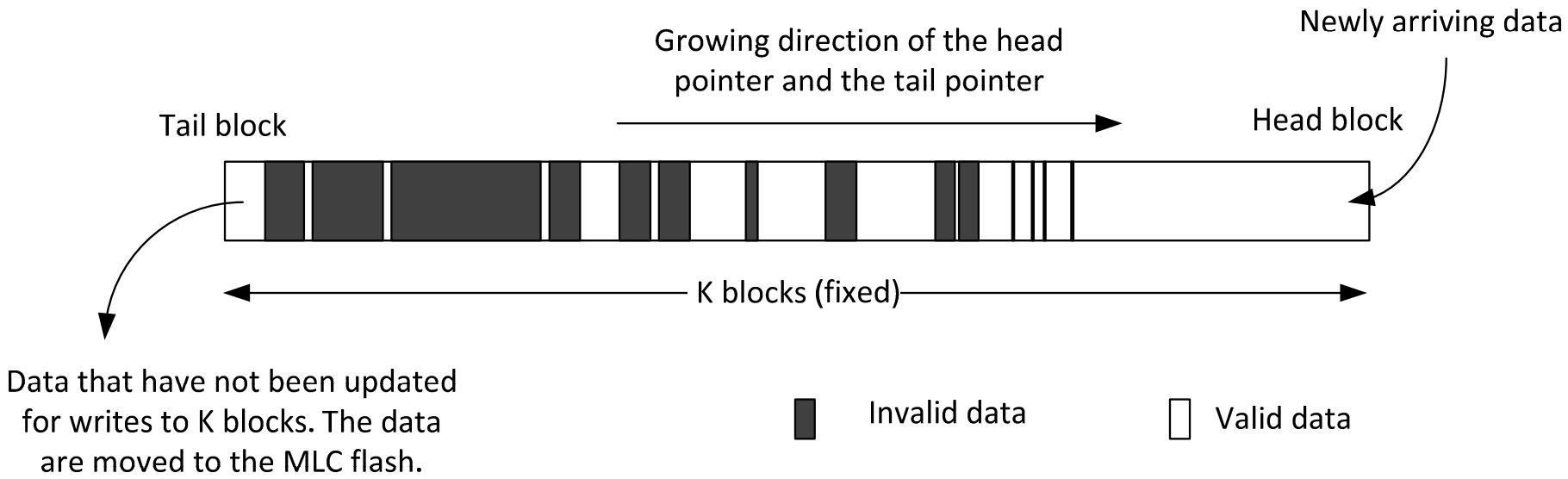
- Data migration
  - If we treat the SLC as a circular logging space…



- Data found valid in the tail block is cold
- Hot data appear close to the head block

