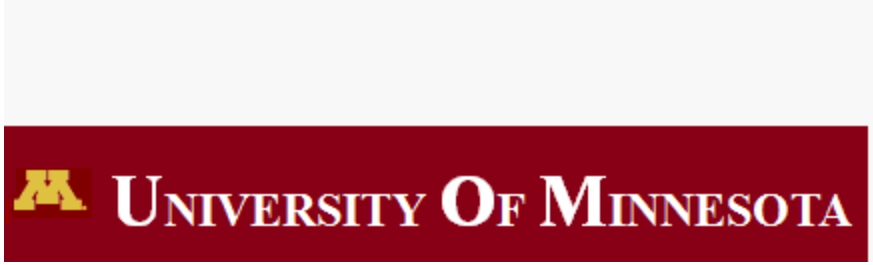


Center for Research in Intelligent Storage

Novel Address Mappings for Shingled Write Disks

Weiping He, David Du

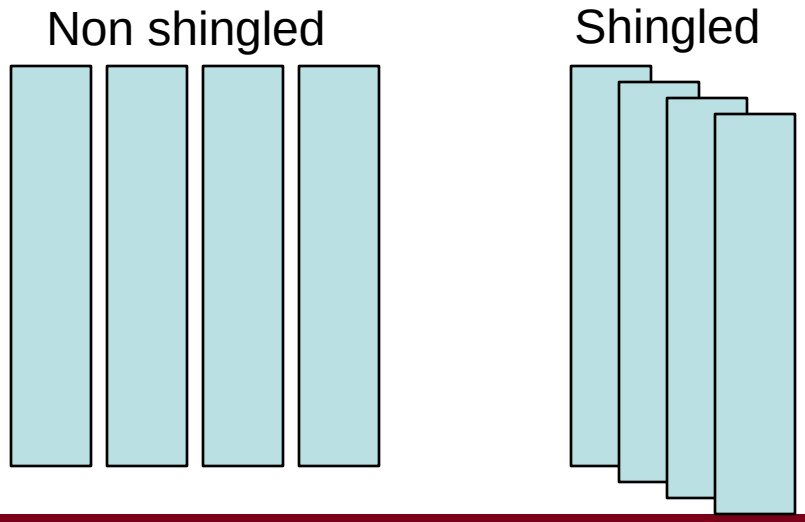
University of Minnesota
Driven to Discover℠

Outline

- Backgrounds
- Motivations/Objectives
- Approaches
- Evaluations

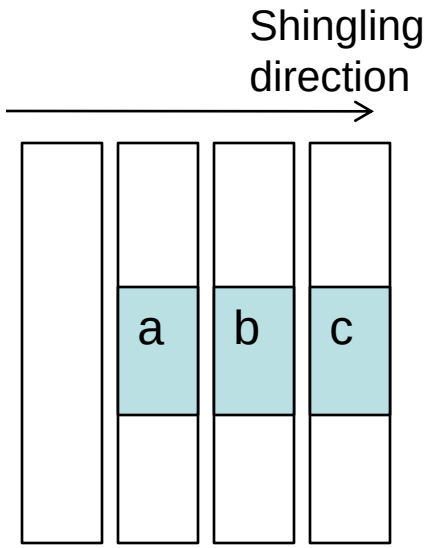
Background

- Traditional HDDs (perpendicular magnetic recording) are reaching areal data density
- Shingled magnetic recording is a new promising technology



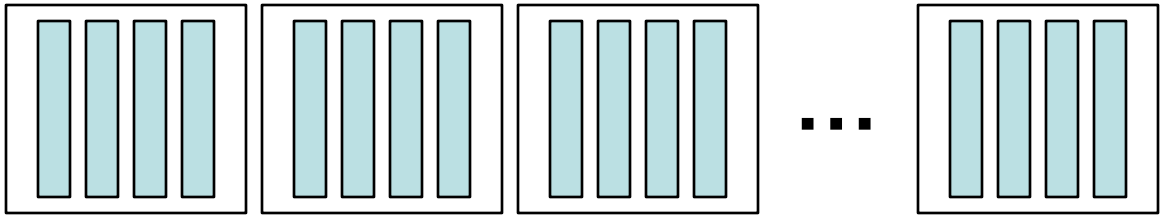
SWD Characteristics

- Sequential write is preferred
- Write/update a block *in place* may destroy the valid data on the subsequent tracks if any
- General approaches to updates:
  - **In-place update:**
    - Extra reads/writes 1 write = 2 reads + 3 writes
  - Out-of-place update:
    - Copy-on-write
    - Mapping table and Garbage collection
- Write amplification
  - Update may incur extra read/write operations



