



SANTA CLARA UNIVERSITY
THE JESUIT UNIVERSITY IN SILICON VALLEY



Design Issues for a Shingled Write Disk System

Ahmed Amer, Darrell Long, Ethan Miller,
Jehan-Francois Paris, Thomas Schwarz



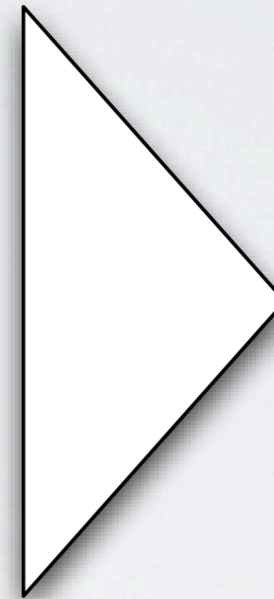
Baskin
Engineering
UC SANTA CRUZ



MEDIA TRILEMMA*

- *Tradeoff dictated by the Superparamagnetic Limit*
- *Three-way Tradeoff*
- *Results in a challenging limit for writes, not reads*

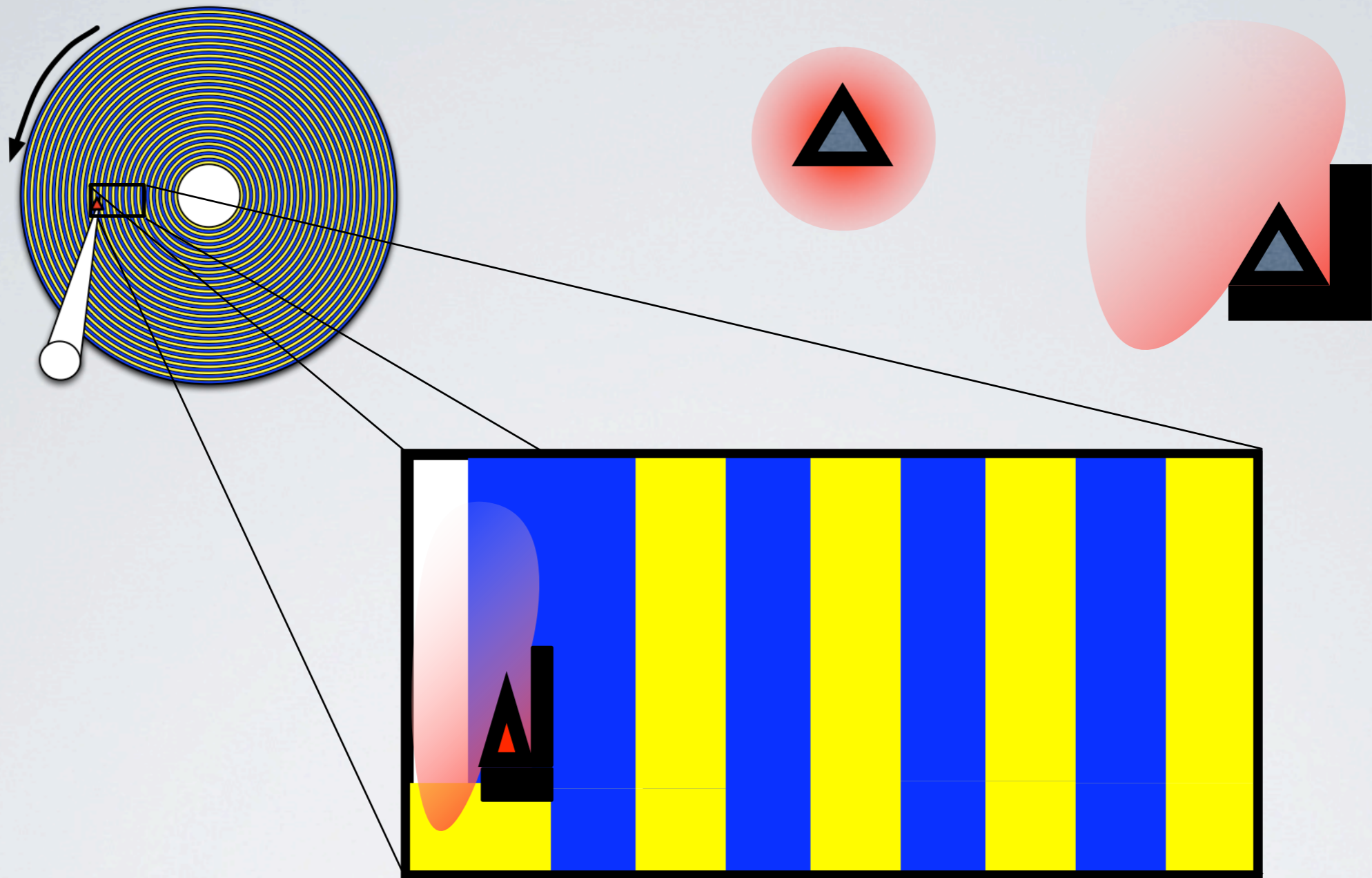