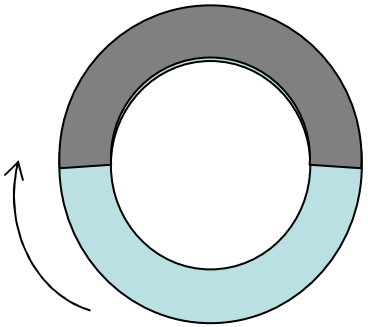
Hybrid Solid-State Disks

- Data migration



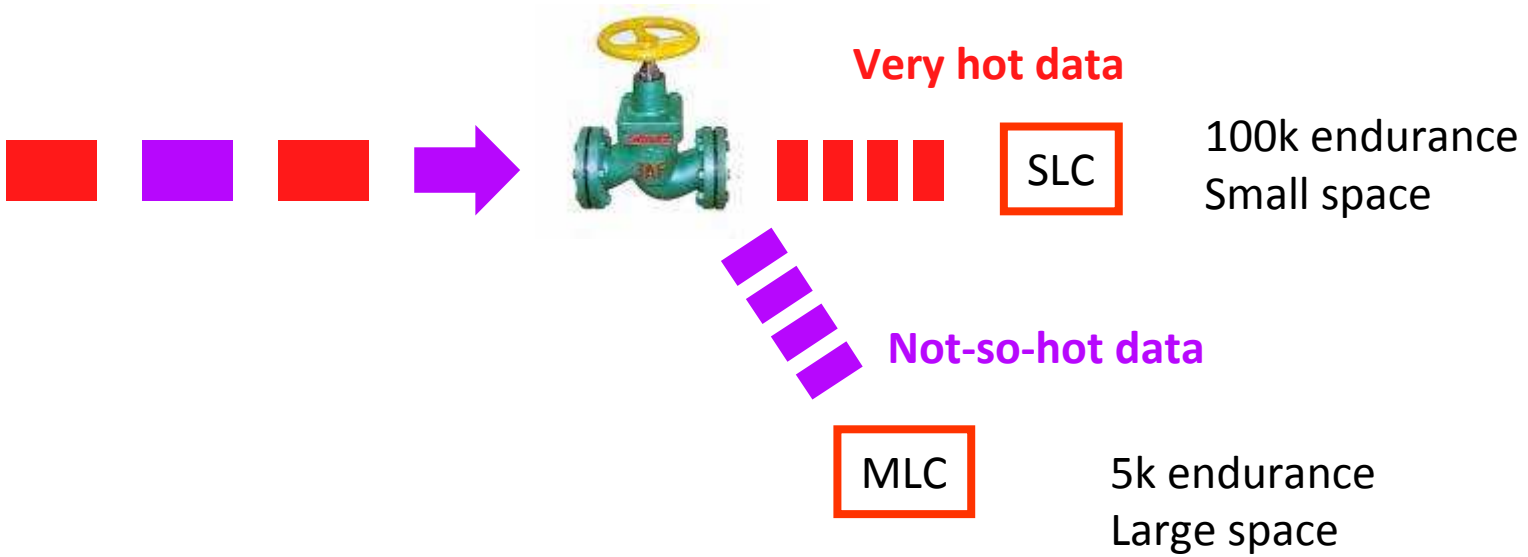
What even better is…

- **Wear leveling over SLC is perfect!**



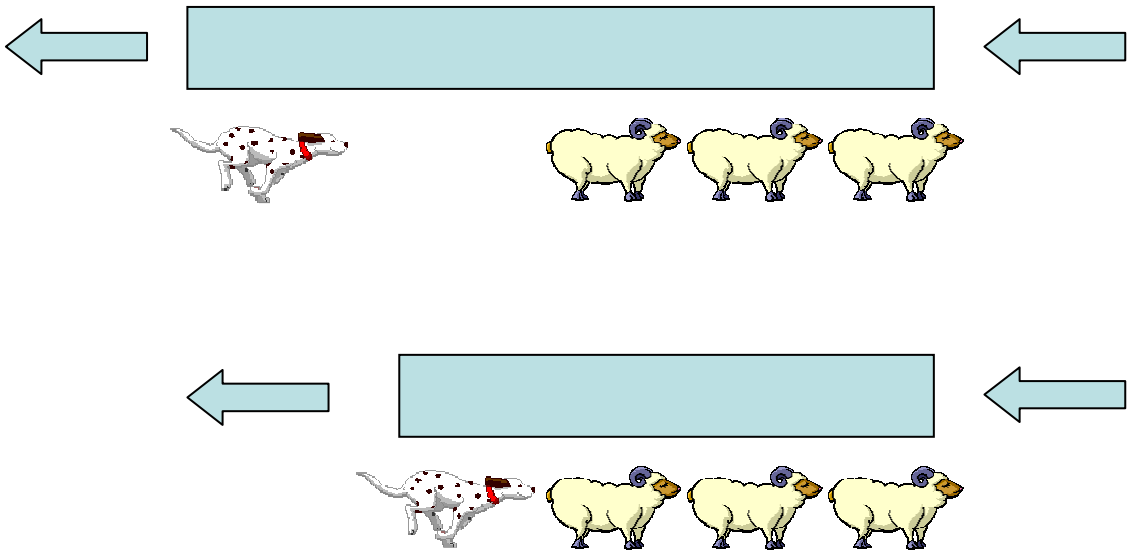
Hybrid Solid-State Disks

- SLC flash and MLC flash have different block endurance
  - SLC is 100k while is MLC 5k
  - When the SLC is overly worn, retrench the data flow into SLC



