Data Management Design for Interlaced Magnetic Recording

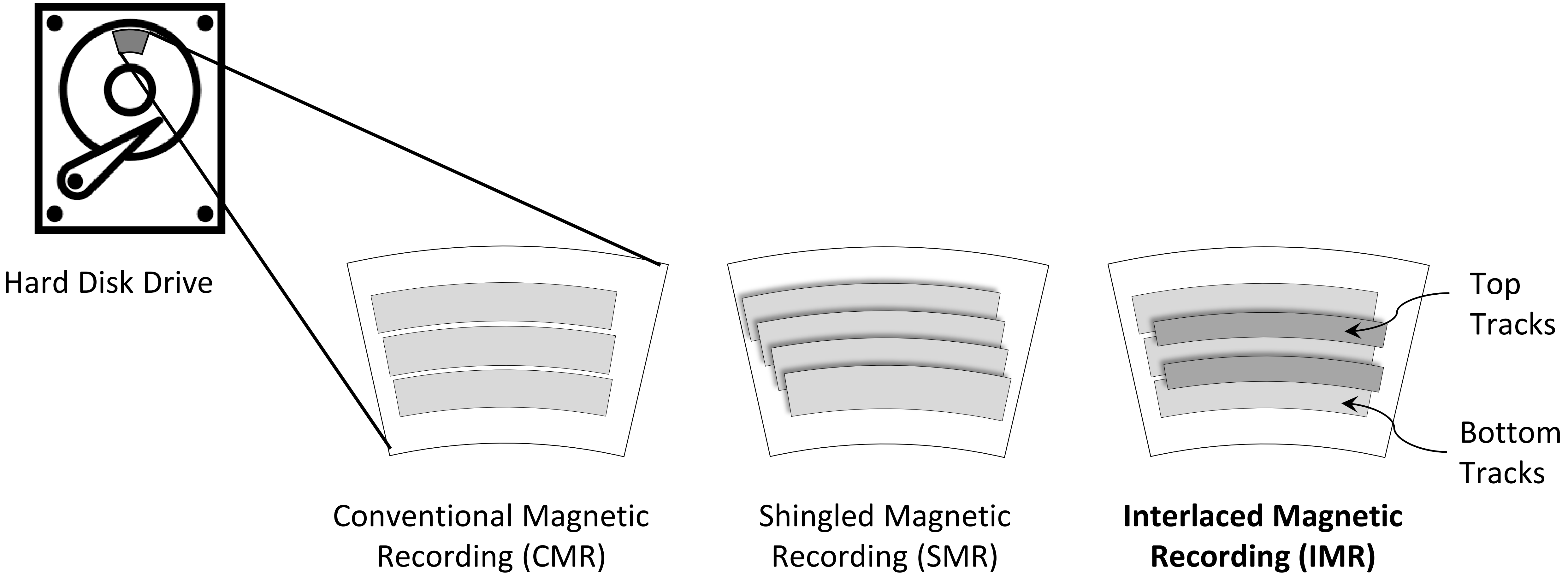
<u>Fenggang Wu</u>, Baoquan Zhang, Zhichao Cao, Hao Wen, Bingzhe Li,
Jim Diehl, Guohua Wang*, David H.C. Du

University of Minnesota, Twin Cities **South China University of Technology*

Center for Research in Intelligent Storage

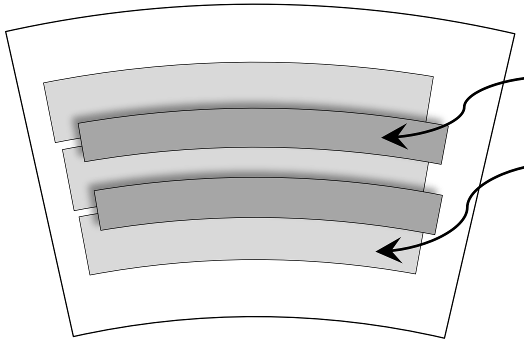
SOUTH CHINA UNIVERSITY OF TECHNOLOGY 1934

UNIVERSITY OF MINNESOTA



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**.

