Log-structured Block Device Driver for Shingled Disk

Lichao Yin, Lei Kang and Remzi Arpaci-Dusseau
University of Wisconsin-Madison

Shingled Disk

- ❑ Read head can be narrower but write head cannot (more strength need to flush the data onto disk)
- ❑ More disk space as the tracks are packed together (25% more disk space)
- ❑ Shingled Disk is divided into bands, no overlap between bands
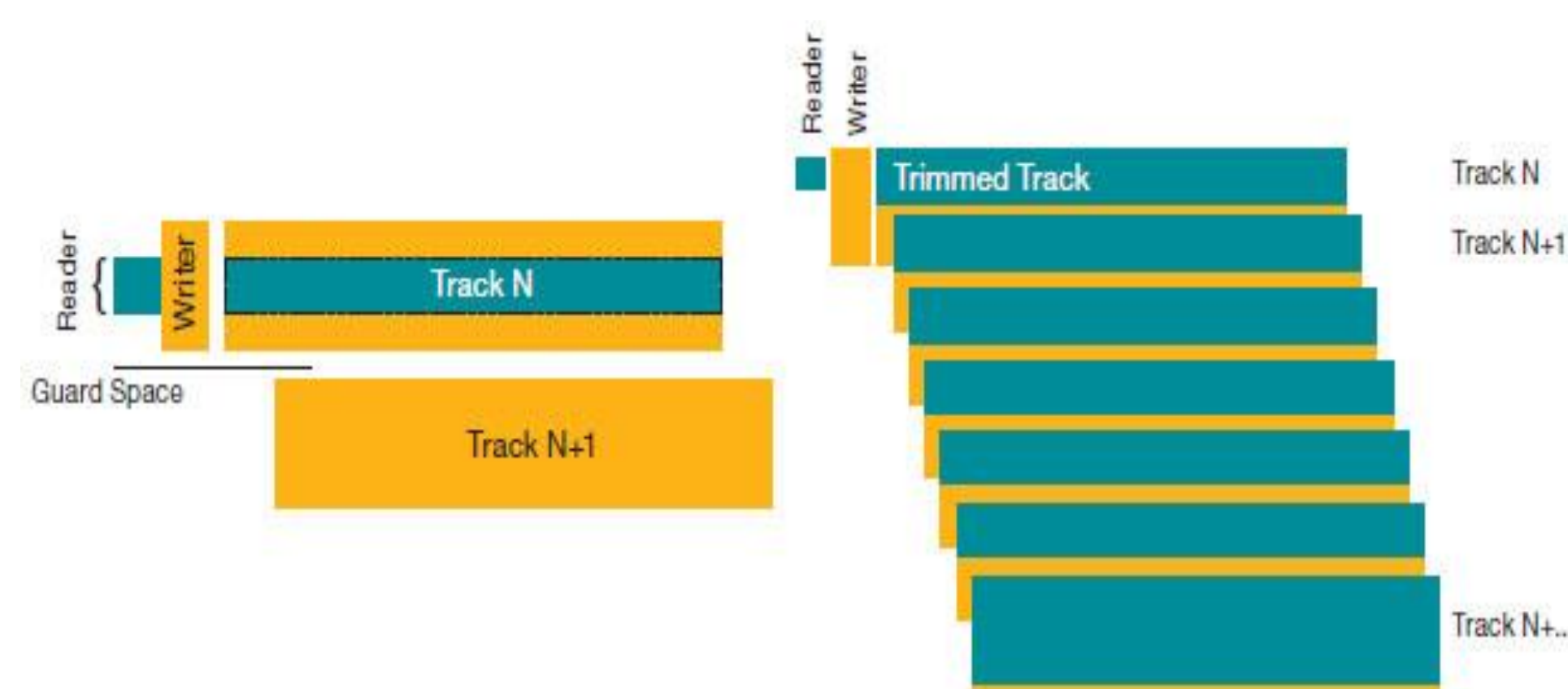- ❑ Each band consists of many tracks, write to one track may overwrite the data in downward tracks