Thermal Stability



SNR/Definition

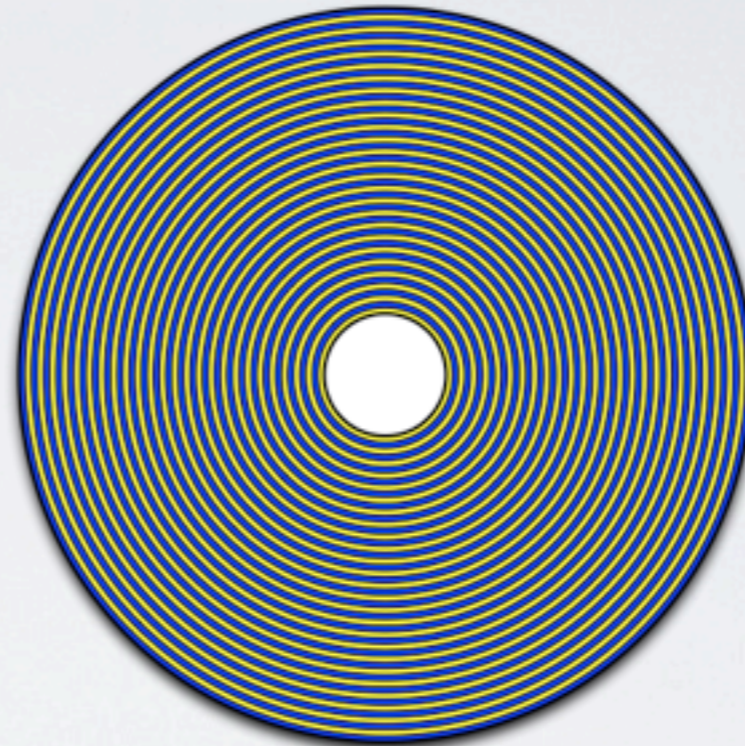
Media Writability

**dubbed by Sann et al.*



SHINGLED WRITING

Thermal Stability vs. Writability



SHINGLED WRITING

Thermal Stability vs. Writability

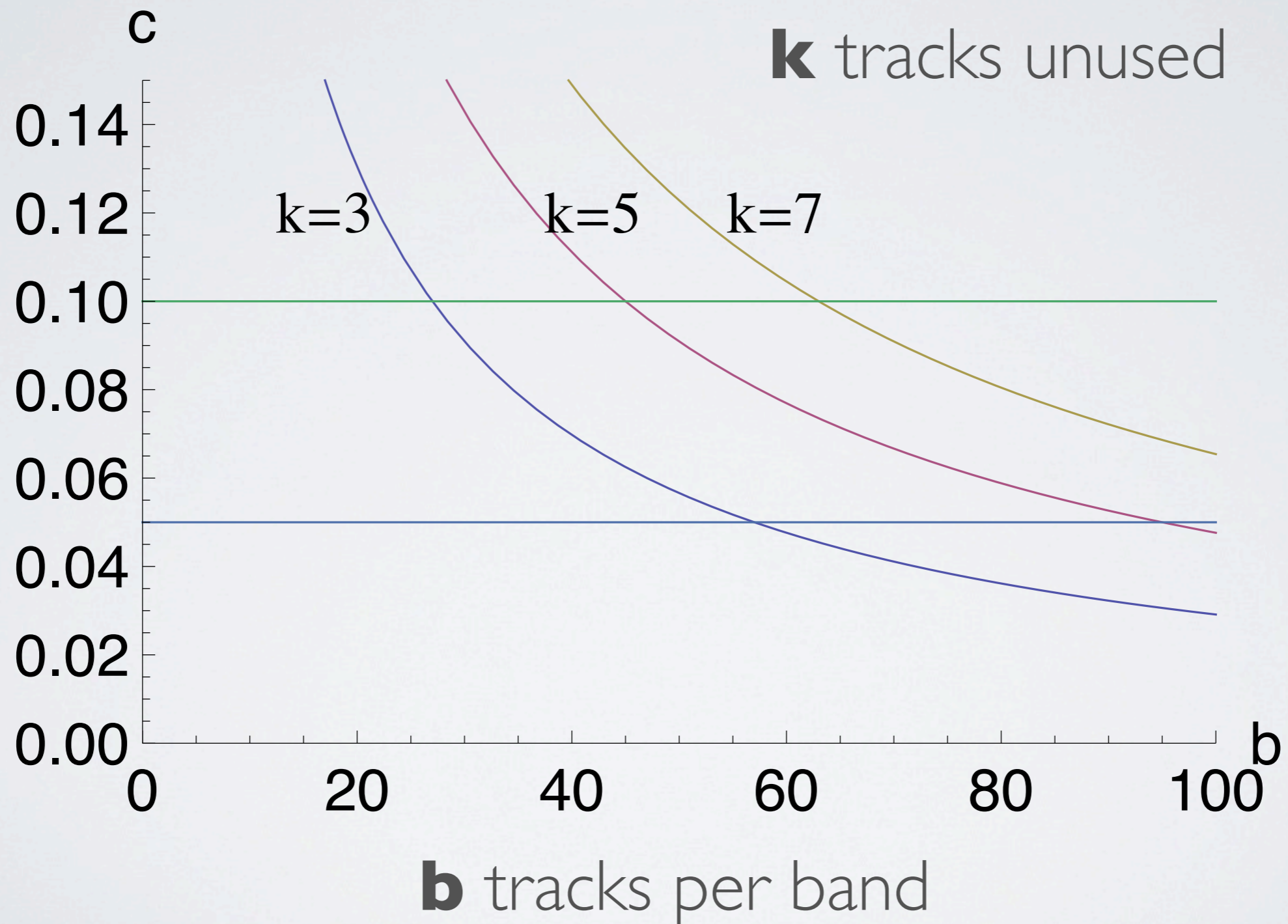
SHINGLED DISK PROJECTIONS

	Seagate 7200.11 (1.5TB SATA Drive)	Shingled Disk Projections
Density (Gb/in ²)	277	2000
BPI (x1000 b/in)	1462	2300
TPI (x1000/in)	192	860