[1]Granz et. al, 2017

The Problem: Data Management Design for IMR

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

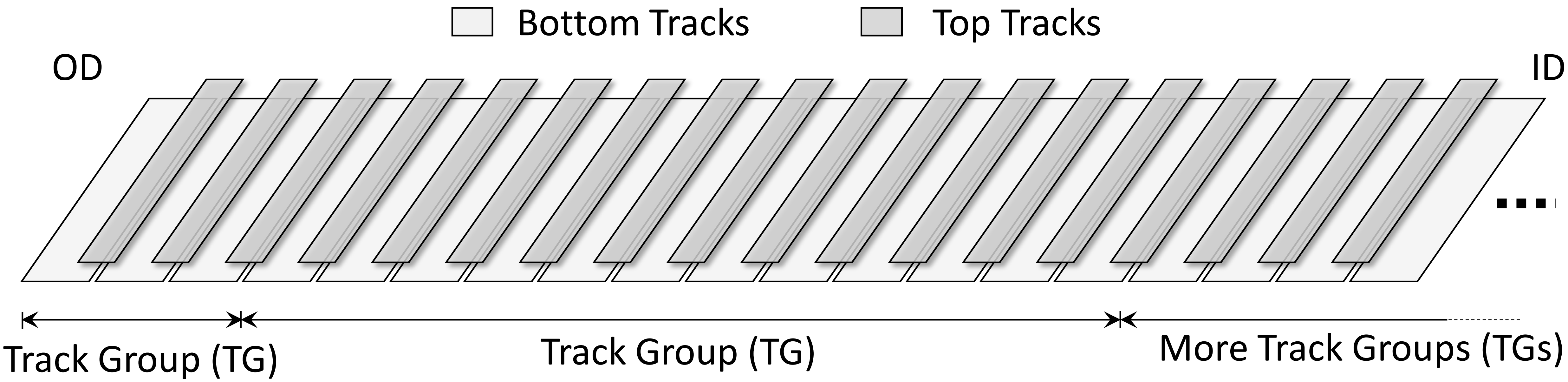
Outline

- *The problem*
- The solutions
  - Baseline design
  - DM-IMR design
- The results
- Future works

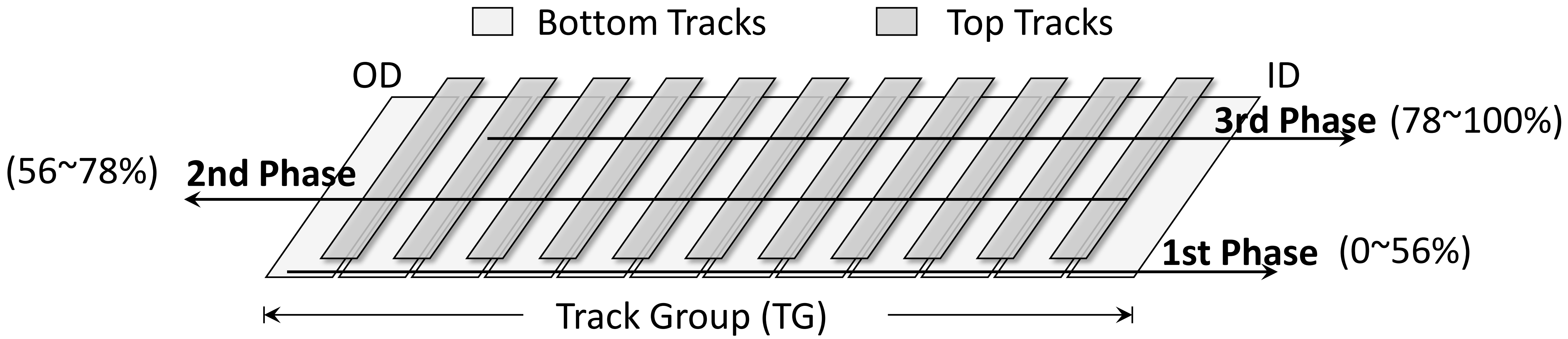
Track Group (TG)

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

This paper only focus on the data allocation and management **within one TG**.