Tradeoff Between Space and Performance

- SWD Layout
  - Tracks are organized into bands
  - There are safety gaps between bands



Write width = W
Band size = N

A good candidate: W=2, N=4
SG = 1.6
WAR = 4

$$SG = W\frac{N}{N+W-1} \quad (1)$$

$$WAR = \frac{1}{N}\sum_{i=0}^{N-1}(1+2i) = N \quad (2)$$

Objective

- Good balance between capacity and performance
  - Reduce the write amplification overhead

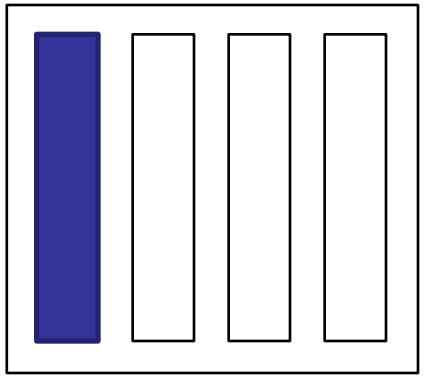
Motivation for New Schemes

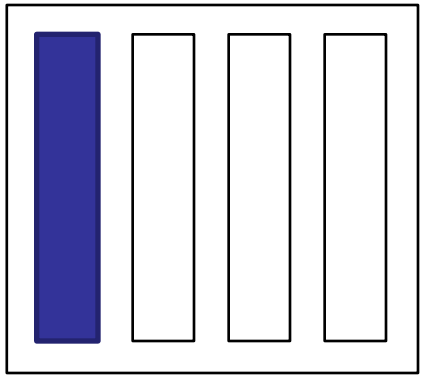
- General rule
  - Delay the use of track(s) in the middle of the bands, e.g., 3rd tracks

	Conventional	Novel
25%	no	no

Affected Tracks

Single Band comparison





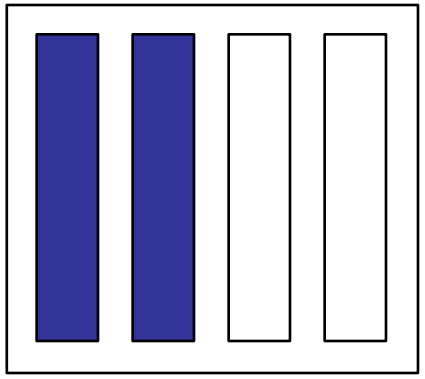
Motivation for New Schemes

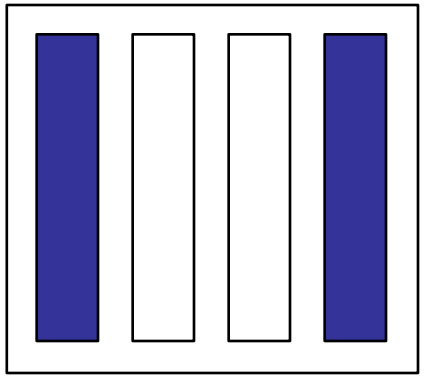
- General rule
  - Delay the use of track(s) in the middle of the bands, e.g., 3rd tracks

