SANTA CLARA UNIVERSITY

The Peril and Promise of Shingled Disk Arrays

(how to avoid two disks being worse than one)

Quoc M. Le JoAnne Holliday Ahmed Amer
Department of Computer Engineering, Santa Clara University, Santa Clara, CA
{qle, jholliday, aamer}@scu.edu

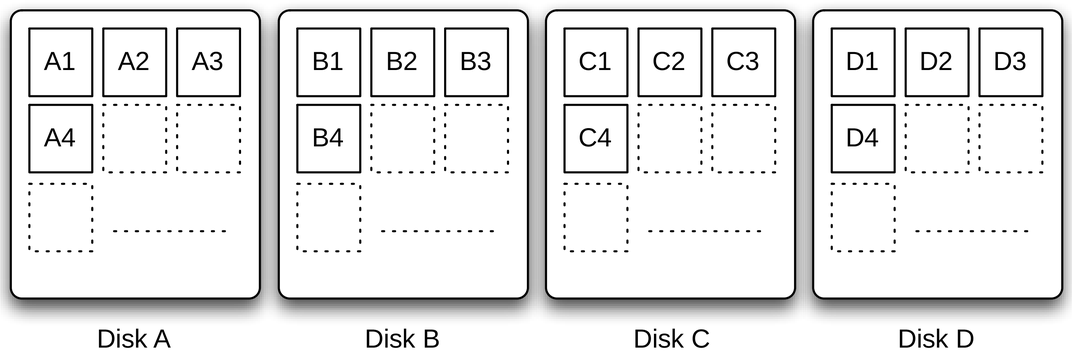
Evaluating the behavior of shingled disks when used in an array configuration or when faced with heavily interleaved workloads from multiple sources.

Initial Findings

- Heavily interleaved workloads can have a dramatic negative impact on disk activity.
- Reducing interleaving has a significant positive effect.

➢ **Disk Layout Options** (in lieu of basic striped arrays):

- Dedicated disks and bands
- Workload differentiation

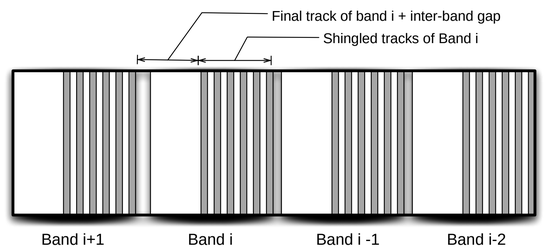


Logical view of a simple array of disks. In the striped arrangement, blocks 0, 1, and 2 are arranged as A1, B1, and C1. In pure arrangements, blocks 0, 1, and 2 are arranged as A1, A2, and A3.

Shingled Disks and Arrays

➢ **Shingled magnetic recording (SMR) potential:**

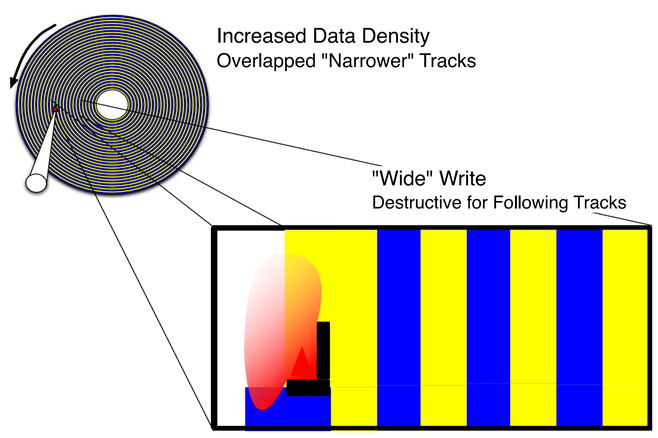
- Current disks offer recording densities of 400Gb/in^2
- With SMR, 1Tb/in^2 is possible [3]