Three-Phase Baseline

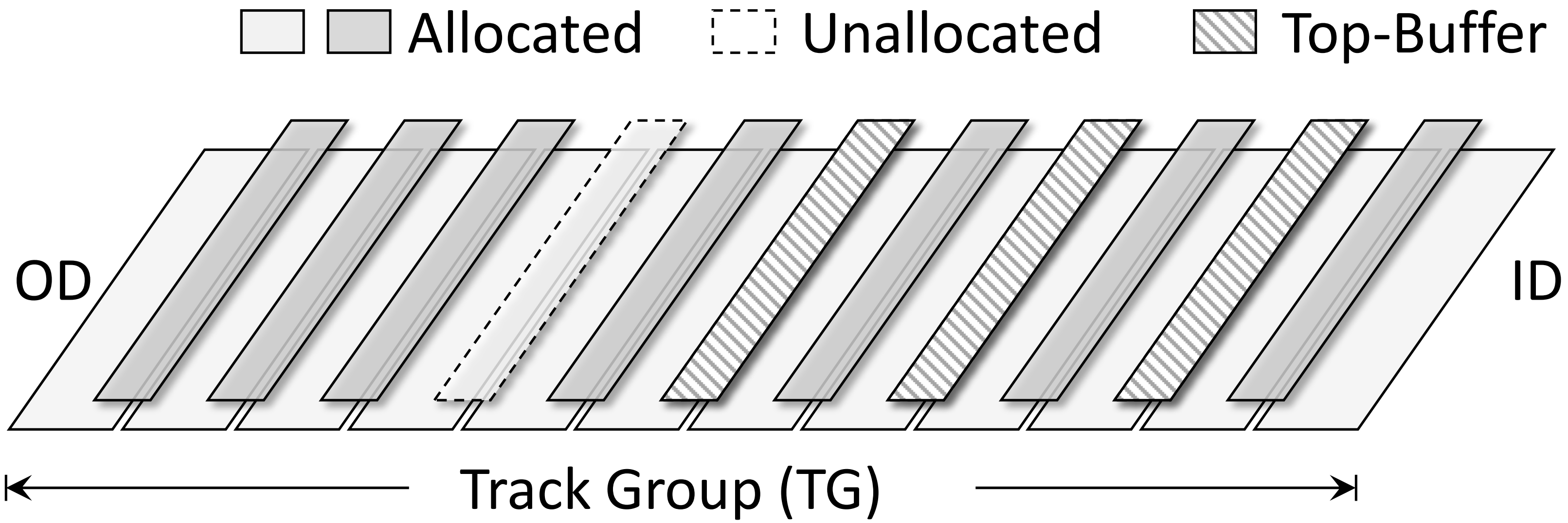


DM-IMR: Data Management for IMR

- 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

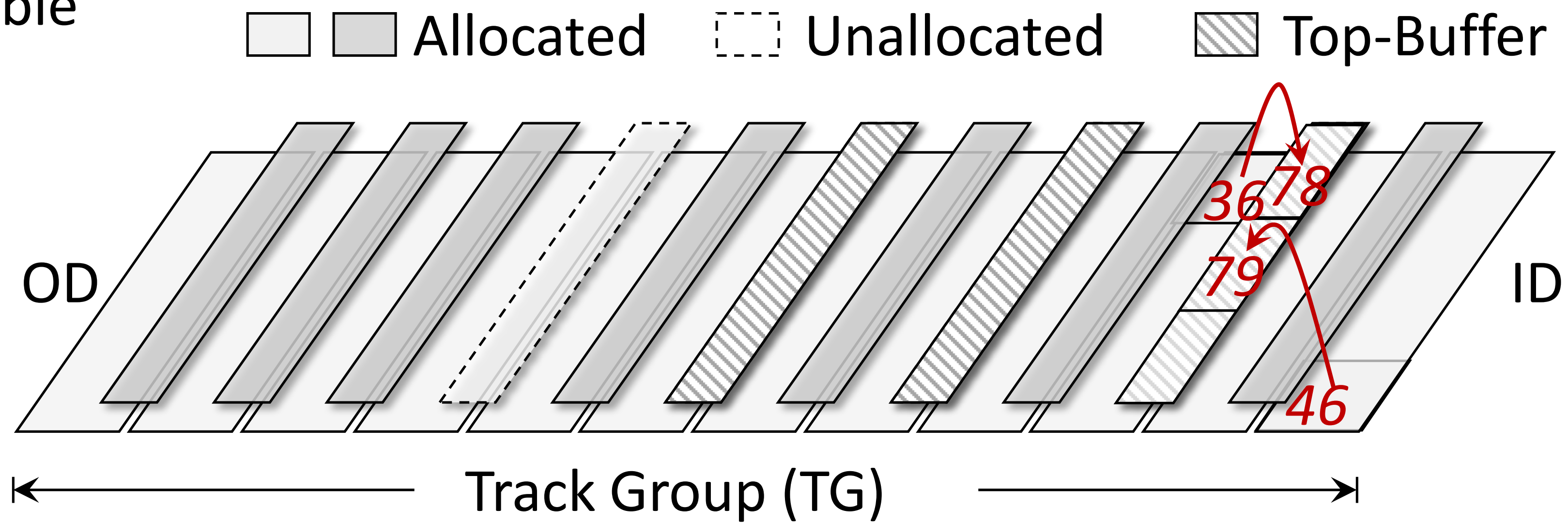
Design choice: user defines the **size budget of the memory table;** memory budget