	Conventional	Novel
25%	no	no
50%	Track1	no

Affected Tracks

Single Band comparison





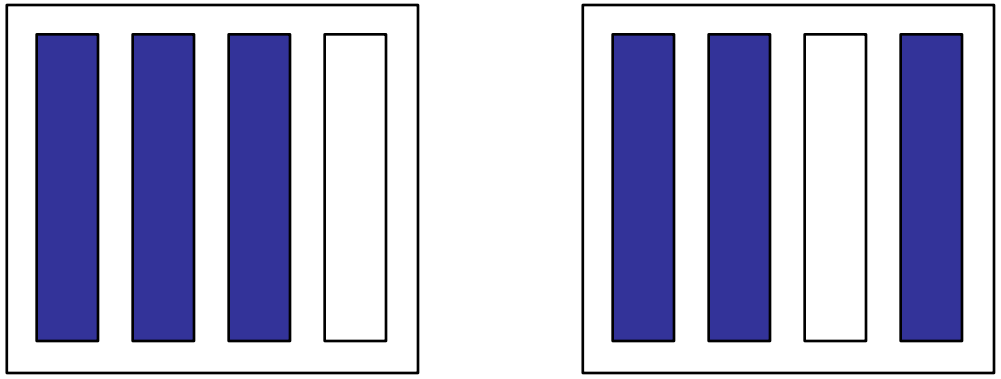
Motivation for New Schemes

- General rule
  - Delay the use of track(s) in the middle of the bands, e.g., 3rd tracks

Affected Tracks

	Conventional	Novel
25%	no	no
50%	Track1	no
75%	Track 1,2	Track1

Single Band comparison



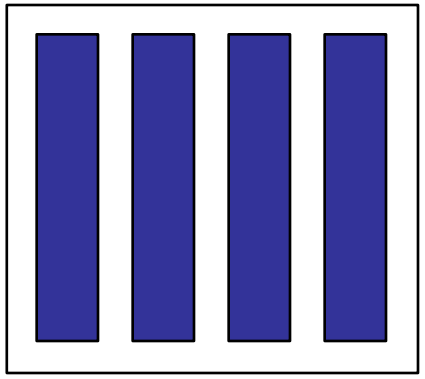
Motivation for New Schemes

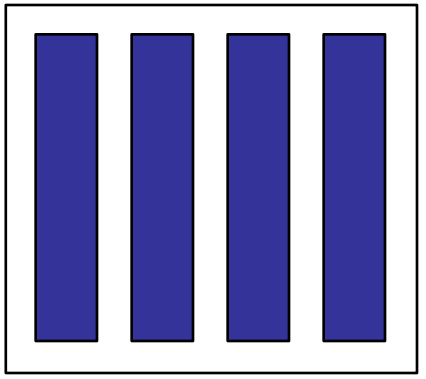
- General rule
  - Delay the use of track(s) in the middle of the bands, e.g., 3rd tracks

Affected Tracks

	Conventional	Novel
25%	no	no
50%	Track1	no
75%	Track 1,2	Track1
100%	Track 1,2,3	Track 1,2,3

Single Band comparison





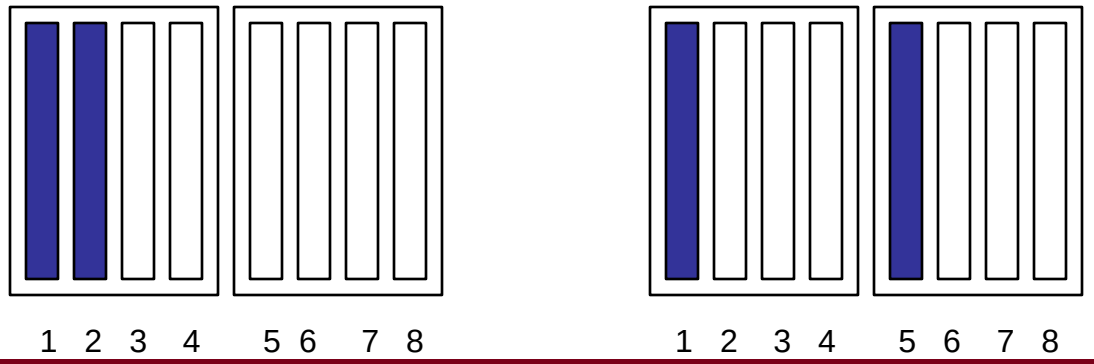
Motivation for New Schemes

- General rule
  - Delay the use of track(s) in the middle of the bands, e.g., 3rd tracks

