Tri-Hybrid SSD with storage class memory (SCM) and MLC/TLC NAND Flash Memories

Chihiro Matsui, Tomoaki Yamada, Yusuke Sugiyama, Yusuke Yamaga, and Ken Takeuchi
Chuo University, Japan
Outline

- Introduction
- Hybrid SSDs with
  1. MLC and TLC NAND flash memories
  2. SCM and MLC NAND flash memory
  3. SCM, MLC and TLC NAND flash memories
- Summary
Outline

- Introduction
- Hybrid SSDs with
  1. MLC and TLC NAND flash memories
  2. SCM and MLC NAND flash memory
  3. SCM, MLC and TLC NAND flash memories
- Summary
SCM and NAND Flash Memory in Memory Hierarchy

- **Access time**
  - 10ms
  - 1us
  - 10ns
  - 1ns

- **Capacity**
  - MLC NAND flash
  - TLC NAND flash
  - DRAM
  - CPU

- **Storage Class Memory (SCM)**

- **MRAM, ReRAM, PRAM, and 3D XPoint**
ReRAM Set/Reset Time

- Short set/reset time if no verify operation

Program Error in ReRAM

- Program BER increases with set/reset cycles

![Graph showing Program BER vs. Set/Reset cycles]

MLC and TLC NAND Flash Latency

- TLC flash has longer read/write latency than MLC flash

![Graph showing write latency comparison between MLC and TLC flash over Write/Erase cycles.]

MLC flash

Write latency [us]

- Upper page
- Lower page

TLC flash

Write latency [us]

- Upper page
- Middle page
- Lower page

1.5x longer

Errors in MLC and TLC NAND Flash

- TLC flash has higher error rate than MLC flash

---

Program disturb error

- TLC flash
- MLC flash

Data retention error

- TLC flash
- MLC flash

---

Outline

- Introduction
- Hybrid SSDs with
  1. MLC and TLC NAND flash memories
  2. SCM and MLC NAND flash memory
  3. SCM, MLC and TLC NAND flash memories
- Summary
Application Characteristics

Application characteristics defined by

- Average overwrite = Total write data size / user data size
- Average read frequency = Total read data size / user data size
- Random: data size is 8KB or less (half of page size)

Data Management Strategy

- Determined by memory characteristics used in hybrid SSD

<table>
<thead>
<tr>
<th>Latency:</th>
<th>Short</th>
<th>Memory characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endurance:</td>
<td>High</td>
<td>Long</td>
</tr>
<tr>
<td>Reliability:</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Capacity:</td>
<td>Small</td>
<td>Large</td>
</tr>
<tr>
<td>Cost:</td>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>

Data characteristics to be stored

- Hot data
- Frozen data
1) MLC and TLC NAND Flash

Hybrid SSD

- Frozen data are collected in TLC NAND flash when MLC Flash GC

Write operation:
1. Write to MLC NAND flash
2. Return “write complete” to host

Frozen data eviction operation:
1. Search frozen and static data during MLC NAND flash GC operation
2. Evict to TLC NAND flash

Read operation:
1. Read from MLC or TLC NAND flash
2. Return “read complete” to host

1) MLC and TLC NAND Flash
Hybrid SSD

- TLC flash improves the hybrid SSD performance and prolongs MLC flash endurance

With proxy_1 (Hot Random)

---

2) SCM and MLC NAND Flash

Hybrid SSD

- SCM stores hot or random data. No data deduplication in between SCM and MLC flash

Hot and random write operation:
1. Write to SCM and update LRU table
2. Return “write complete” to host

Cold data eviction operation:
1. Search cold and sequential data in SCM
2. Evict to MLC NAND flash

Cold and sequential write operation:
1. Write to MLC NAND flash
2. Return “write complete” to host

2) SCM and MLC NAND Flash Hybrid SSD

with SCM read/write latency 100ns

---

Normalized IOPS to MLC flash only SSD

<table>
<thead>
<tr>
<th>SCM capacity ratio</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>prxy_0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>proj_0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hm_0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>web_1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>prxy_1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>proj_3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>src2_1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Write intensive**

- prxy_0: Hot Random
- proj_0: Hot Seq.
- hm_0: Cold Random
- web_1: Cold Seq.
- prxy_1: Hot Random
- proj_3: Cold Random
- src2_1: Cold Seq.

**Read intensive**

3) SCM and MLC/TLC NAND Flash Tri-hybrid SSD

- SCM stores hot or random data, and TLC stores frozen data

**Hot and random write operation:**
1. Read from SCM or MLC NAND flash
2. Return “write complete” to host

**Cold and sequential write operation:**
1. Write to MLC NAND flash
2. Return “write complete” to host

**Frozen data eviction algorithm:**
1. Search frozen and static data during MLC NAND flash GC operation
2. Evict to TLC NAND flash

---

3) SCM and MLC/TLC NAND Flash Tri-hybrid SSD

- SCM improves SSD performance of hot applications

prxy_0 workload 
*(Hot Random)*

proj_0 workload 
*(Hot Seq.)*

3) Optimal Memory Capacity in Tri-hybrid SSD with 1.1-times Cost

w/ SCM read/write latency 100ns

![Diagram showing capacity ratio for different workloads and storage classes.]

<table>
<thead>
<tr>
<th>Storage class memory</th>
<th>Bit cost ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLC flash</td>
<td>1</td>
</tr>
<tr>
<td>TLC flash</td>
<td>2/3</td>
</tr>
</tbody>
</table>

3) Tri-hybrid SSD Performance with 1.1-times Cost

w/ SCM read/write latency 100ns

Outline

- Introduction
- Hybrid SSDs with
  1. MLC and TLC NAND flash memories
  2. SCM and MLC NAND flash memory
  3. SCM, MLC and TLC NAND flash memories
- Summary
Summary

- SCM and TLC flash store hot and frozen data, respectively in tri-hybrid SSD
- If 10% SSD cost increase is allowed:
  - Small capacity of SCM boosts SSD performance by 6 times for hot random application
  - Hot sequential or cold random applications require mix of SCM, MLC and TLC NAND flash memories
  - Small SCM capacity is enough for cold sequential application
Acknowledgement

This presentation is based on results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).