NAND Flash-based Disk Cache Using SLC/MLC Combined Flash Memory

Seongcheol Hong and Dongkun Shin
Sungkyunkwan University, Korea
{adonis0101,dongkun}@skku.edu

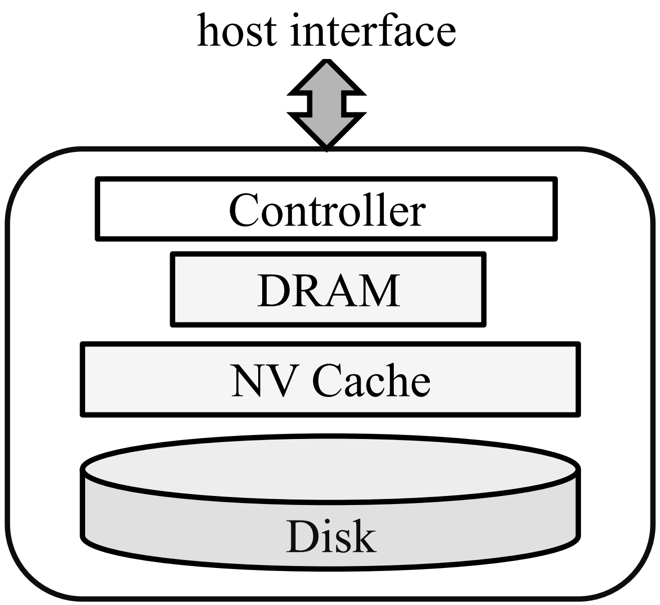
SNAPI 2010

Introduction♪

- Objective
  - Propose an effective management scheme for SLC/MLC comb ined flash memory
  - Determine the optimal proportion between the two regions
    - Maximize performance and energy reduction
    - Guarantee the lifespan constraint
- Keyword
  - Hybrid HDD
  - SLC/MLC Combined Flash Memory

Background♪

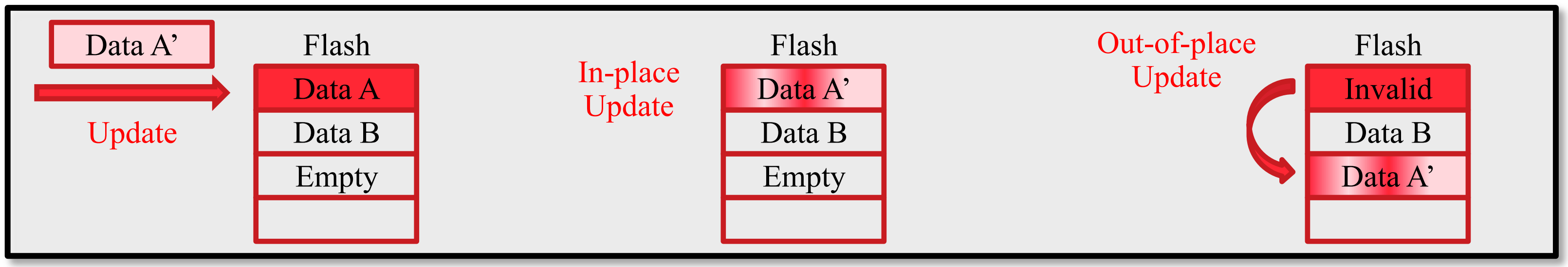
- Hybrid HDD



- NV Cache stores data blocks that are likely to be accessed in t he near future.
  - Faster I/O performance
  - Higher energy efficiency

Background♪

- Flash memory
  - A non-volatile memory able to be electrically erased and prog rammed.
  - Advantages
    - Cheaper than DRAM
    - Faster than HDD
  - Disadvantages
    - Limited P/E cycle
    - Write a "page", but erase a "block"
    - Only the "out-of-place" update is possible



Background♪

- Two types of flash memories
  - SLC(single-level-cell)
    - More reliable
    - Longer lifespan
    - Faster than MLC
  - MLC(multi-level-cell)
    - Cheaper than SLC(larger storage capacity)

- If the NVC of hybrid HDD is
  - MLC flash memory,
    - High performance and low energy consumption
    - But, short lifespan
  - SLC flash memory,
    - Long lifespan
    - But, low performance and high energy consumption

Background♪

- SLC/MLC Combined Flash Memory
  - It has both SLC blocks and MLC blocks in a single chip.
    - By programming only the LSB of a cell in the MLC flash m emory, the cell can be used as an SLC
  - The flash memory blocks can be divided into two regions for ease of management.
    - Depending on the size of each region, the total storage ca pacity of the flash memory is determined.
  - Samsung semiconductor released SLC/MLC combined flash m emory, Flex-OneNAND, in 2007

	SLC	MLC
Page size	4KB	
Block size	256KB (64pages)	512KB (128pages)
Page read	45us	50us
Page write	240us	1ms
Block erase	500us	
P/E cycles	50K	10K

Background♪

- eg) 1024 blocks in a flash memory chip
  SLC block size : 256KB
  MLC block size : 512KB

SLC (512 blocks)	MLC (512 blocks)

  – 256KB X 512 blocks + 512KB X 512 blocks = 384MB

SLC (256 blocks)	MLC (768 blocks)