	Conventional	Novel
25%	Track 1	no

Affected Tracks

Two Bands comparison



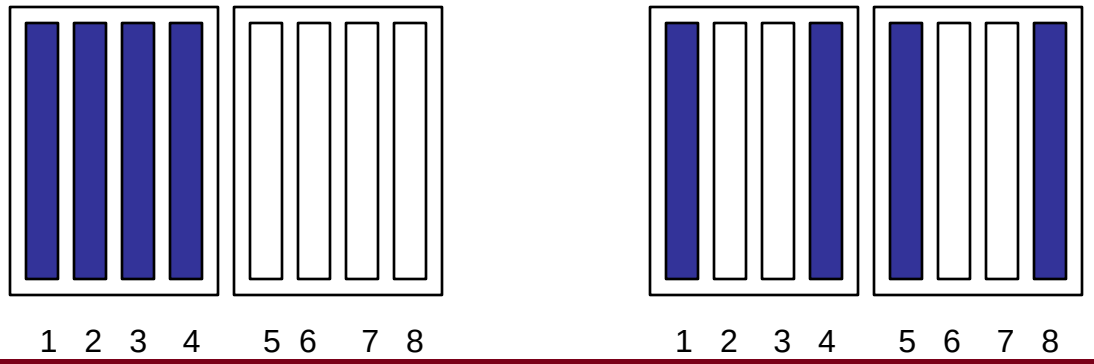
Motivation for New Schemes

- General rule
  - Delay the use of track(s) in the middle of the bands, e.g., 3rd tracks

	Conventional	Novel
25%	Track 1	no
50%	Track 1,2,3	no

Affected Tracks

Two Bands comparison



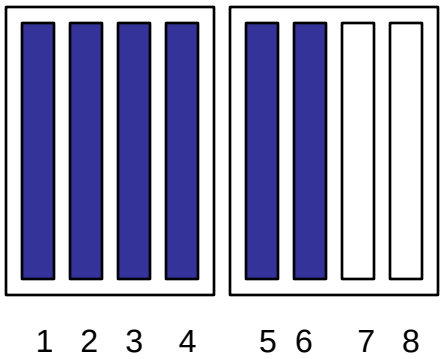
Motivation for New Schemes

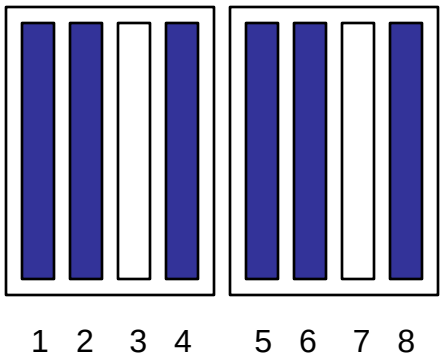
- General rule
  - Delay the use of track(s) in the middle of the bands, e.g., 3rd tracks

Affected Tracks

	Conventional	Novel
25%	Track 1	no
50%	Track 1,2,3	no
75%	Track 1,2,3,5	Track1,5

Two Bands comparison





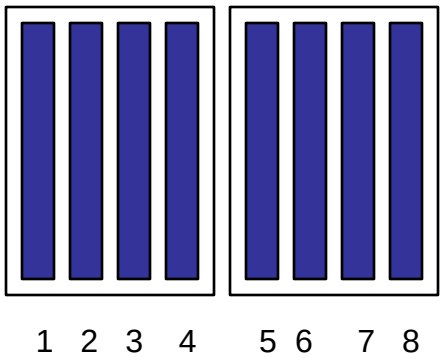
Motivation for New Schemes

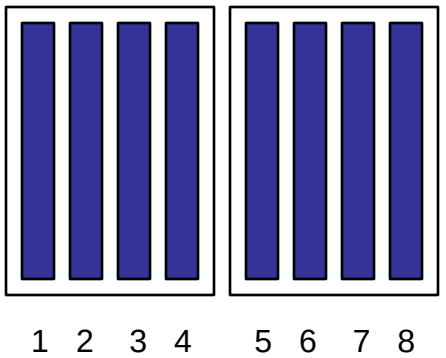
- General rule
  - Delay the use of track(s) in the middle of the bands, e.g., 3rd tracks