Hybrid Solid-State Disks

- To retrench the utilization throttle,
  - aggressively phasing out data by decreasing “k”
  - SLC accepts only updates to existing data

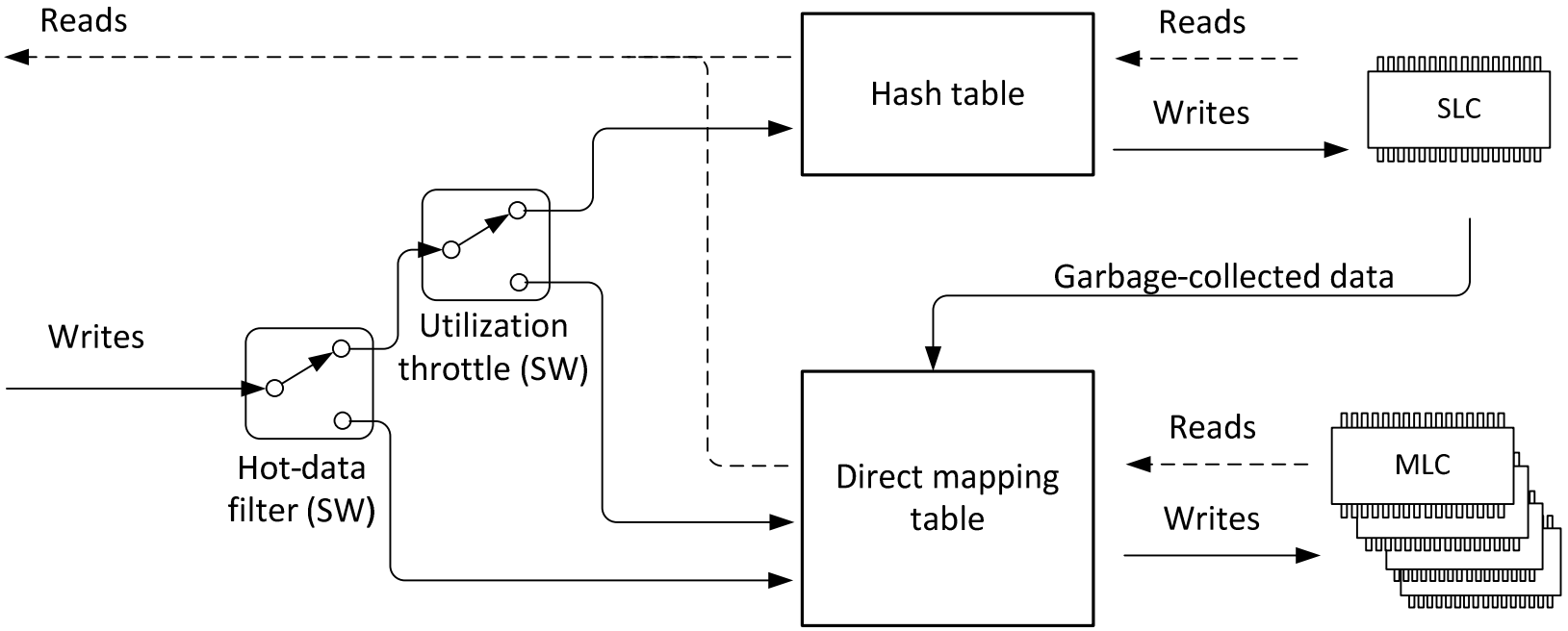


Hybrid Solid-State Disks

- Managing the MLC flash
  - MLC flash is the final repository of data
  - MLC flash is managed as if SLC flash does not exist
  - Existing techniques, such as NFTL, are considered