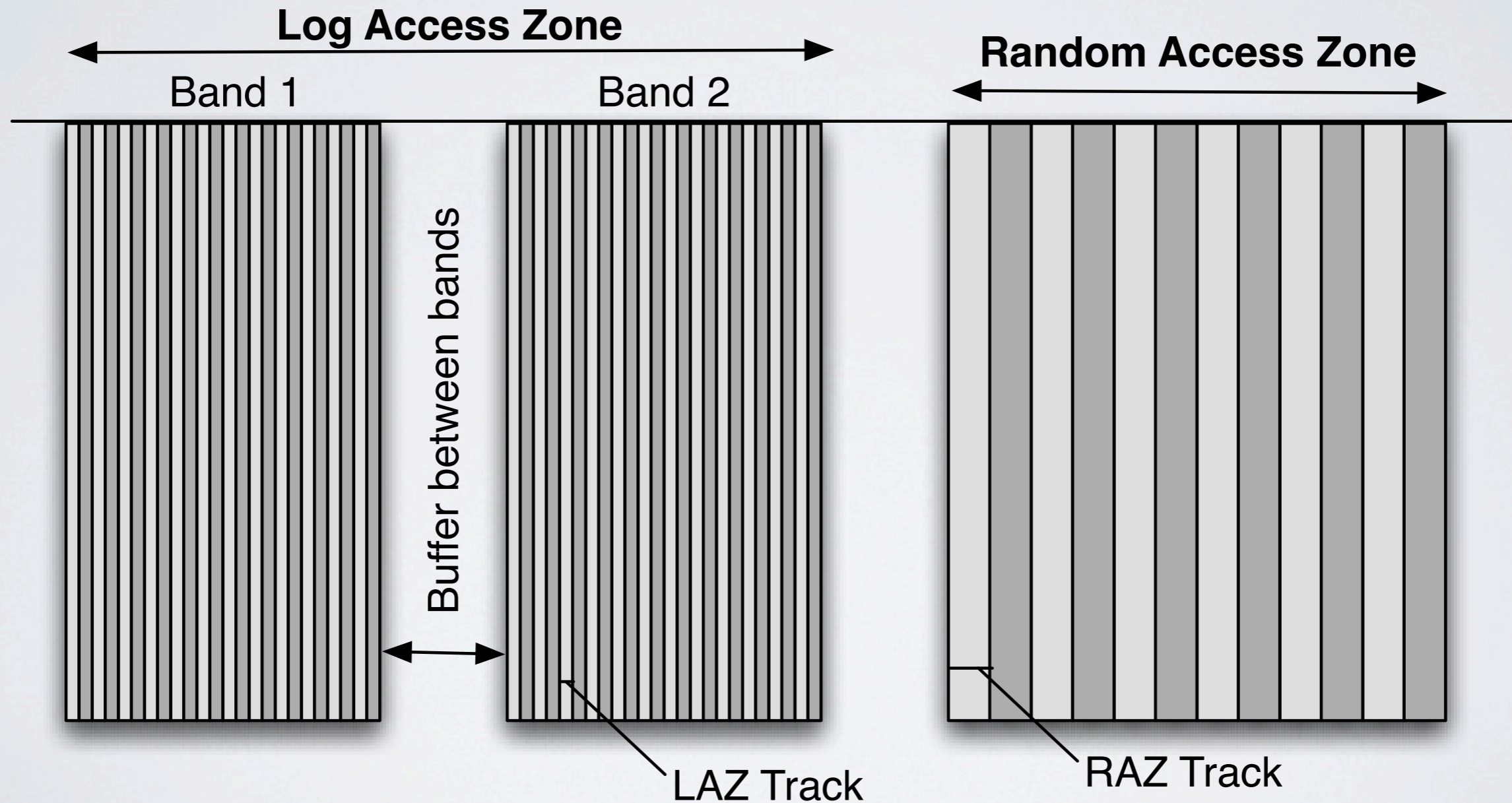
LOGICAL VIEW

- Append-only or destructive overwrite
- Divide into **bands**
- Allowing for an intra-band gap of **k** tracks
 - allows update-in-place of the band
- With adjustable bands, an SWD is logically similar to
 - tape with very flexible partition sizes
 - flash with variable bank sizes