Affected Tracks

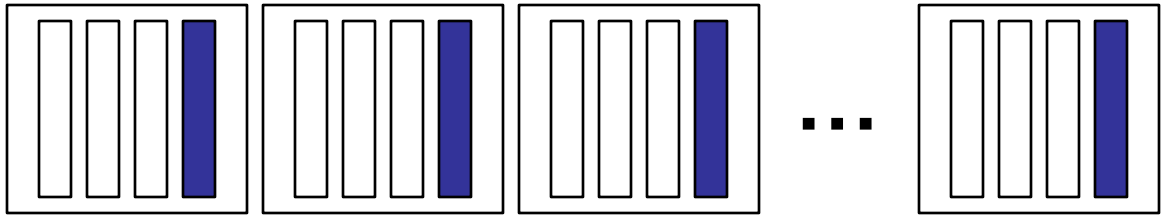
	Conventional	Novel
25%	Track 1	no
50%	Track 1,2,3	no
75%	Track 1,2,3,5	Track1,5
100%	Track 1,2,3,5,6,7	Track 1,2,3,5,6,7

Two Bands comparison



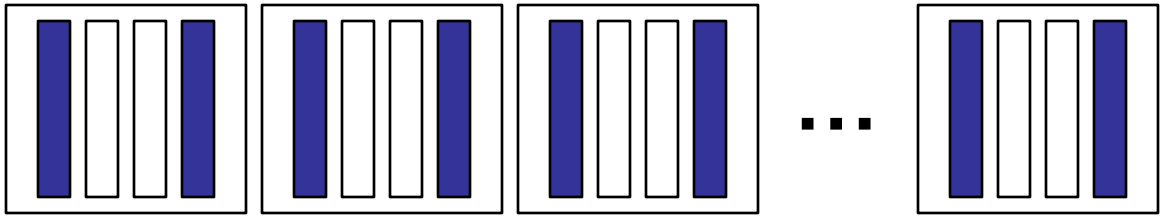


Scheme1: R(4123)



1st 25%: 4th tracks

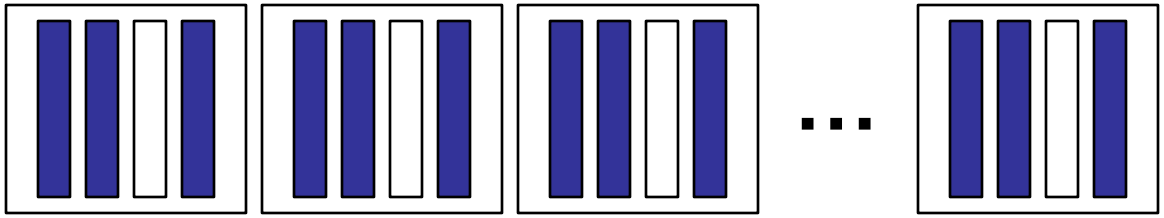
Scheme1: R(4123)



1st 25%: 4th tracks

2nd 25%: 1st tracks

Scheme1: R(4123)

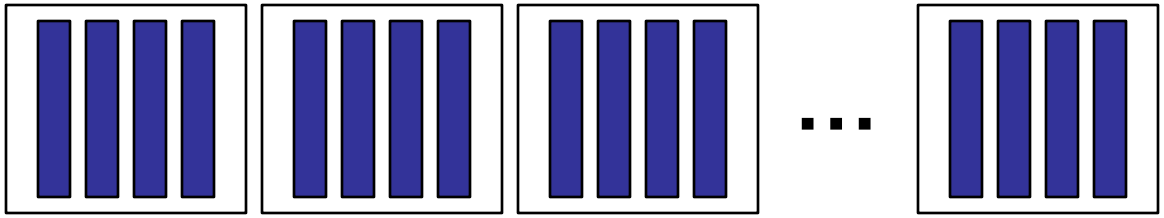


1st 25%: 4th tracks

2nd 25%: 1st tracks

3rd 25%: 2nd tracks

Scheme1: R(4123)



