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

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

• Objective
  • Propose an effective management scheme for SLC/MLC combined 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 the near future.
    - Faster I/O performance
    - Higher energy efficiency
• Flash memory
  • A non-volatile memory able to be electrically erased and programmed.
  • 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
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
• **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 memory, 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 capacity of the flash memory is determined.
• Samsung semiconductor released SLC/MLC combined flash memory, Flex-OneNAND, in 2007

<table>
<thead>
<tr>
<th></th>
<th>SLC</th>
<th>MLC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Page size</strong></td>
<td>4KB</td>
<td></td>
</tr>
<tr>
<td><strong>Block size</strong></td>
<td>256KB (64pages)</td>
<td>512KB (128pages)</td>
</tr>
<tr>
<td><strong>Page read</strong></td>
<td>45us</td>
<td>50us</td>
</tr>
<tr>
<td><strong>Page write</strong></td>
<td>240us</td>
<td>1ms</td>
</tr>
<tr>
<td><strong>Block erase</strong></td>
<td>500us</td>
<td></td>
</tr>
<tr>
<td><strong>P/E cycles</strong></td>
<td>50K</td>
<td>10K</td>
</tr>
</tbody>
</table>
Background

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

- $256\text{KB} \times 512 \text{ blocks} + 512\text{KB} \times 512 \text{ blocks} = 384\text{MB}$

- $256\text{KB} \times 256 \text{ blocks} + 512\text{KB} \times 768 \text{ blocks} = 448\text{MB}$
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.
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

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.

e. Migration to MLC
d. GC in SLC

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.

e. Migration to MLC
Overall Architecture

f. GC in MLC

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.

g. Migration to HDD

Embedded Software Lab.
f. GC in MLC

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%)

g. Migration to HDD
**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
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

- 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

- 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.

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.
Experiments

• Alternative policies

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 each region moves cold pages into lower level storage without internal migrations.
• Alternative policies

Policy 2 strategy underestimates Policy 3 and overestimates Policy 4 slightly. This is due to the fact that Policy 3 tends to be more aggressive in its write policy compared to Policy 4.

Policy 5 shows the worst results for read latency and energy consumption since it invokes many requests on the hard disk.

Policy 6 results in slightly worse energy consumption compared to Policy 4, especially when the size of SLC region is small.
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 buffer for hot data and the MLC region as a second-level write buffer for cold data.
• Our hybrid HDD can maximize performance and energy reduction guaranteeing the lifespan constraint.
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

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

- GC algorithm for MLC region

1: if (HDD is spin-up) then
2:   if (0.6 ≤ portion of invalid pages in cold region) then
3:     all valid pages migrate within MLC region;
4:   else
5:     all valid pages are flushed into HDD;
6: else /* HDD is spin-down */
7:   if (0.4 ≤ portion of invalid pages in cold region)
8:     all valid pages migrate within MLC region;
9:   else
10: all valid pages are flushed into HDD;
Garbage Collection

- Power state transition diagram of an HDD

![Power state transition diagram of an HDD](image-url)