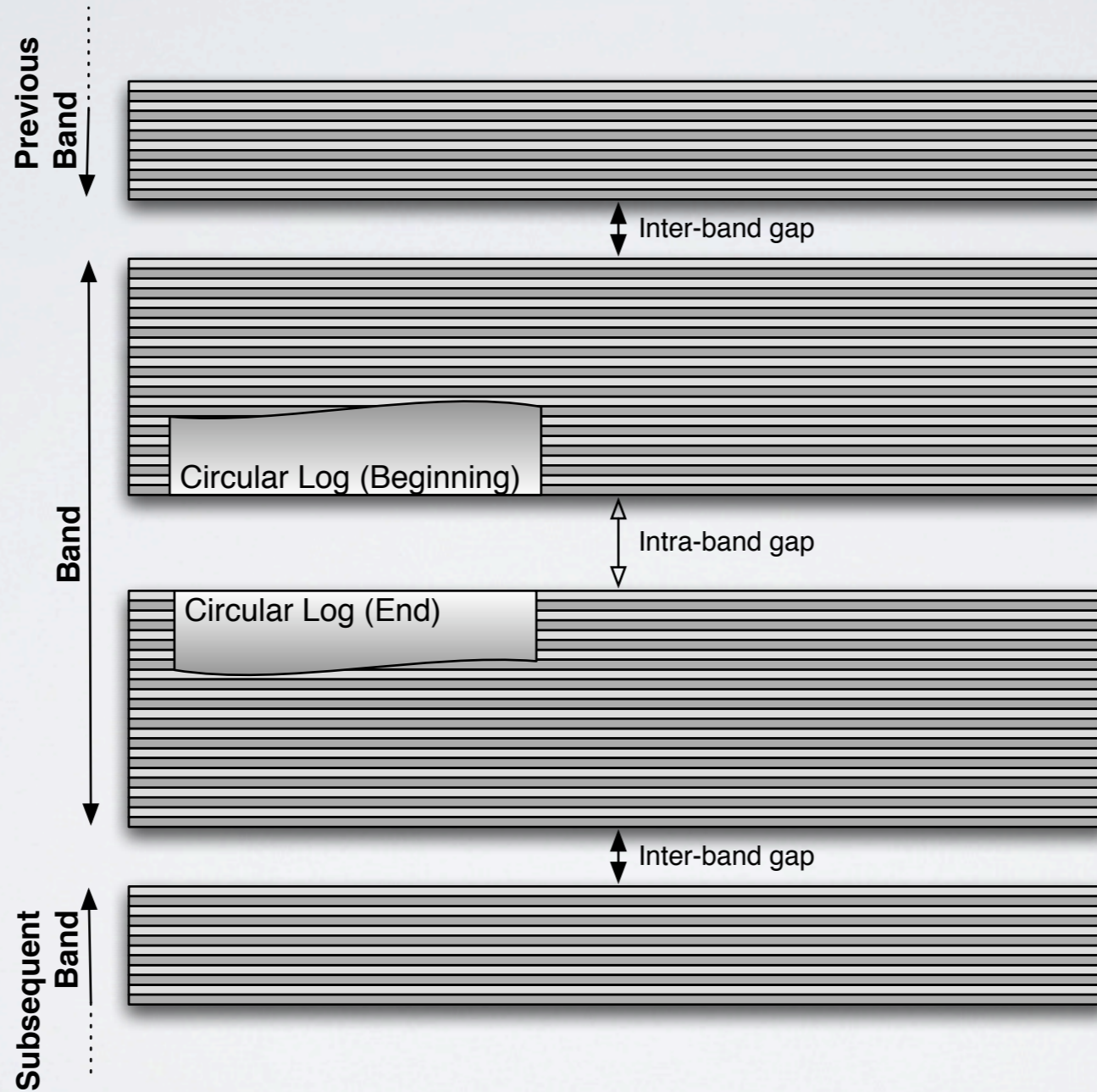
PROPORTIONAL CAPACITY LOSS



LOG-ACCESS & RANDOM ACCESS ZONES

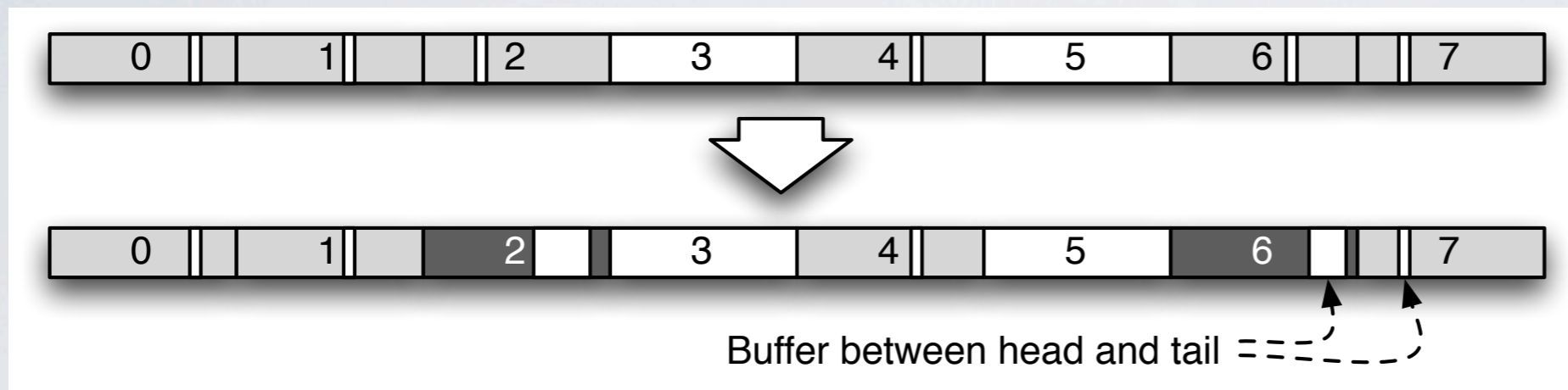


CIRCULAR LOG

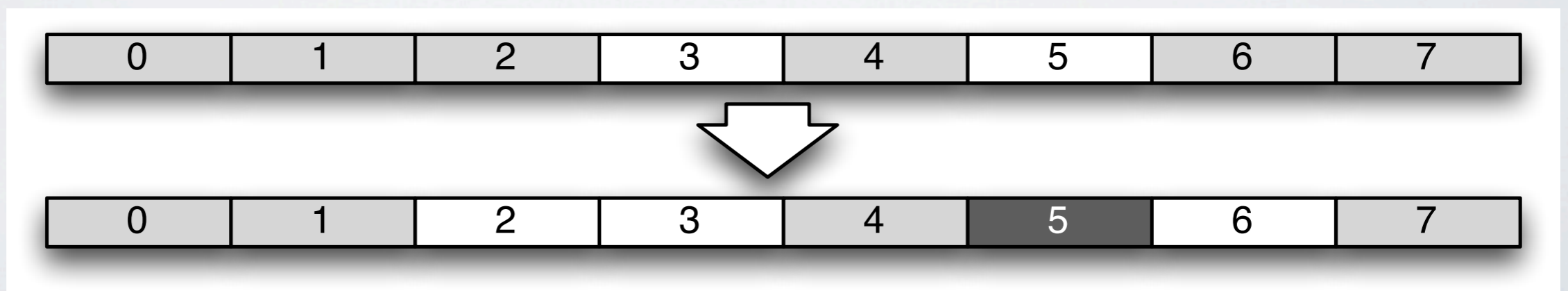


LOGS VS SEGMENTS

Example: Compacting Bands 2 & 6



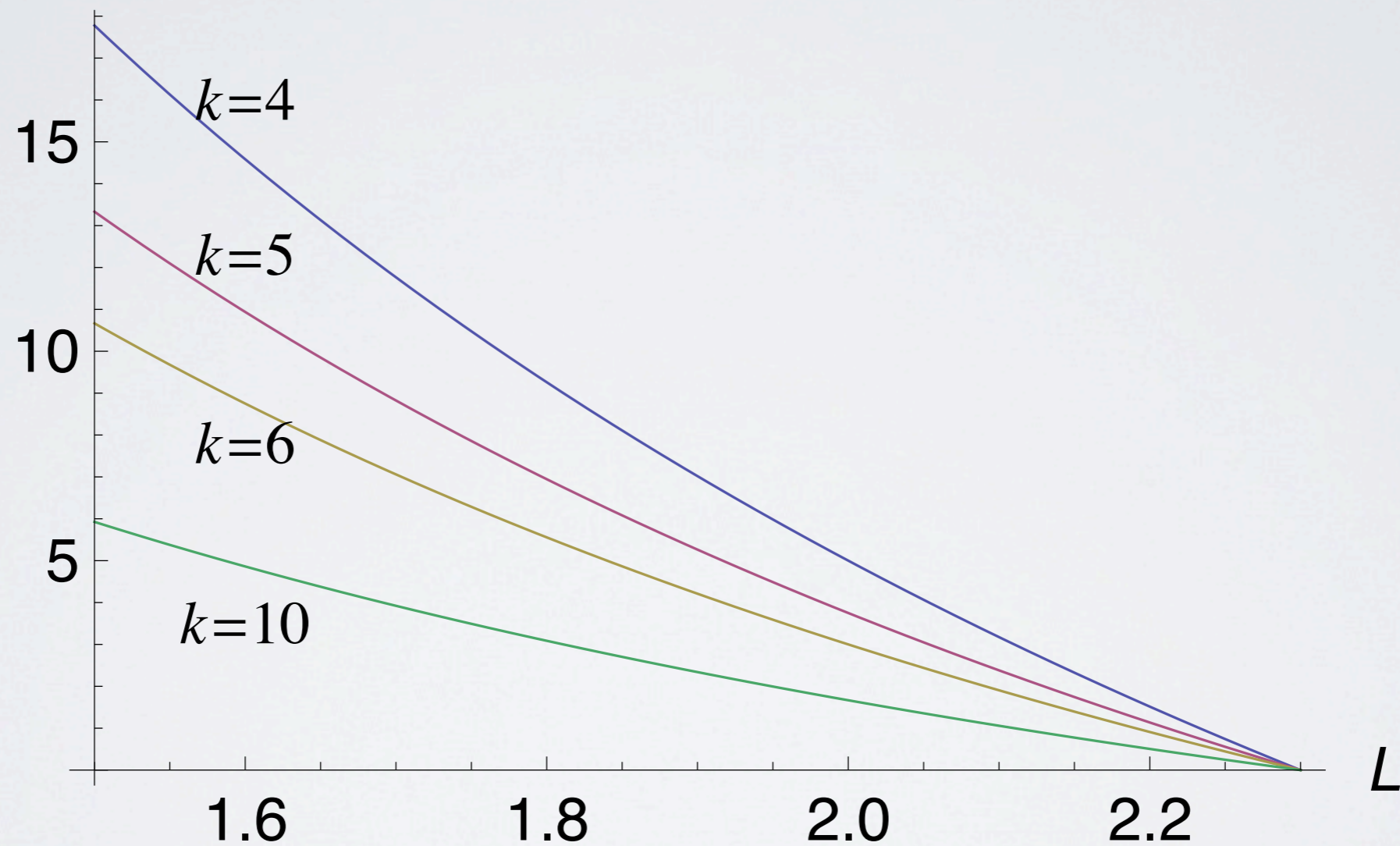
In-band compaction with circular logs



Compaction of bands to new segments

RAZ CAPACITY OVERHEAD

% Capacity RAZ



L-fold capacity increase (assuming $\alpha=2.3$)

GENERAL PURPOSE USE

- Redirection and NVRAM-augmentation - workload dependent
- Desirable Workload Properties
 - Low update frequency
 - Low update volume
 - Minimal **re**writes (updates of previously written data)
- Expected for archival and similar workloads, but general purpose use?