1st 25%: 4th tracks

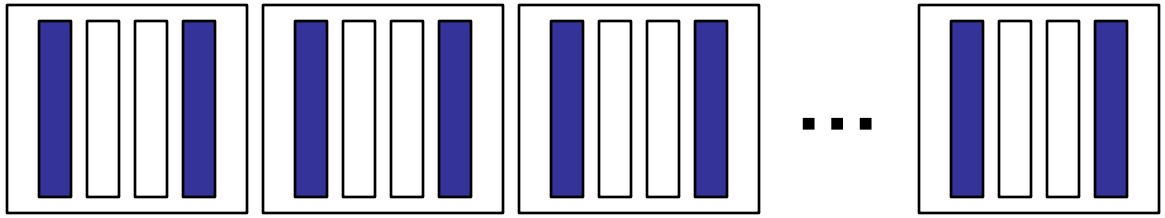
2nd 25%: 1st tracks

3rd 25%: 2nd tracks

4th 25%: 3rd tracks

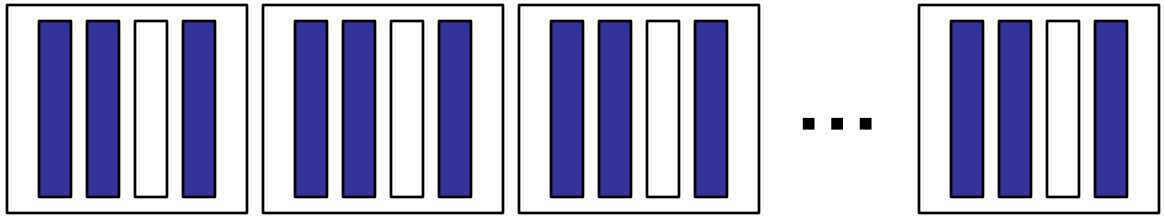
Similarly, R(1423) can be adopted.

Scheme2: 14R(23)



1st 50%: 1st and 4th tracks

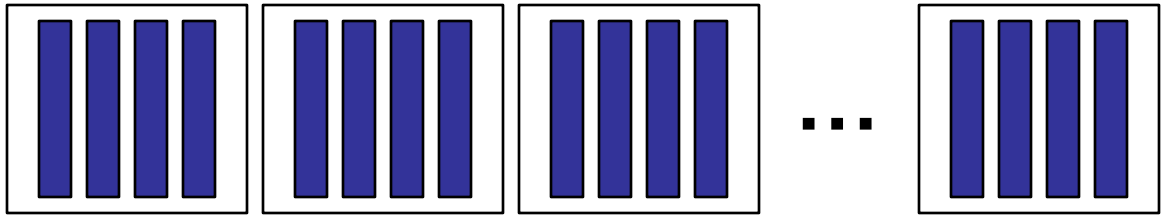
Scheme2: 14R(23)



1st 50%: 1st and 4th tracks

The following 25%: 2nd tracks

Scheme2: 14R(23)

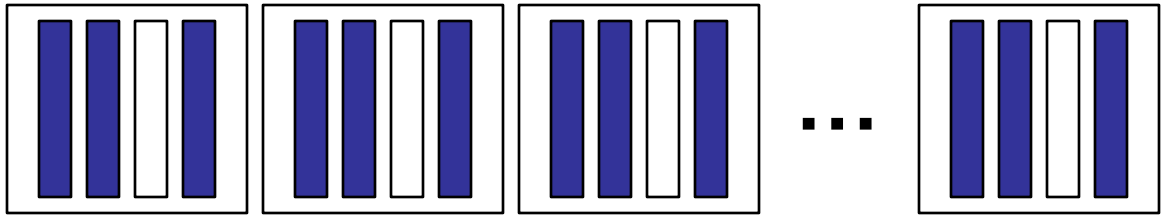


1st 50%: 1st and 4th tracks

The following 25%: 2nd tracks

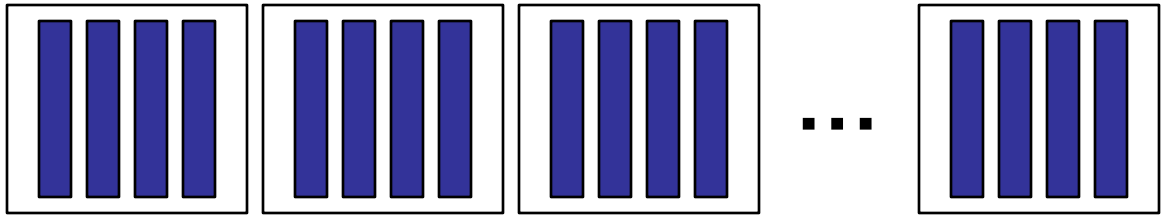
The last 25%: 3rd tracks

Scheme3: 124R(3)