Hybrid Solid-State Disks

- Putting things together…



Performance Evaluation

- Workload
  - Disk traces collected from a UMPC
  - One-month real-life use
- Primary performance metrics
  - Response ratio (RS)
  - Throughput ratio (TS)

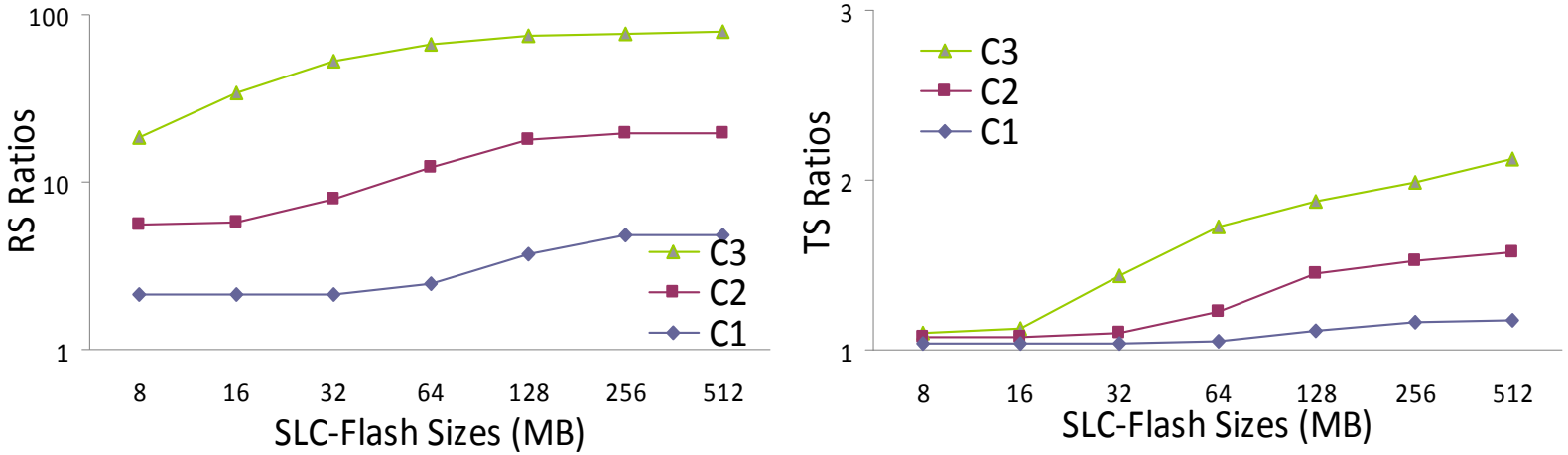
Performance Evaluation

- Let triple (x,y,z) denote

(page size, block size, block endurance)