  – 256KB X 256 blocks + 512KB X 768 blocks = 448MB

Overall Architecture♪

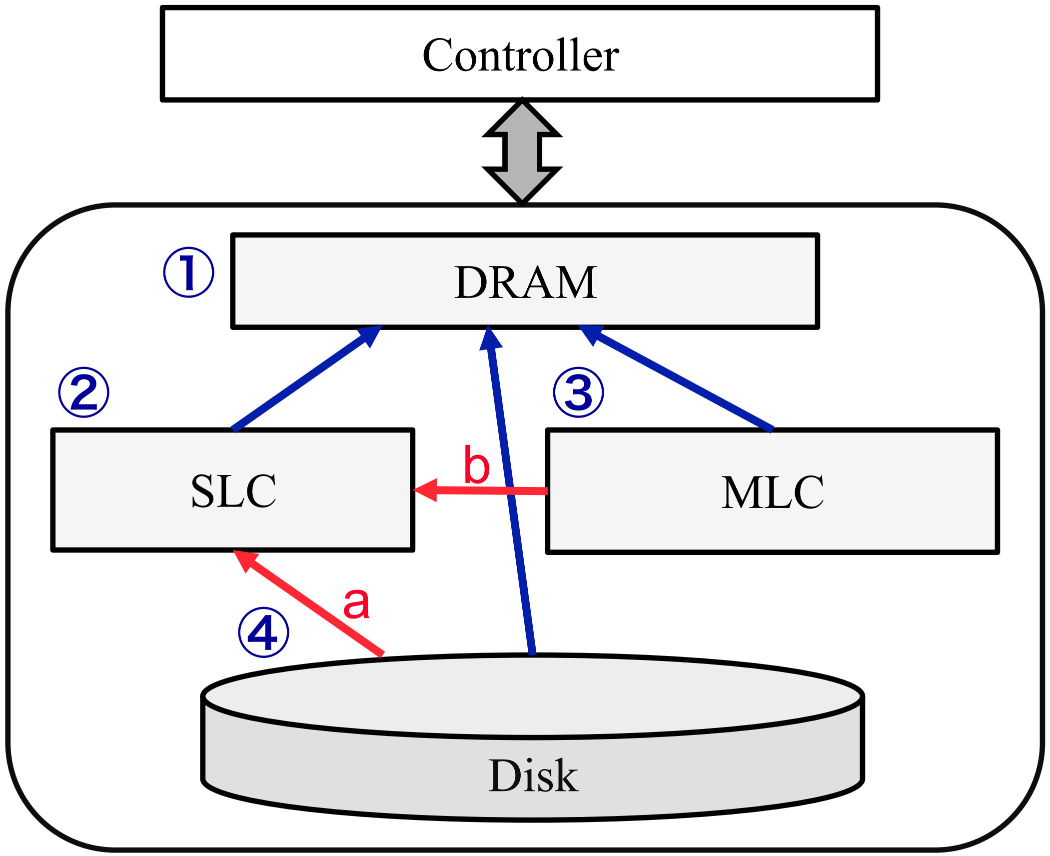
a. Copy from HDD

b. Copy from MLC

For read requests, the system searches the data in the DRAM buffer, NVC, and hard disk

If the data is found in the flash memory or hard disk, it is copied into the DRAM buffer.

To minimize DRAM read miss penalty, when the data is found in the MLC region or the hard disk, it is also copied into the SLC region.

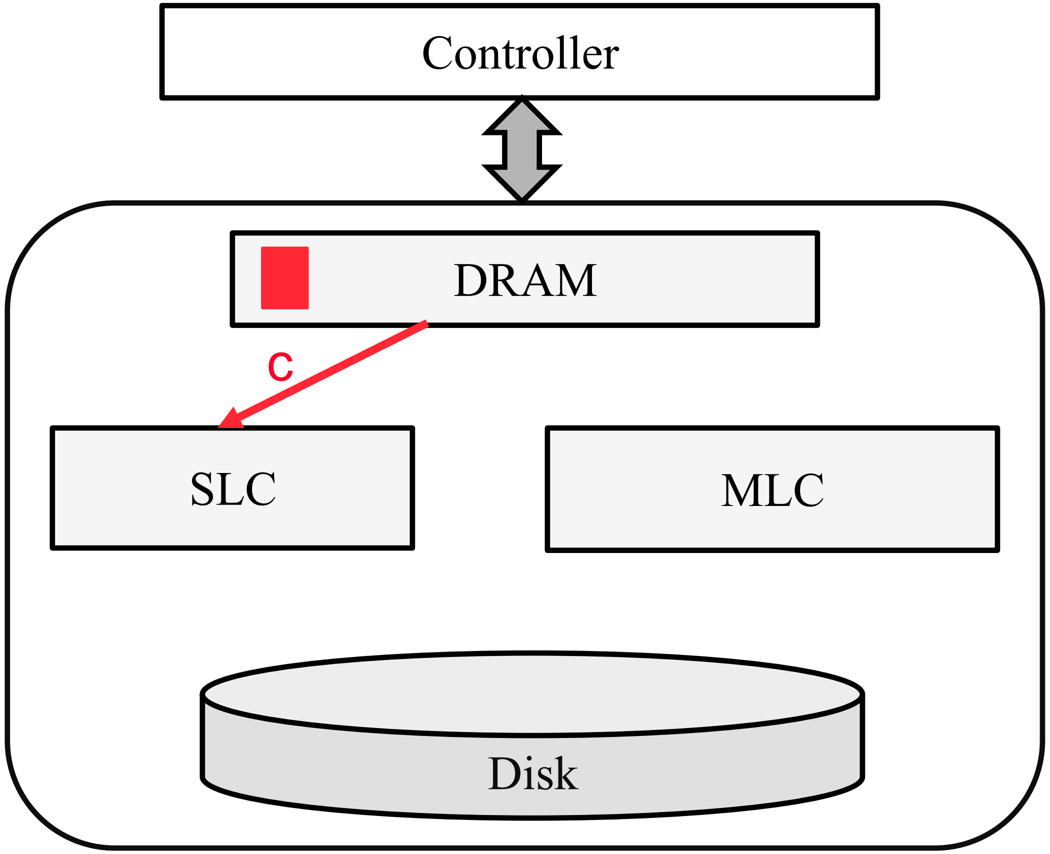


Overall Architecture♪

c. Evict to SLC

For write requests from the host, all data is first written to the DRAM buffer.