1st 75%: 1st, 2nd and 4th tracks

Scheme3: 124R(3)



1st 75%: 1st, 2nd and 4th tracks

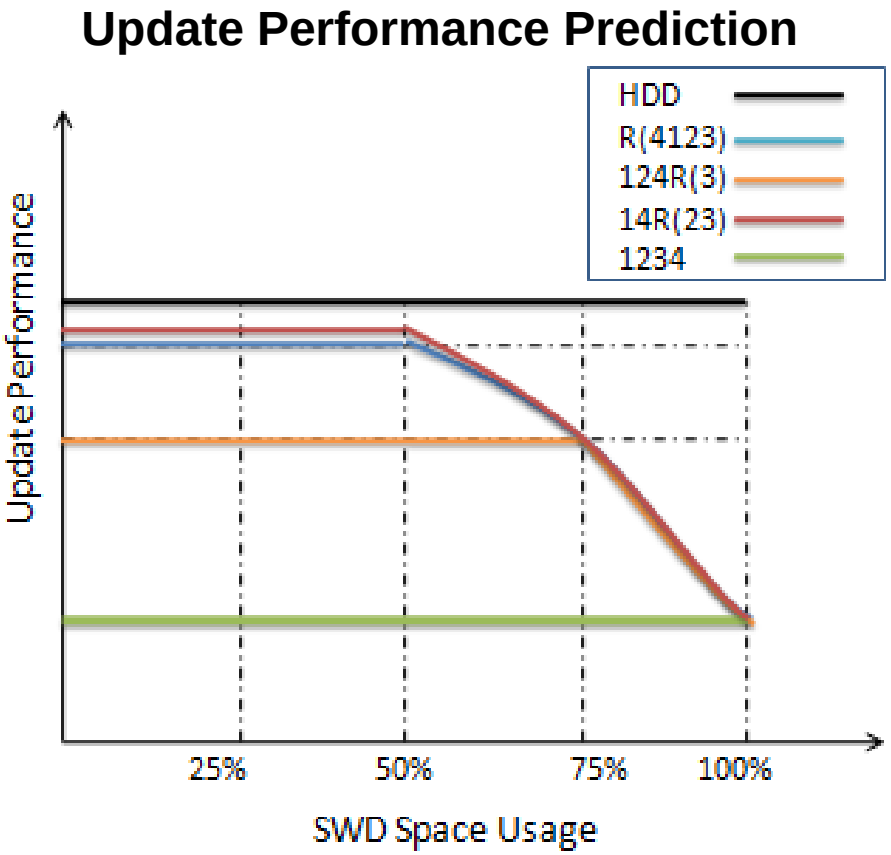
The last 25%: 3rd tracks

Scheme Comparisons

- SWD setup:
  - W = 2
  - N = 4

scheme	Spatial Locality	Write Amplification
HDD	5	5
R(4123)	2	4
14R(23)	3	3
124R(3)	4	2
1234	5	1

5 is best, 1 is worst



Evaluation Design (1/2)