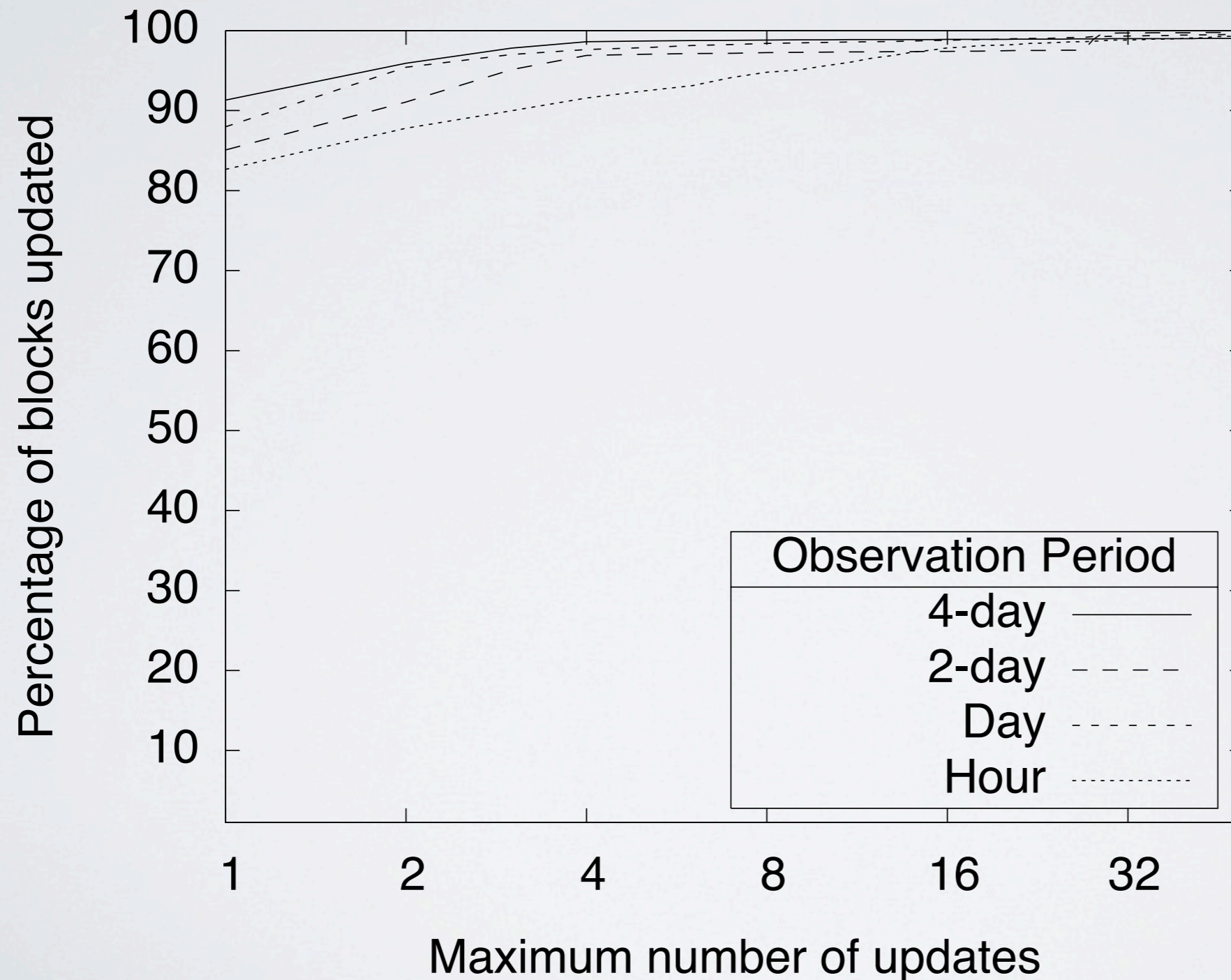
SAMPLE WORKLOADS

- OSX 10.4, HFS+
- Collected from student and faculty laptops
- November and December of 2007
- Continuous collection
 - interruptions due to occasional system reboots

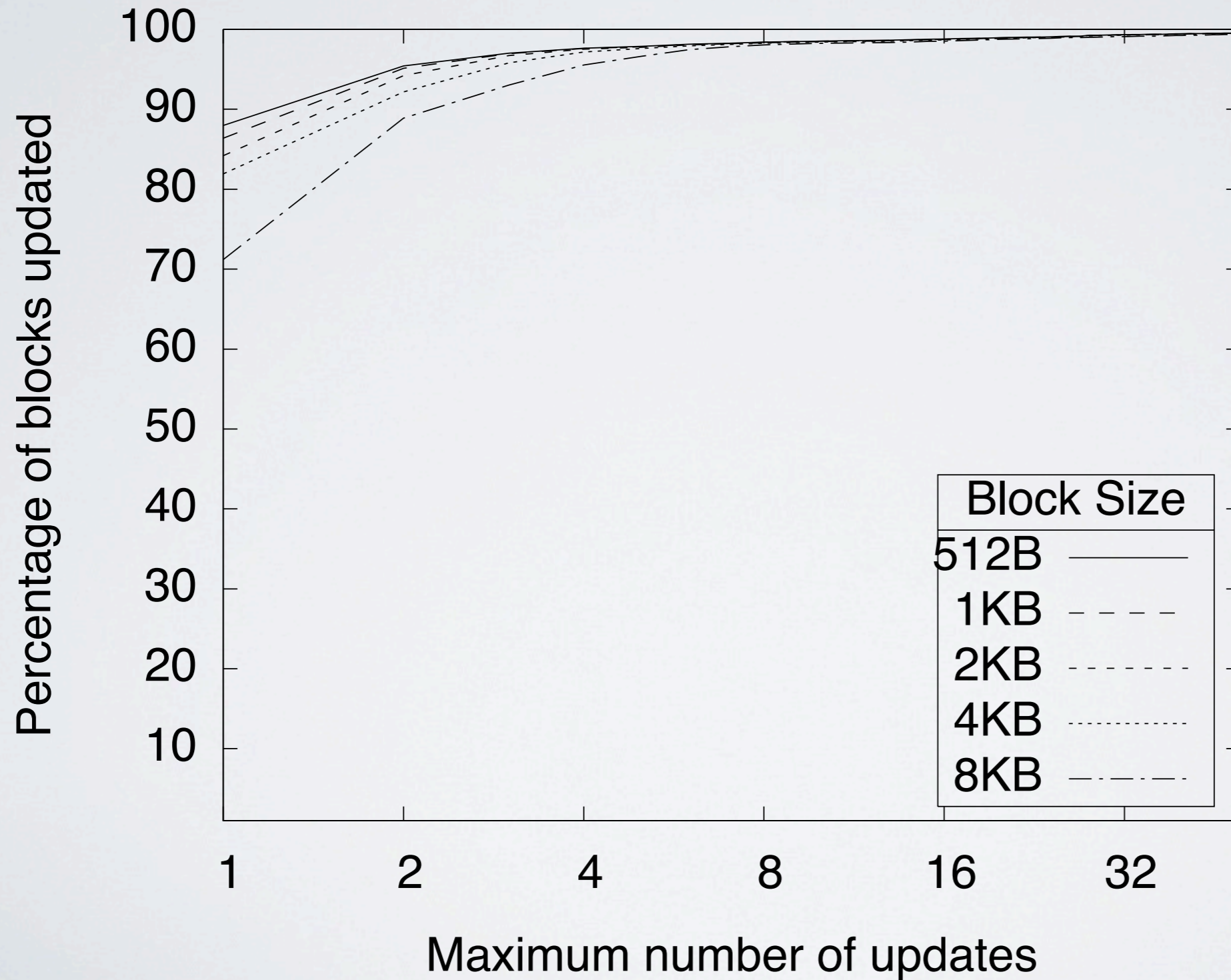
WORKLOAD ANALYSIS

- Evaluated the relative volume of updated blocks (not timing dependencies)
- Varied
 - Observation periods
 - Block sizes tracked
 - Separated Data and Metadata updates
 - Primary application
- Individual process workloads unseparated