Log-structured Driver

- ❑ Only appending on empty band, avoid overwrite existing data
- ❑ No changes to existing file systems
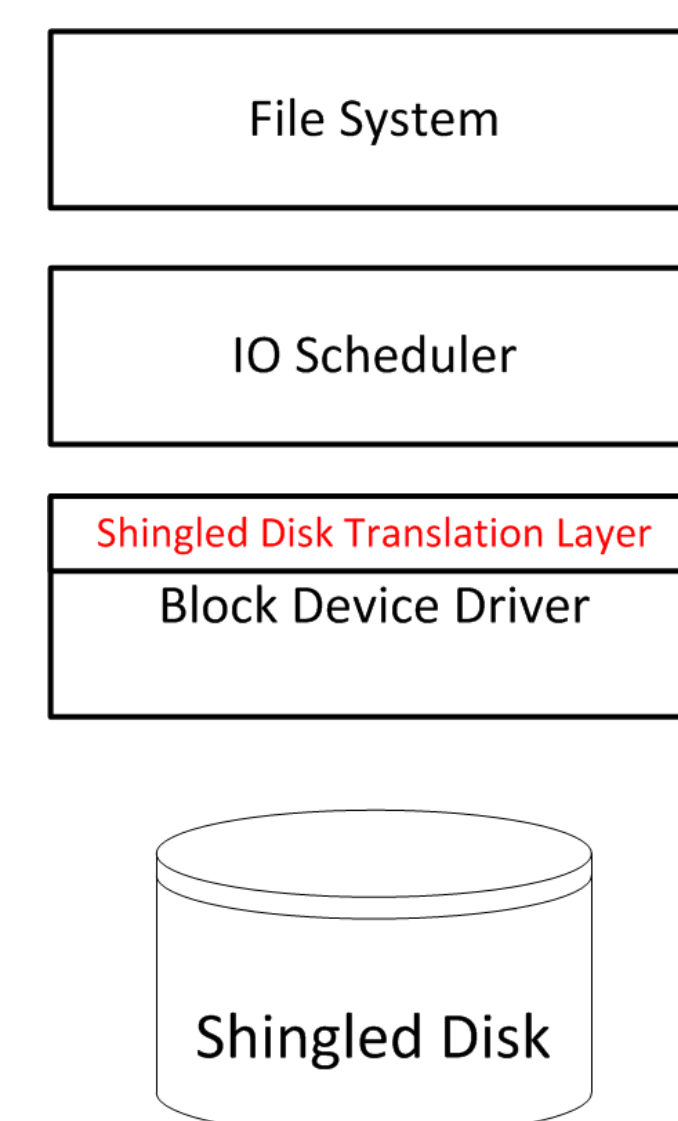- ❑ Address translation and garbage collection (or clean-up) are required