And it is sent to the SLC region by a replacement policy of the DRAM buffer.



Overall Architecture

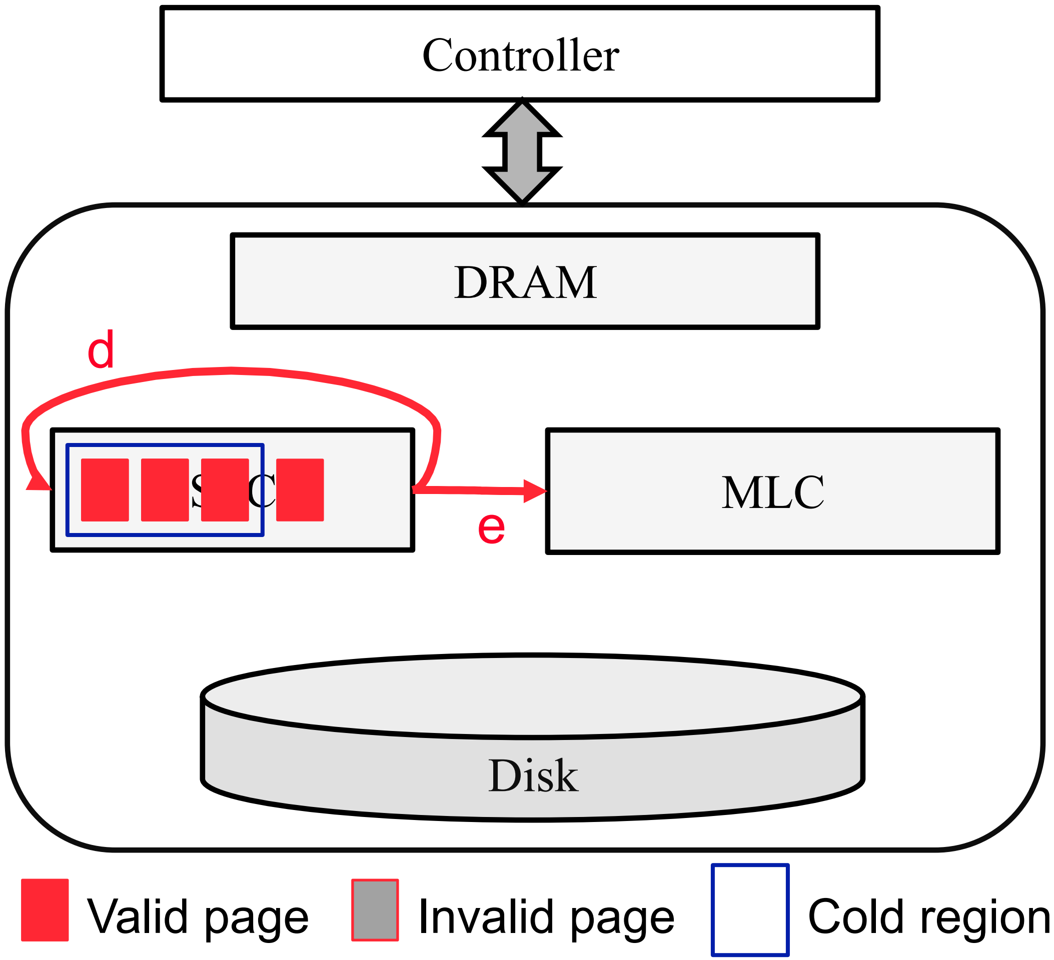
d. GC in SLC

e. Migration to MLC

If the SLC region needs more free space, the garbage collection is invoked.

If the cold region of SLC has sufficient invalid pages(>=70%), the valid pages are moved into other blocks within SLC region.

However, if there are only a few invalid pages (<40%), they migrate to the MLC region.