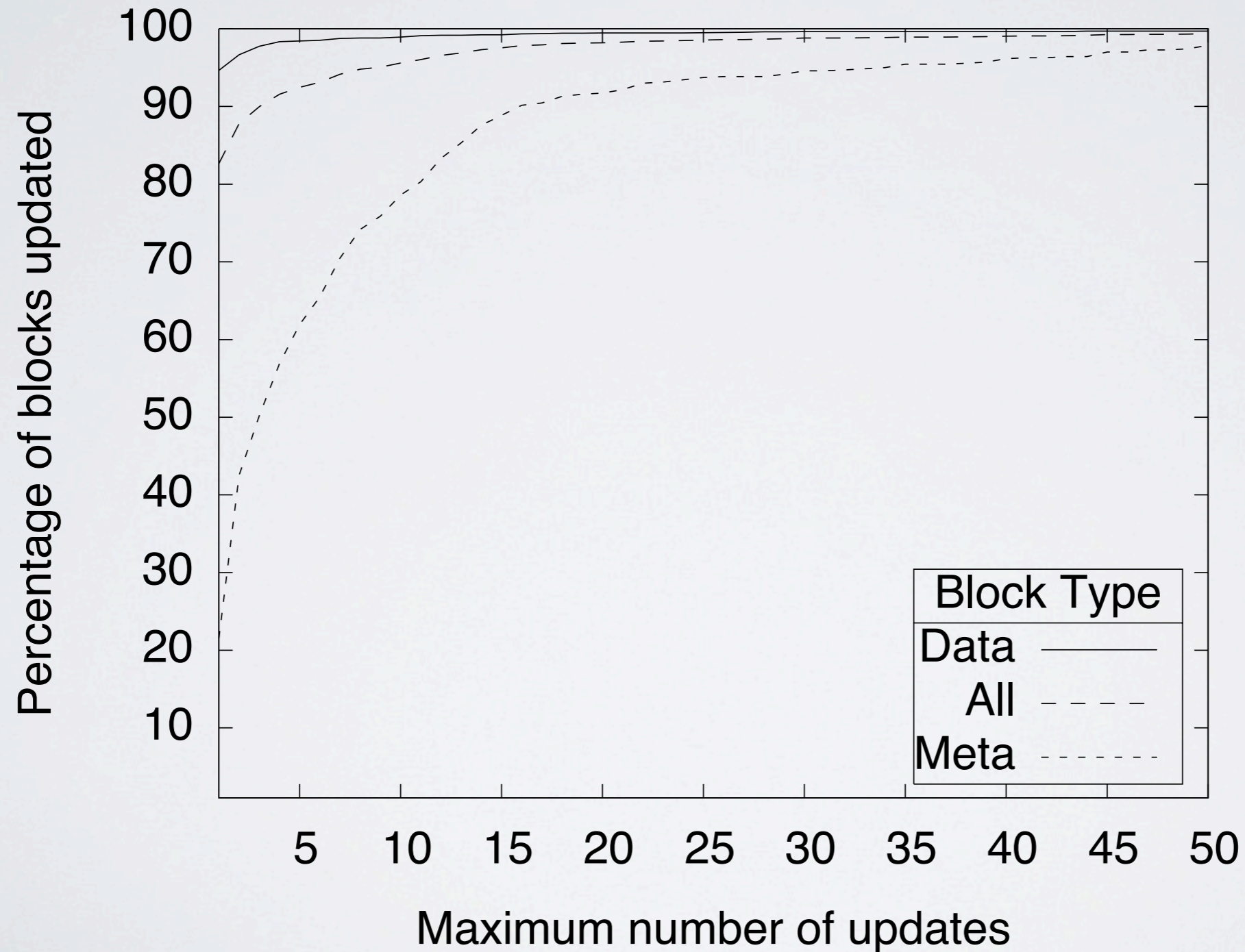
OBSERVATION PERIOD



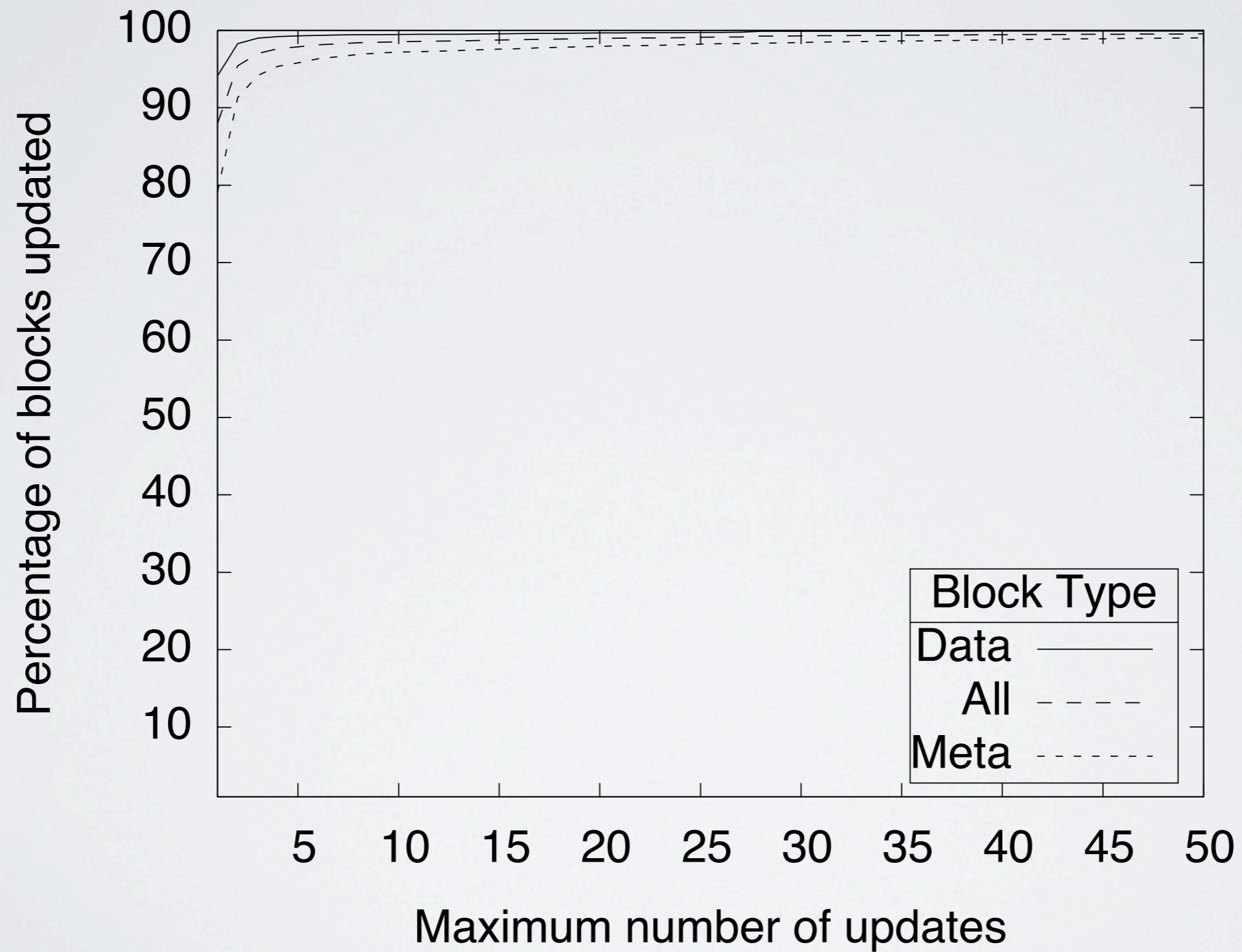