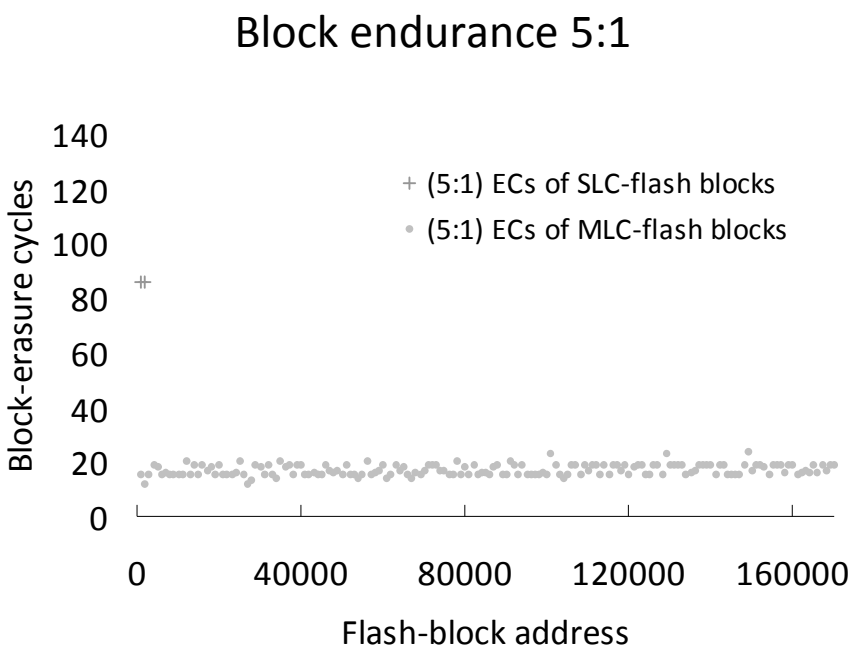
- Three configurations were evaluated
  - C1: SLC (2k,128k,100k)+MLC (2k,128k,10k)
  - C2: SLC (2k,128k,100k)+MLC (4k,256k,10k)
  - C3: SLC (2k,128k,100k)+MLC (4k,512k,10k)

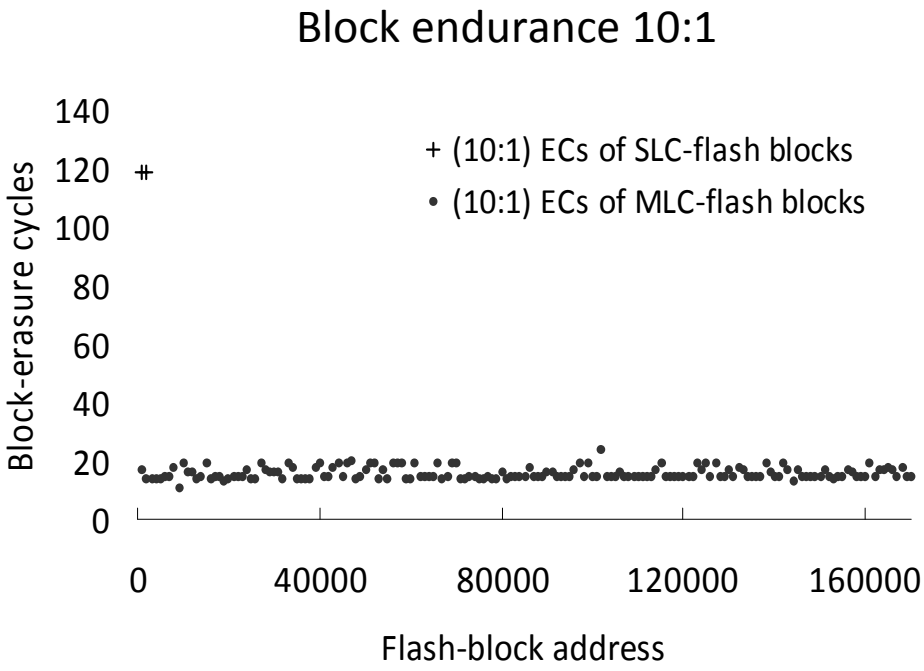
Performance Evaluation



- Response is greatly improved
  - The majority of requests go to hot data
- Throughput is slightly improved
  - The SLC flash is of little use for bulk/sequential writes
- Using a 128-MB SLC flash is a good choice

Performance Evaluation





- Wear-leveling for MLC uses the dual-pool algorithm (previously proposed in ACM SAC 2007)
- The wearing of the SLC flash is properly controlled

Conclusion

- The price issue largely concerns the success of SSD in the recent years
- To combine heterogeneous NAND flash in SSD is to achieve a good balance between price and performance
- Our next step is to investigate the use of new non-volatile memory such as MRAM and PCRAM in hybrid SSDs

Thank you!

Questions?