Overall Architecture♪

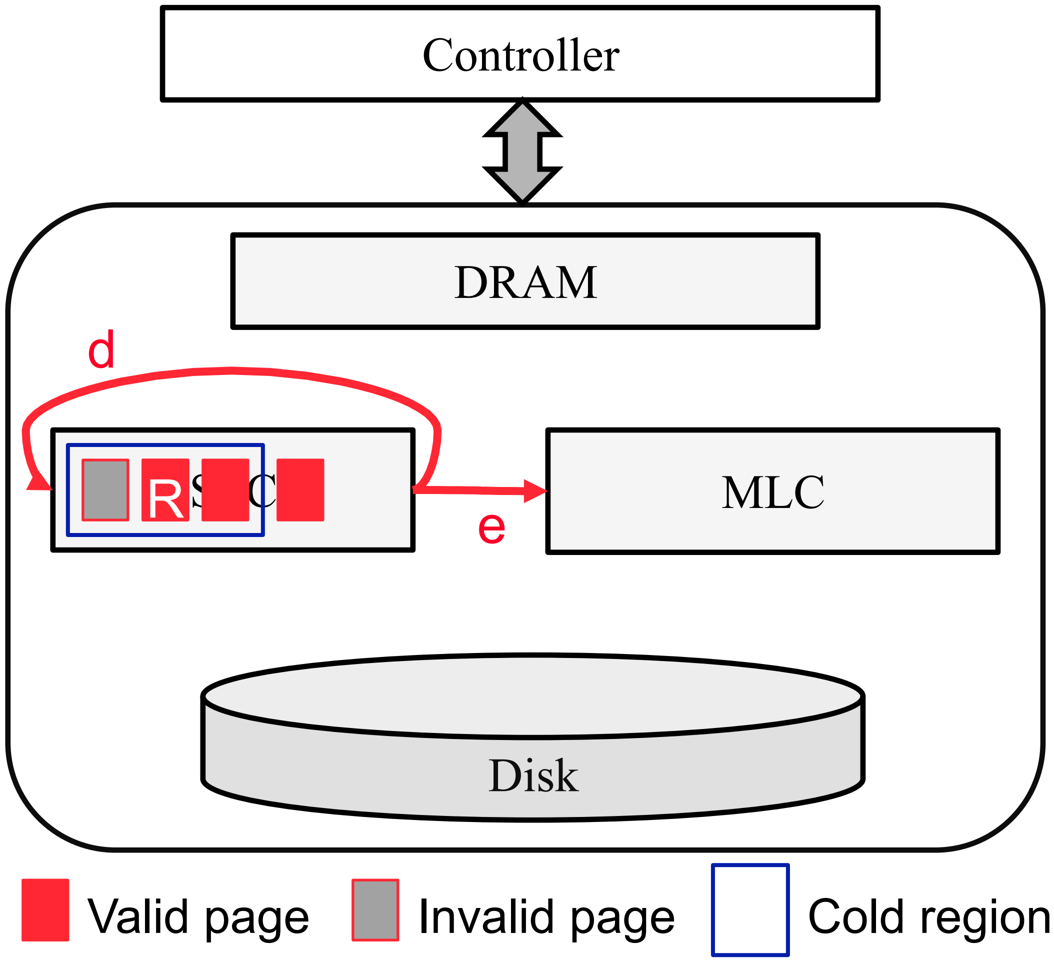
d. GC in SLC

e. Migration to MLC

Since the SLC region is used to reduce DRAM read miss penalty, we gave a chance to be in SLC region to read pages.

When a page is read, the read hit mark is set.

If there are some valid and invalid pages(<70%, >=40%), read hit marked pages are moved into SLC region and other valid pages migrate to MLC region.



Overall Architecture♪

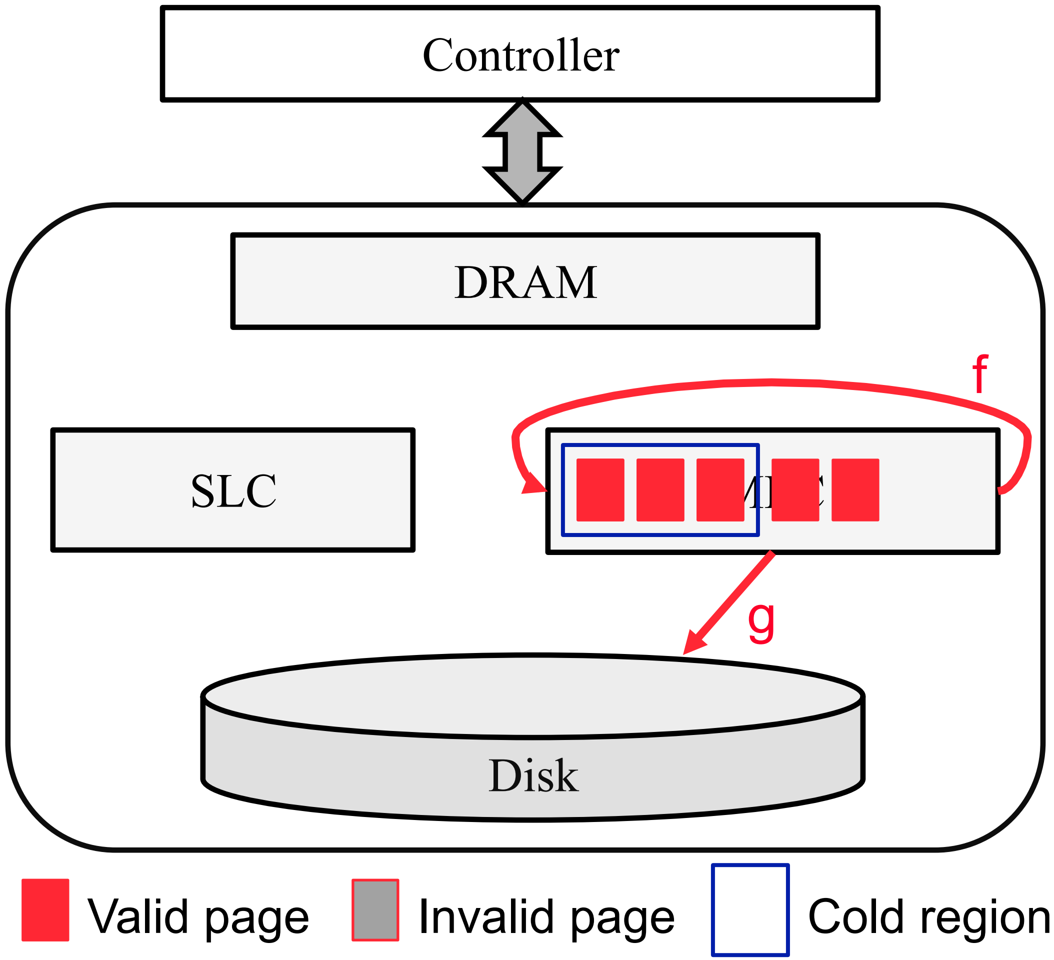
f. GC in MLC

g. Migration to HDD

If the MLC region needs more free space, the garbage collection is invoked.