determines the max number of tracks Top-Buffer may have.

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

bounded memory budget

lba	pba
36	78
46	79



Top-Buffer

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

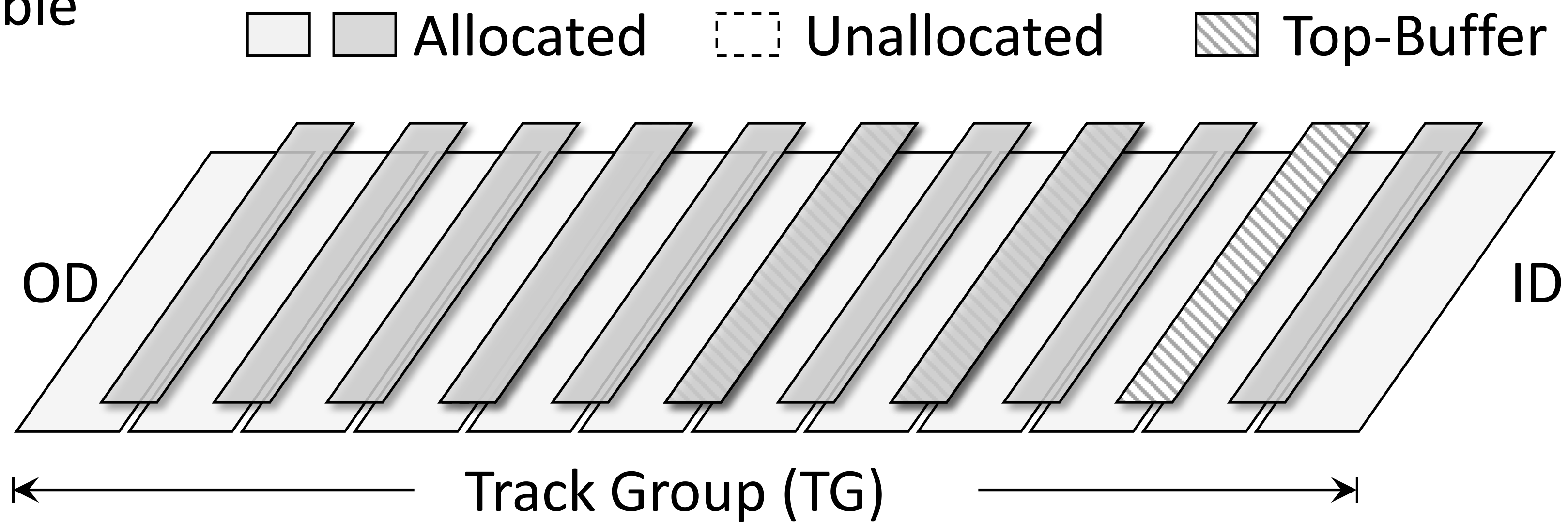
Problem:

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

Memory Mapping Table

bounded memory budget

lba	pba
X1	Y1
X2	Y2
X3	Y3
~~X4~~	~~Y4~~
~~X5~~	~~Y5~~



Block-Swap

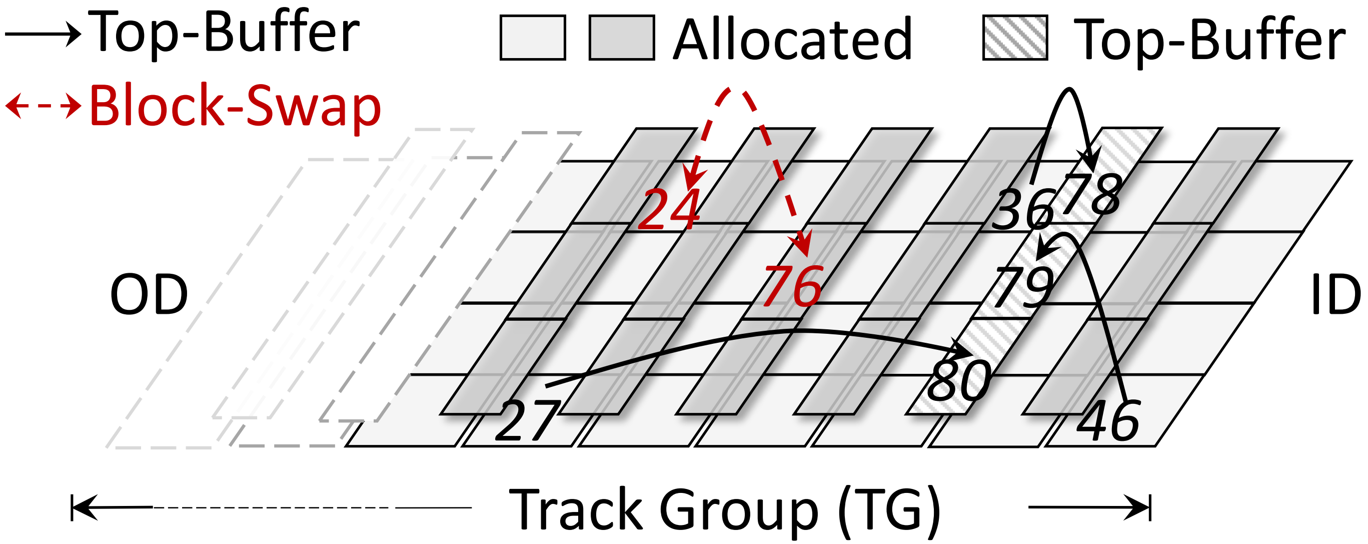
The idea: progressively swap hot data in bottom tracks with cold data in top tracks.

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).