Address Translation

- ❑ Logical Sectors
  - ❑ Flat address space what the file system see
  - ❑ Map to the disk sectors
- ❑ Disk Sectors
  - ❑ Track disk sector status
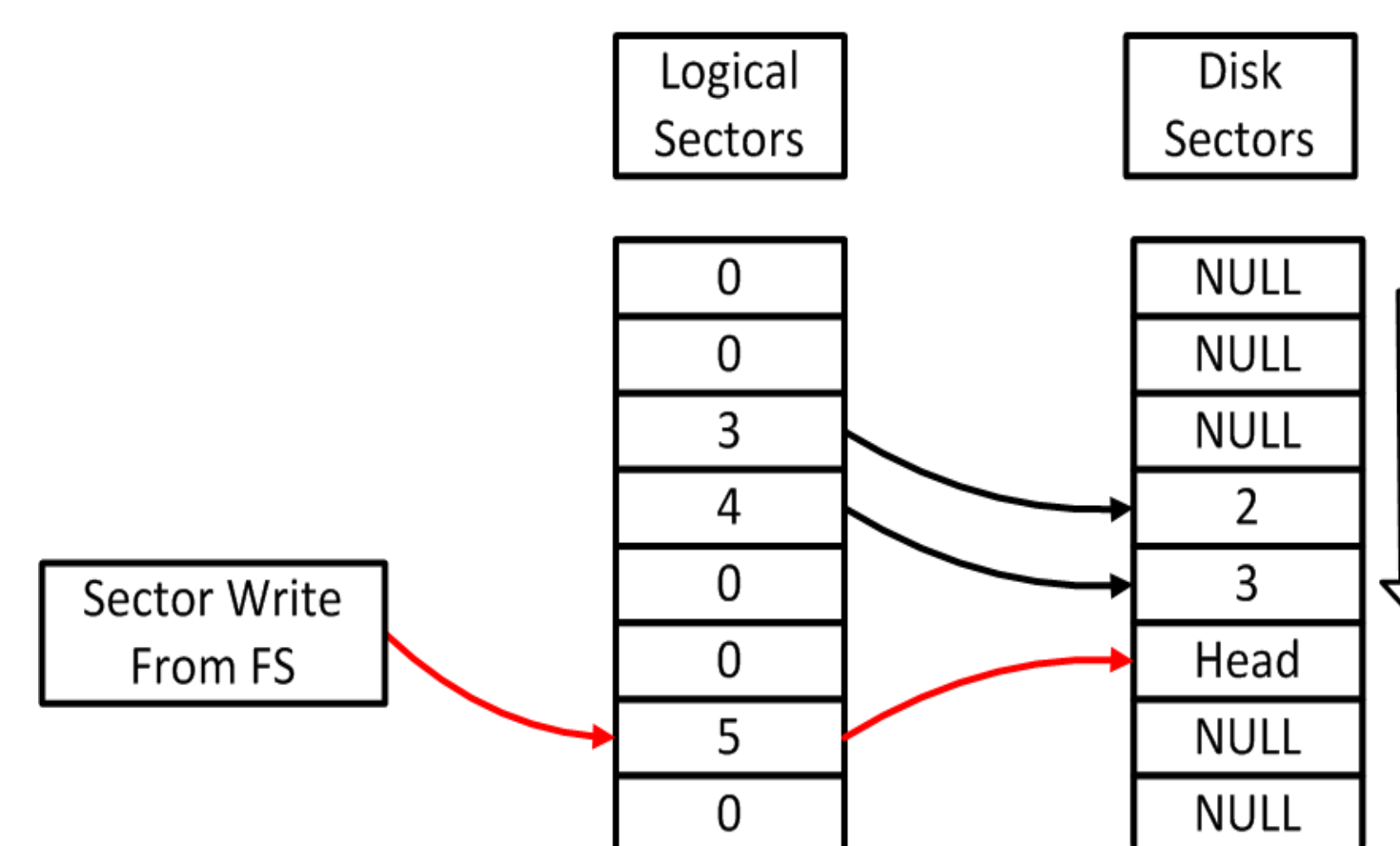 - ❑ Map back to the logical sectors for clean-up
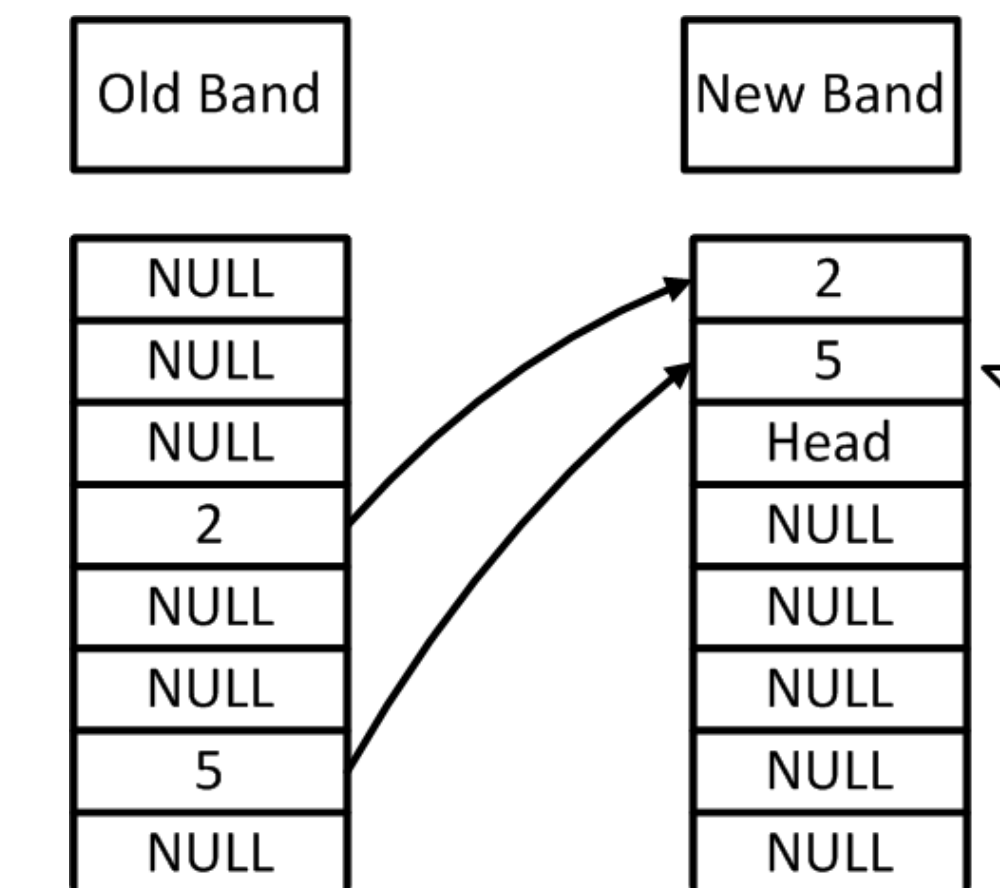
Traditional Disk vs. Shingled Disk
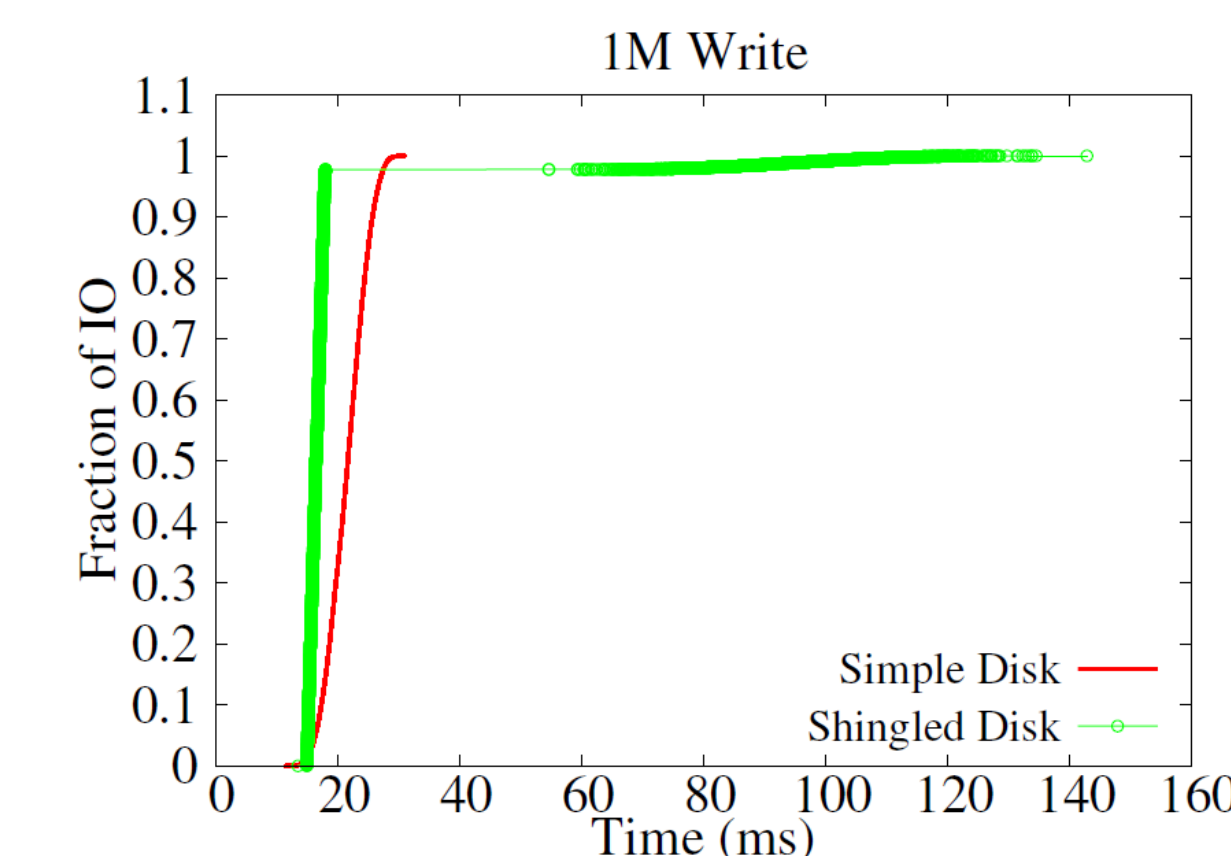