If the MLC region has sufficient invalid pages, the valid pages are moved into other blocks within MLC region.

However, if there are too many valid pages, they migrate to the HDD.



Overall Architecture♪

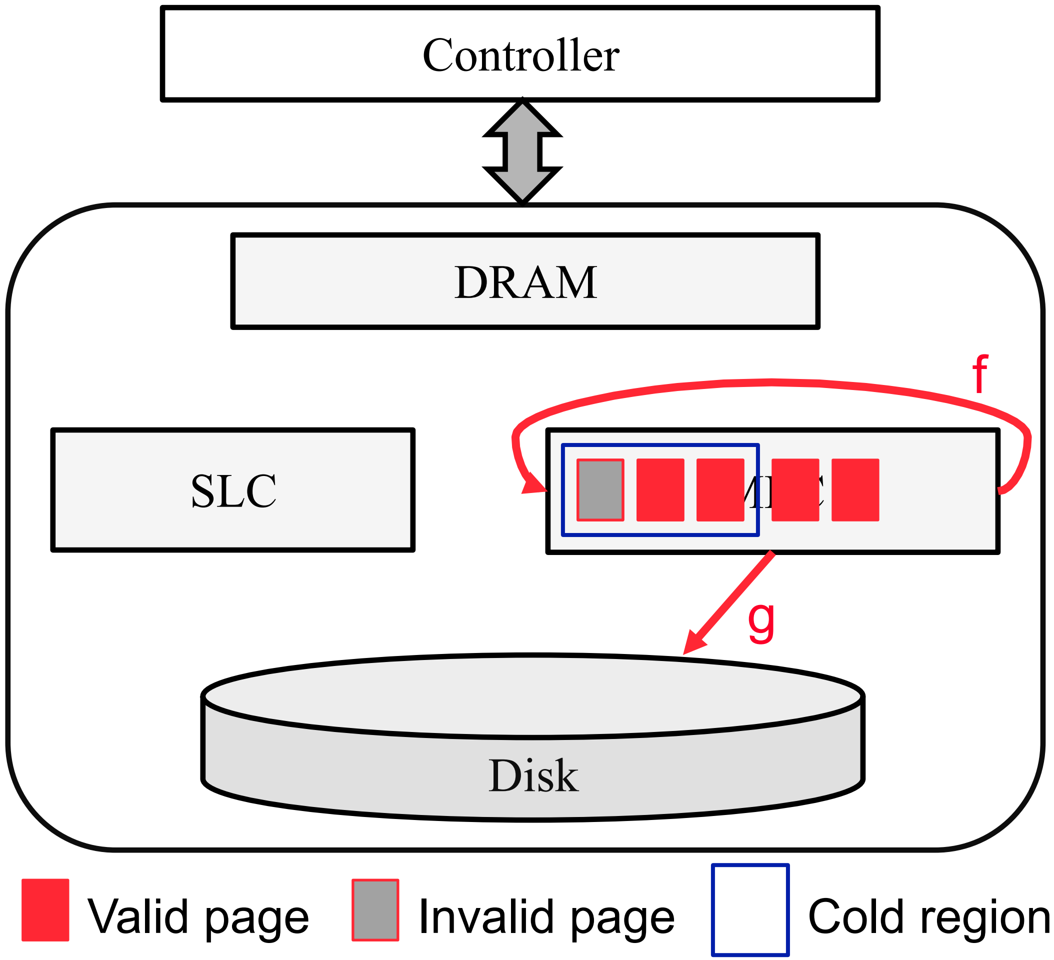
f. GC in MLC

g. Migration to HDD

Since the GC in MLC region can cause the disk accesses, disk power state should be considered.

If the disk is spin-up, it is easy to migration to HDD. (number of invalid pages<60%)

If the disk is spin-down, it is hard to migration to HDD. (number of invalid pages<40%)



Experiments♪

- We implemented a hybrid HDD simulator.
  - DRAM size : 512KB
  - SLC/MLC combined flash memory capacity
    - SLC only : 512MB / MLC only : 1GB
  - HDD model : Samsung's HM080H1

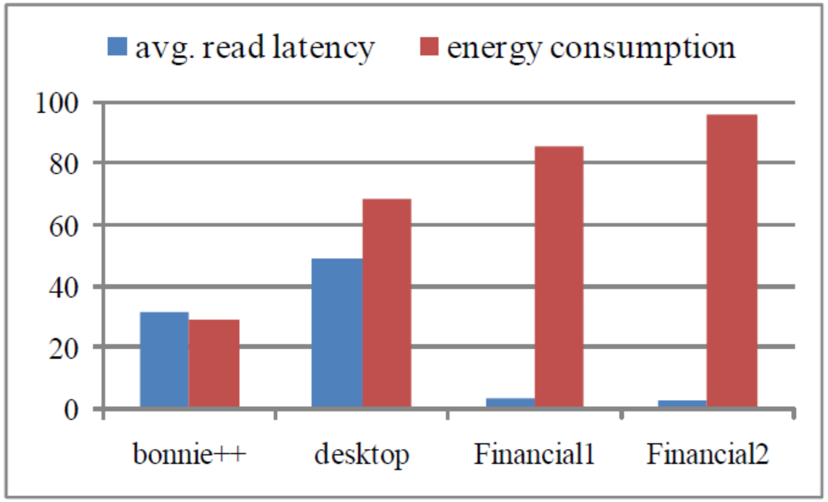
- Trace
  - bonnie++
    - Popular storage benchmark program
  - Desktop
    - Real I/O trace collected executing desktop applications.
  - Financial1, Financial2
    - OLTP application traces
    - http://traces.cs.umass.edu/index.php/Storage/Storage