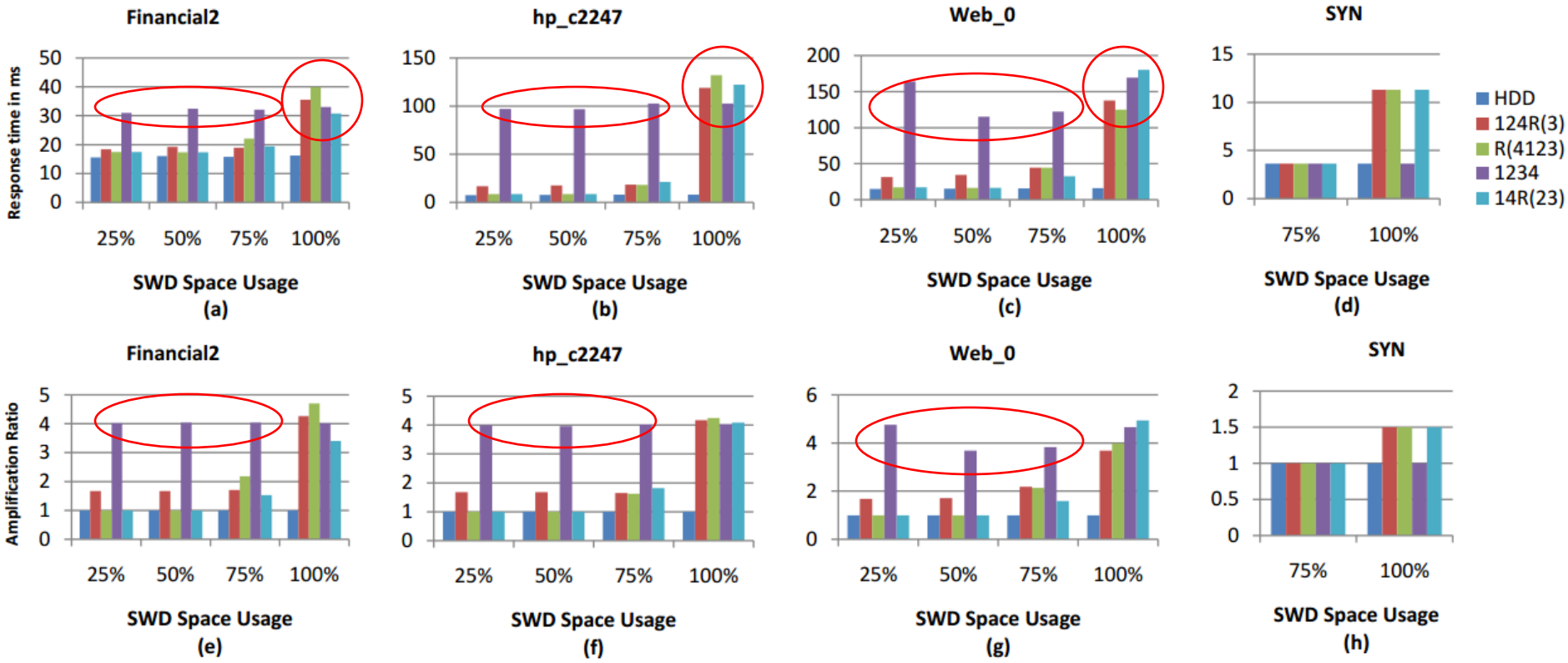
- SWD simulation
  - Disksim, with hp_c3323a disk model
  - **Address mapper**: translate LBAs to PBAs
  - **Write amplifier**: convert an update into several reads/writes accordingly
- SWD setup:
  - 3000 cylinders
  - 1000 blocks per cylinder
  - Band size = 4
  - Write width = 2

Trace	Write %	Notes
Web_0	0.70123	Update intensive
Financial_0	0.096978	Update light
hp_c2247	0.488449	Update moderate
SYN	1	Sequential write, average size 8 blocks, IAT=(mean 50ms, std. dev 10ms). No update.

Evaluation Design (2/2)

- Performance test points:
  - 25%, 50%, 75%, 100%
- For each test point, “pre-fill” the SWD space to the corresponding percentage to logically convert writes to updates
  - This is done to pass the percentage to the *write amplifier*

Results


Financial2
hp_c2247
Web_0
SYN
Response time in ms
Amplification Ratio
HDD
124R(3)
R(4123)
1234
14R(23)
25%
50%
75%
100%
SWD Space Usage
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)

Summary

- Shingled magnetic recording drives
- Write amplification problem
- Achieve good space gain and performance balance with new static address mappings
  - R(4123) or R(1423)
  - 14R(23)
  - 124R(3)

Future Work

- Comparison to out-of-place update SWD
- SWD file system designs
- Construct storage system with SWDs, e.g., RAID and erasure codes
- Hybrid SWDs

Questions?