Translation Layer

File System

IO Scheduler

Shingled Disk Translation Layer
Block Device Driver

Shingled Disk

Sector Address Translation

Logical Sectors → Disk Sectors

Logical Sectors	Disk Sectors
0	NULL
0	NULL
3	NULL
4	2
0	3
0	Head
5	NULL
0	NULL

Sector Write From FS

Band Cleanup

Old Band	New Band
NULL	2
NULL	5
NULL	Head
2	NULL
NULL	NULL
NULL	NULL
5	NULL
NULL	NULL

Write Performance

Band Clean-up

- ❑ New (Clean) Band:
  - ❑ Assume there is at least one clean band
- ❑ Old Band:
  - ❑ Least used band
- ❑ Read sectors from old band into memory
- ❑ Flush all the valid data into new (clean) band

Disk Model From David

- ❑ Seek time: $a + b * \sqrt{tracks}$
- ❑ Rotate time: $sectors * us\_per\_sector$
- ❑ Transfer time:
  $tracks * switch\_time + us\_per\_sector * sectors$
- ❑ Bus time: $data/bus\_speed$

Evaluation

- ❑ Random Write is Faster in 99% cases, slow on band clean-up
- ❑ Random Read is slightly slower, due to mapping on each sector

Table 1: Shingled Disk vs. Simple Disk

Disk type	Random 4k Read	Random 4k Write	Random 1M Read	Random 1M Write
Simple Disk	9.94ms	9.92ms	21.61ms	21.61ms
Shingled Disk	11.45ms	0.14ms	31.39ms	18.23ms