Experiments♪

- We experimented
  - to evaluate the performance of the hybrid HDD
  - to determine the optimal proportion between the two regions
  - to compare with alternative policies

Experiments♪

- Performance of the hybrid HDD

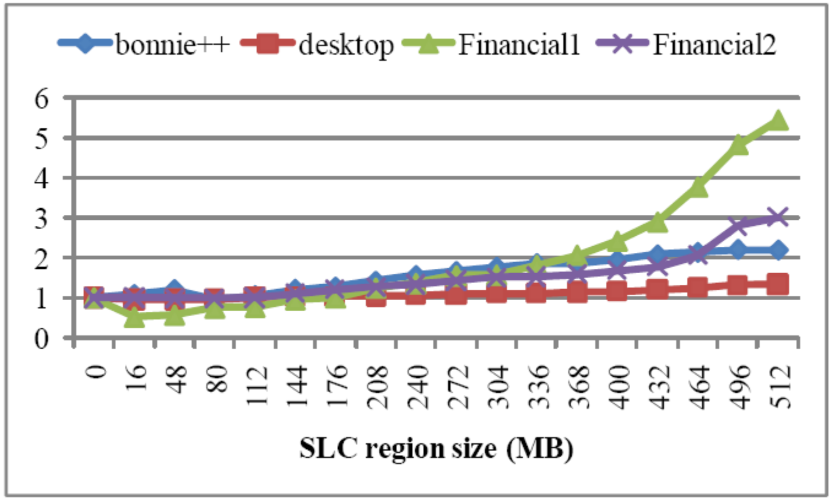


(a) comparison with normal HDD

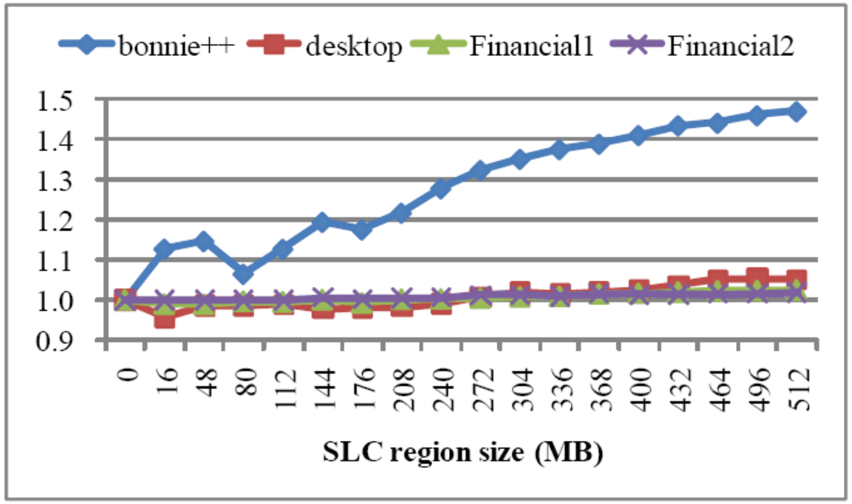
  - Read latencies are reduced by 52~97%
  - Energy consumptions are reduced by 4~71%

Experiments♪

- Changes by varying the size of the SLC region



(a) change of avg. read latency



(c) change of energy consumption

- As the size of the SLC region increases, the read latency and energy consumption increase.
  - The total size of the NVC decreases.

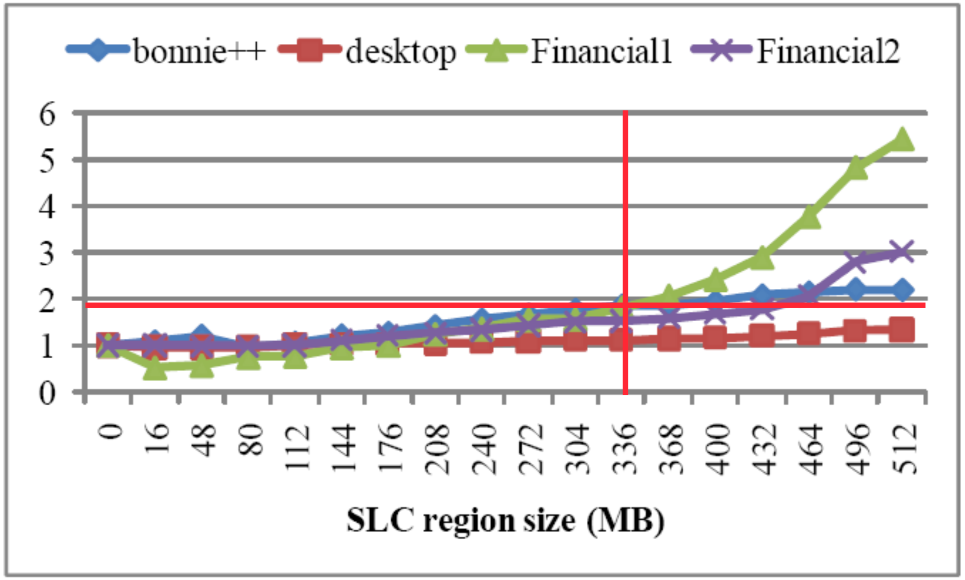
Experiments♪

- Changes by varying the size of the SLC region

Better performance than the SLC-only NVC and a longer lifespan than an MLC-only NVC

For example, if we have the constraint that the lifespan, the minimum SLC region size is identified from the result in this figure.

By selecting the SLC region size, the read latency is reduced by 67% in comparison to the SLC-only NVC.