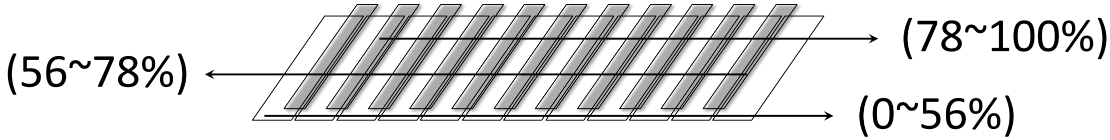
Memory Mapping Table

bounded memory budget

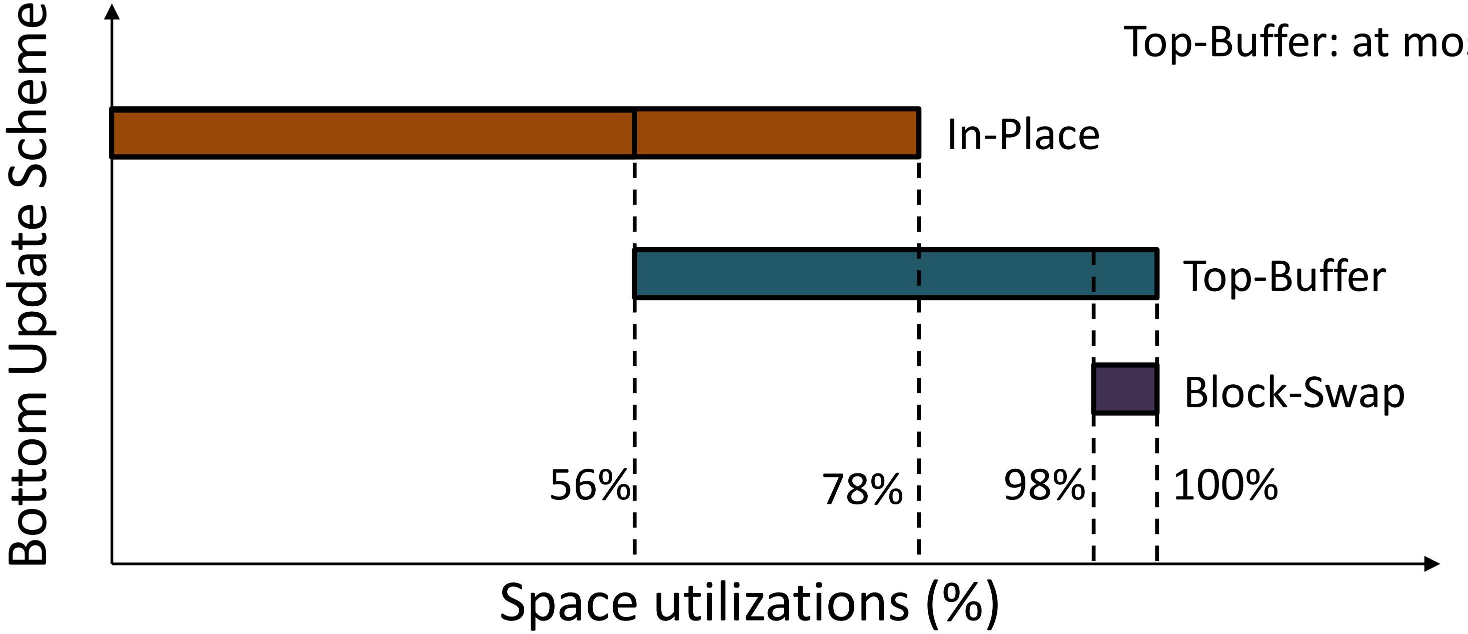
lba	pba
36	78
46	79
27	80
24	76
76	24



DM-IMR: Put it together



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



(more design details in paper)

