Data Management Design for Interlaced Magnetic Recording

**Fenggang Wu**, Baoquan Zhang, Zhichao Cao, Hao Wen, Bingzhe Li, Jim Diehl, Guohua Wang*, David H.C. Du

*University of Minnesota, Twin Cities*  
*South China University of Technology*
IMR: Higher areal data density than CMR, lower write amplification (WA) than SMR.
IMR Tracks | Width | Laser Power | Data Density | Data Rate | Track Capacity
---|---|---|---|---|---
Bottom Tracks | wider | higher | higher(+27%)[1] | higher | higher
Top Tracks | narrower | lower | lower | lower | lower

Updating top tracks has no penalty
Updating bottom tracks causes Write Amplification (WA)
Only using bottom tracks when disk is not full may reduce WA.
I/O Performance depends on disk usage, and layout design.

The Problem: Data Management Design for IMR

- Adapt to disk usage.
- Reduce write amplification.
- Bound memory budget.
Outline

- The problem
- The solutions
  - Baseline design
  - DM-IMR design
- The results
- Future works
Track Group (TG)

Track Group (TG): an interlaced set of consecutive physical top and bottom tracks.

This paper only focuses on the data allocation and management within one TG.
Three-Phase Baseline

1st Phase (0~56%)

2nd Phase (56~78%)

3rd Phase (78~100%)

Bottom Tracks

Top Tracks

Track Group (TG)
DM-IMR: Data Management for IMR

- Top-Buffer
- Block-Swap
Top-Buffer

The idea: opportunistically buffer bottom-write requests into unallocated top tracks; accumulate multiple updates and write to bottom only once.
Top-Buffer

Design choice: user defines the size budget of the memory table; memory budget determines the max number of tracks Top-Buffer may have.

E.g., If the user bounds the memory table size to be 0.004% of the disk capacity, the max size of the Top-Buffer will be 2% of the disk capacity.

Memory Mapping Table

<table>
<thead>
<tr>
<th>lba</th>
<th>pba</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>78</td>
</tr>
<tr>
<td>46</td>
<td>79</td>
</tr>
</tbody>
</table>

bounded memory budget

E.g., If the user bounds the memory table size to be 0.004% of the disk capacity, the max size of the Top-Buffer will be 2% of the disk capacity.
Top-Buffer

Top-Buffer capacity also depends on available unallocated top tracks.

Problem:
- Extremely small Top-Buffer brings little benefit.
- Top-Buffer cannot function when usage=100%.
Block-Swap

The idea: progressively swap hot data in bottom tracks with cold data in top tracks.

Design choice: Top-Buffer and Block-Swap **share the memory budget**; Block-Swap will kick in when Top-Buffer cannot fully use the mapping table (i.e. usage is high).

<table>
<thead>
<tr>
<th>lba</th>
<th>pba</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>78</td>
</tr>
<tr>
<td>46</td>
<td>79</td>
</tr>
<tr>
<td>27</td>
<td>80</td>
</tr>
<tr>
<td>24</td>
<td>76</td>
</tr>
<tr>
<td>76</td>
<td>24</td>
</tr>
</tbody>
</table>
DM-IMR: Put it together

Top-Buffer: at most 2% of the whole space

Bottom Update Scheme

In-Place

Top-Buffer

Block-Swap

(56~78%)

(78~100%)

(0~56%)

Space utilizations (%)

56%  78%  98%  100%

(more design details in paper)
Evaluation

- IMR Sim
- MSR Cambridge Trace Replay
- Competing Schemes

### Table 1: IMR disk configuration.

<table>
<thead>
<tr>
<th>Basic Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Track pitch</td>
<td>820KTP</td>
</tr>
<tr>
<td>Median top track density</td>
<td>1640KB</td>
</tr>
<tr>
<td>Median bottom track density</td>
<td>2030KB</td>
</tr>
<tr>
<td>RPM</td>
<td>5400</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Derived Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>#tracks (N)</td>
<td>1045800</td>
</tr>
<tr>
<td>Average bottom track size</td>
<td>2MB</td>
</tr>
<tr>
<td>Average top track size</td>
<td>1.6MB</td>
</tr>
</tbody>
</table>

### Space Utilizations (%)

<table>
<thead>
<tr>
<th>Bottom Update Scheme</th>
<th>56%</th>
<th>78%</th>
<th>98%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Place</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Three-Phase Baseline**

**Buffer-Only**
Average Throughput with Varying Usage

- Buffer-Only and DM-IMR both can increase throughput.
- DM-IMR outperforms Buffer-Only after 98% because Block-Swap starts to kick in.

More results in the paper
Summary

• Problem: data management for IMR.

• Two approaches are proposed:
  – Three-Phase baseline
  – DM-IMR, which uses Top-Buffer and Block-Swap to improve from the Three-Phase baseline.

• Results show DM-IMR can increase throughput and reduce write amplification.

• Future work: space manager design for TGs, eviction algorithms of Top-Buffer and Block-Swap, computation optimization, etc.
Data Management Design for Interlaced Magnetic Recording

Thank you! Comments/Questions?