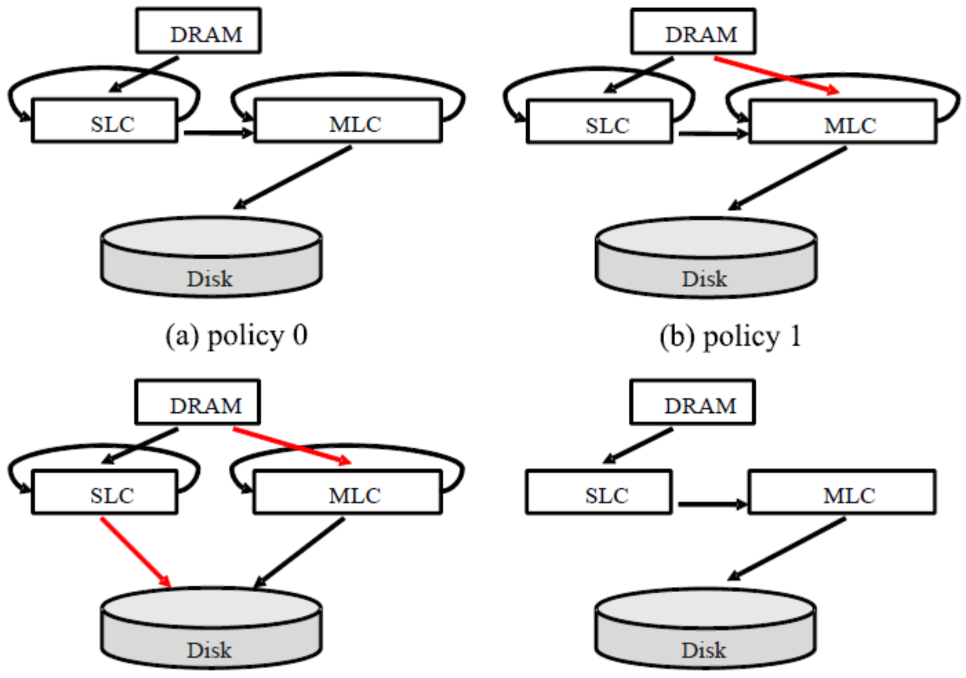


(a) change of avg. read latency

- As the size of the SLC region increases, the lifespan increases .
  - The number of P/E cycles of an SLC block is five times that of an MLC block.

Experiments♪

- Alternative policies



(a) policy 0 (b) policy 1 (c) policy 2 (d) policy 3

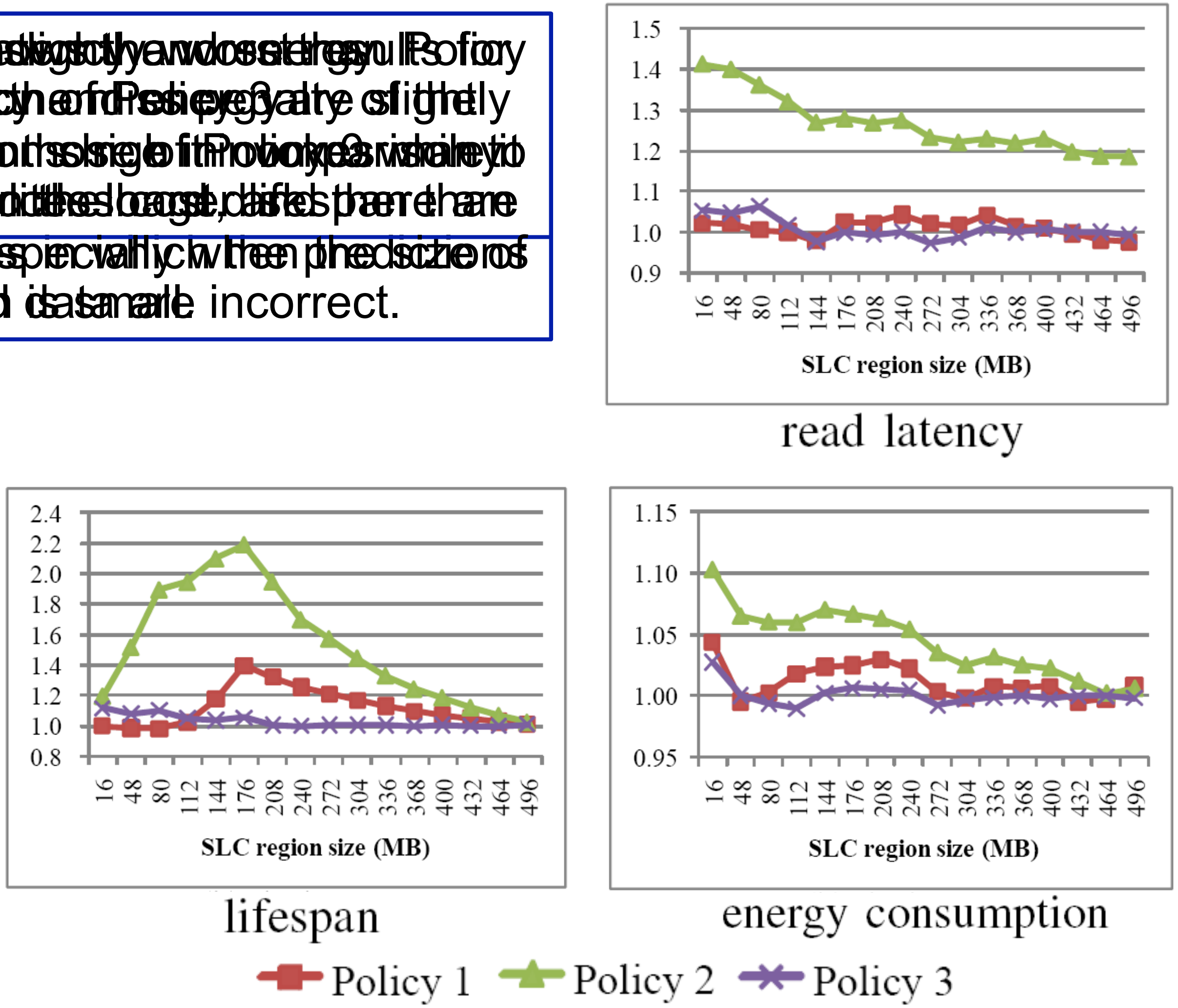
- Policy 0:proposed
- Policy 1: same as Policy 0 except that the large-sized write requests bypass the SLC region, assuming the large data will be cold data.
- Policy 2: the SLC and MLC regions are managed separately
- Policy 3: same as Policy 0 except that the garbage collection for eac h region moves cold pages into lower level storage without internal migrations.