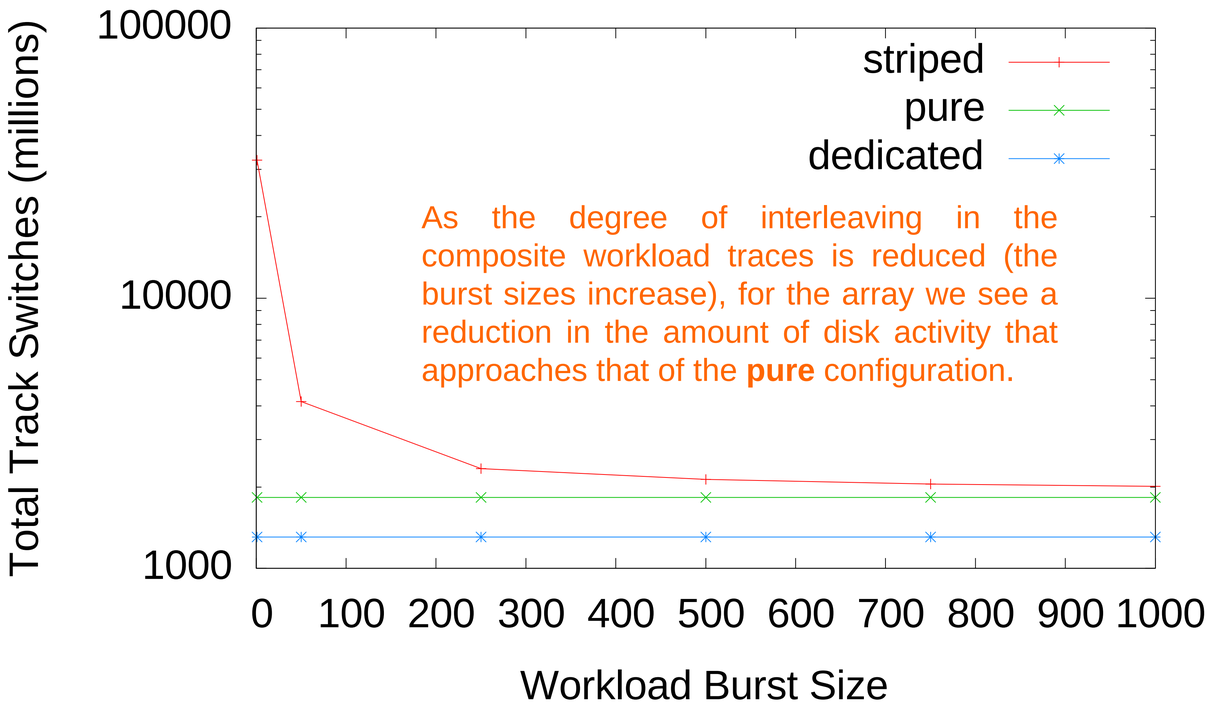


Logical view of a shingled write disk divided into bands, allowing the in-place update of a band, although at the expense of a destructive track write within an individual band.

➢ **SMR functional considerations:**

- Consecutive tracks are overlapped.
- Updates to individual tracks result in overwrites to any overlapped tracks.
- Overlapped tracks can be grouped into bands or logs [1].
- Log-structuring approaches can defer the need to update in-place [2]
- Alternative design parameters, interface models, and file systems-based solutions are possible [1].





Disk activity when replaying multi-source traces against a simulated array of shingled write disks. Based on a shingled write disk utilizing a log-structured write scheme to minimize in-place updates.

Workload-Based Evaluation

➢ **Striped workload**

- using composite of four workloads
- workload mix varied by adjusting a random interleave

➢ **Pure workload**

- total activity across four disks
- disks arranged in sequence
- time-varying workload, but not interleaving

➢ **Dedicated workload**

- total activity across four disks
- disks dedicated to individual workload sources
- unlike "pure" and "striped" workloads: no interleaving per-disk

References

[1] A. Amer, D. E. Long, E. L. Miller, J.-F. Paris, and T. Schwarz, “Design issues for a shingled write disk system,” Proceedings of IEEE MSST, 2010.

[2] Y. Casutto, M. Sanvido, C. Guyot, D. Hall, and Z. Bandic, “Indirection systems for shingled-recording disk drives,” Proceedings of IEEE MSST, 2010.

[3] I. Tagawa and M.Williams,“High density data-storage using shingle-write,” Proceedings of IEEE INTER-MAG, 2009.