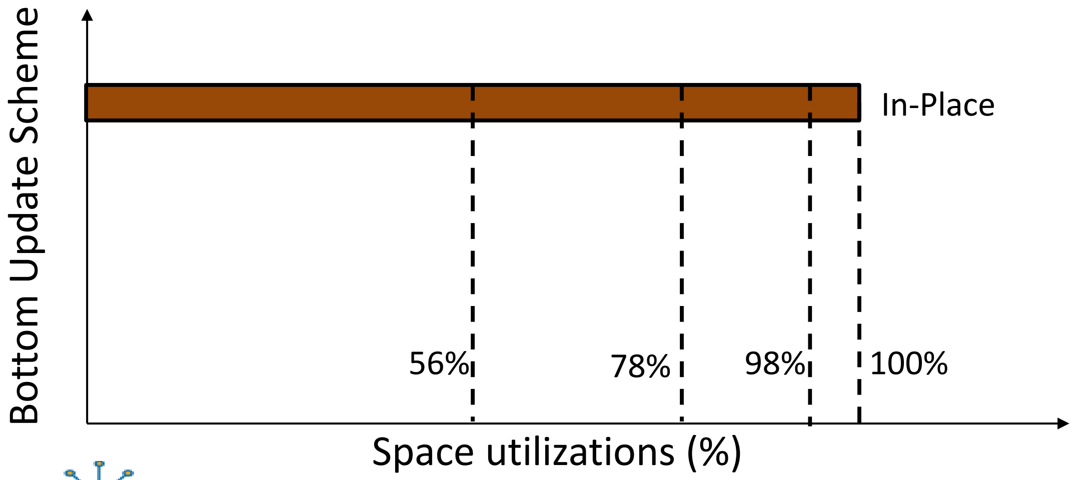
Evaluation

- IMR Sim
- MSR Cambridge Trace Replay
- Competing Schemes

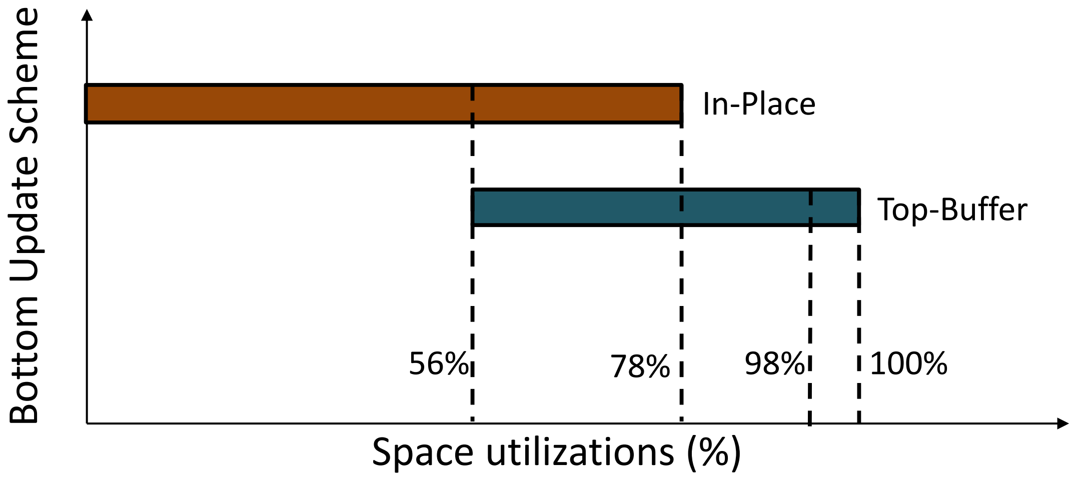
Table 1: IMR disk configuration.