Experiments

- Alternative policies

Policy 2 is slightly worse than Policy 0 in read latency and energy because of the MLC region usage. But, Policy 2 shows better lifespan than Policy 0, especially when the size of SLC region is small.

The read latency and energy consumption of Policy 3 are slightly worse than those of Policy 0 which requires additional cost, but there are many cases in which the predictions for the cold data are incorrect.



Conclusion♪

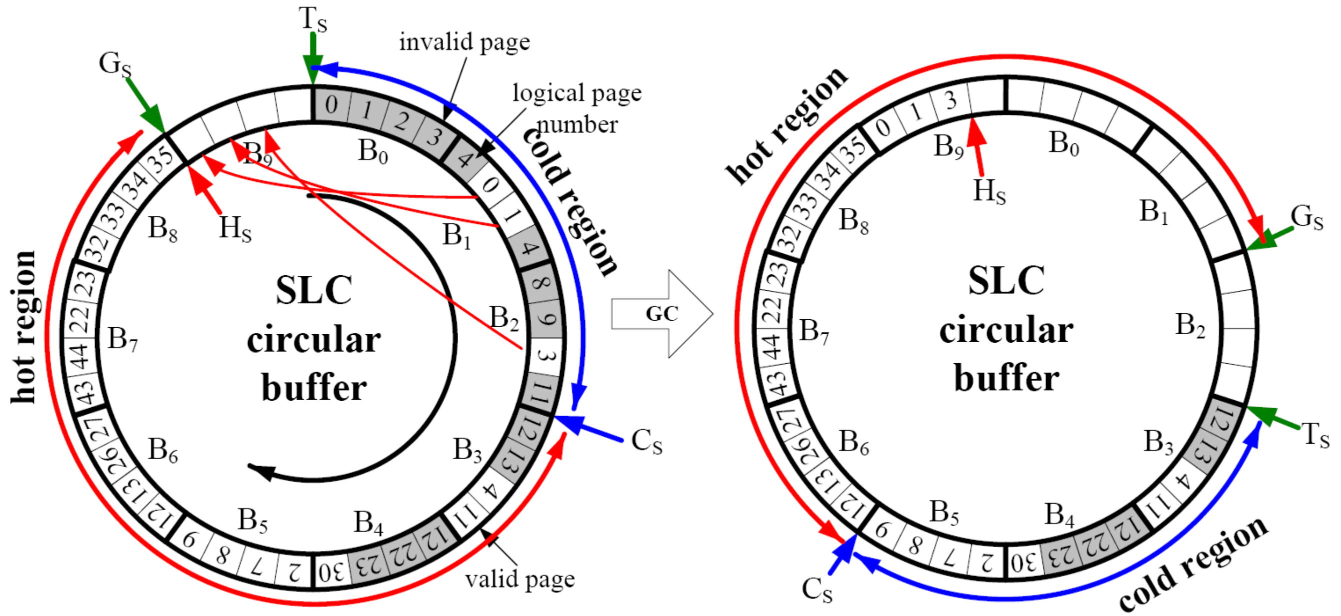
- We proposed an effective management scheme for SLC /MLC combined flash memory in hybrid HDD.
  - We utilized the SLC region as a first-level write buff er for hot data and the MLC region as a second-lev el write buffer for cold data.
- Our hybrid HDD can maximize performance and energy reduction guaranteeing the lifespan constraint.

Q&A♪

Thank you

Garbage Collection♪

- The SLC and MLC regions are maintained as circular buffers.
  - Tail pointer : points to the oldest page
  - Head pointer : points to the youngest page
  - Cold pointer : separates cold and hot region
  - GC pointer : triggers garbage collection



Garbage Collection♪

- GC algorithm for SLC region

```
if (0.7 ≤ portion of invalid pages in cold region ) then
    all valid pages migrate within SLC region;
else if (0.4 ≤portion of invalid pages in cold region < 0.7 )
    only read-hit valid pages migrate within SLC region
      and other valid pages migrate to MLC region;
else  /* portion of invalid pages in cold region < 0.4 */
    all valid pages migrate to MLC region;
```

- GC algorithm for MLC region

```
if (HDD is spin-up) then
    if (0.6 ≤ portion of invalid pages in cold region ) then
        all valid pages migrate within MLC region;
    else
        all valid pages are flushed into HDD;
else /* HDD is spin-down */
    if (0.4 ≤portion of invalid pages in cold region )
        all valid pages migrate within MLC region;
    else
        all valid pages are flushed into HDD;
```

Garbage Collection♪

- Power state transition diagram of an HDD