Basic Parameters	
Median Track pitch	820KTPI
Median top track density	1640KBPI
Median bottom track density	2030 KBPI
RPM	5400
Derived Parameters	
#tracks (N)	1045800
Average bottom track size	2MB
Average top track size	1.6MB

Three-Phase Baseline

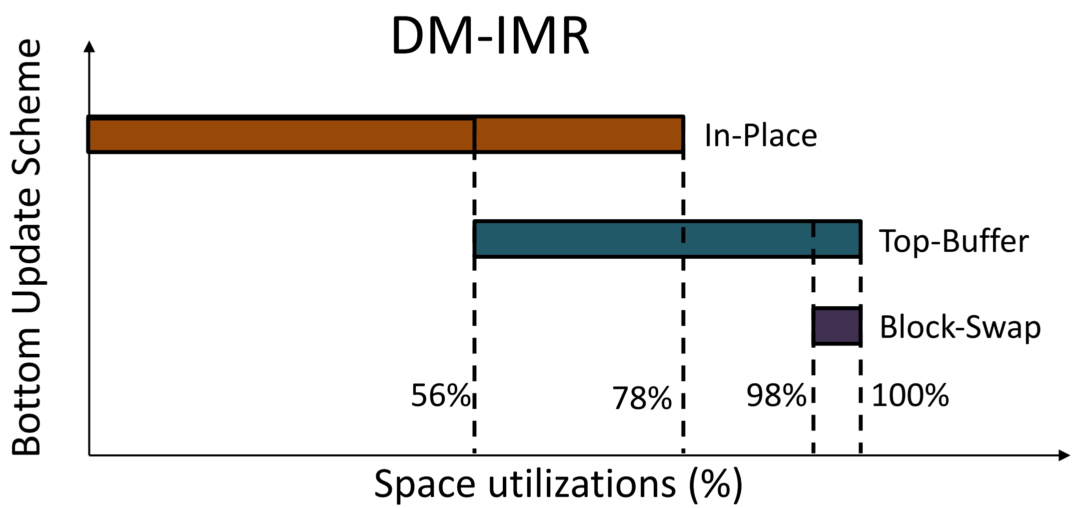


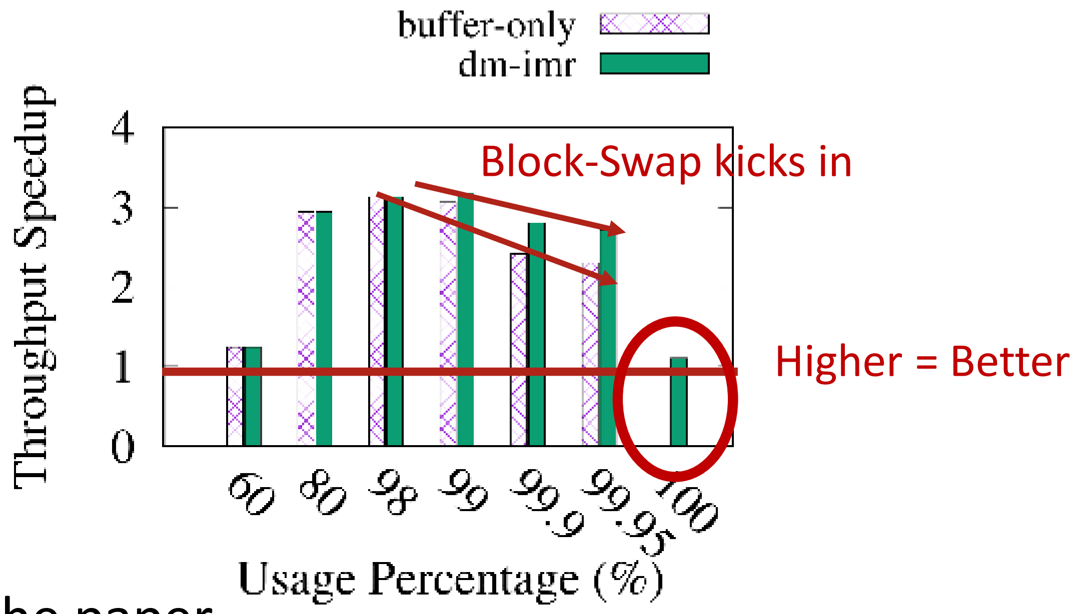
Buffer-Only



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?

CRIS
Center for Research in Intelligent Storage

UNIVERSITY OF MINNESOTA