BLOCKSIZE



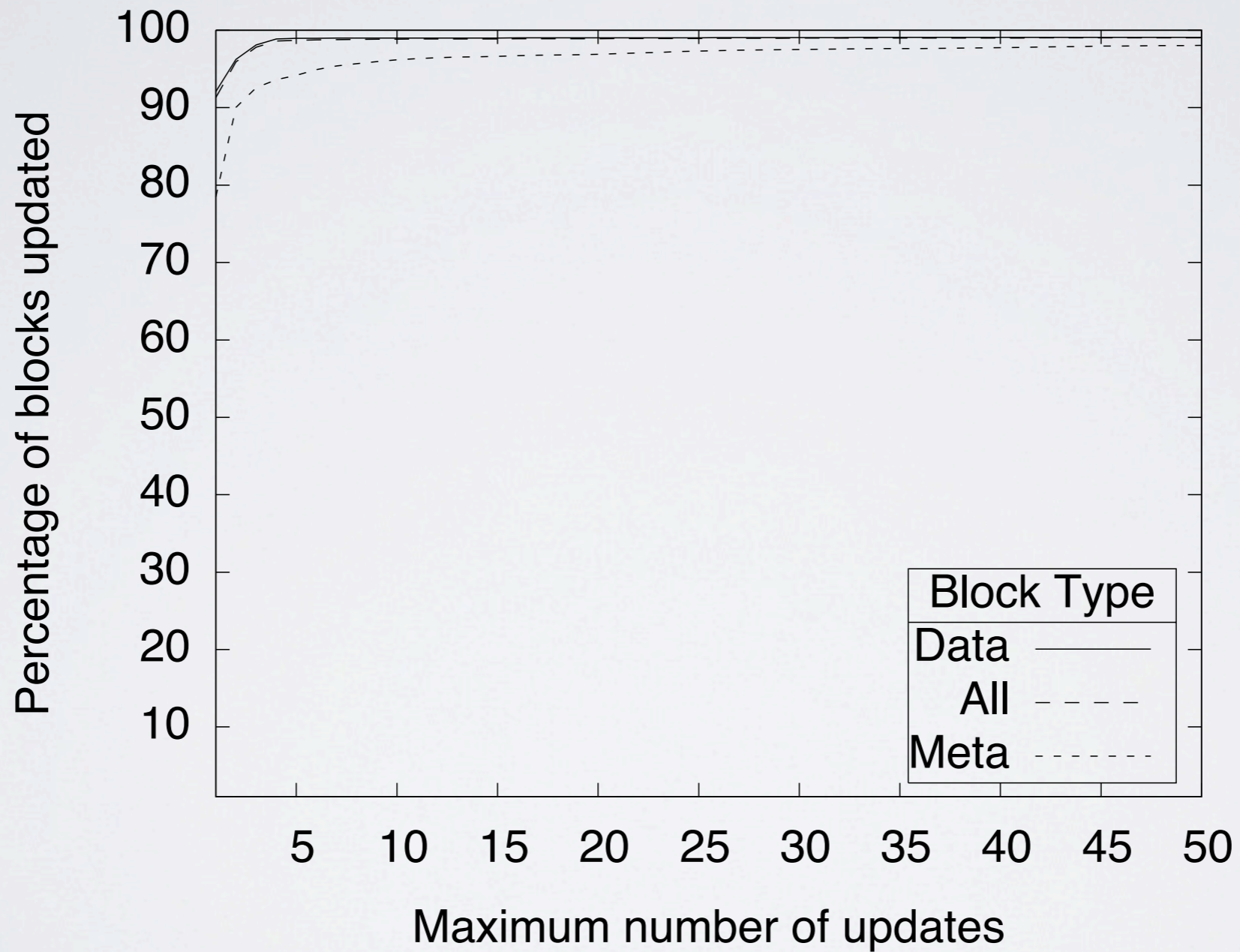
TYPE: HOUR



TYPE: DAY



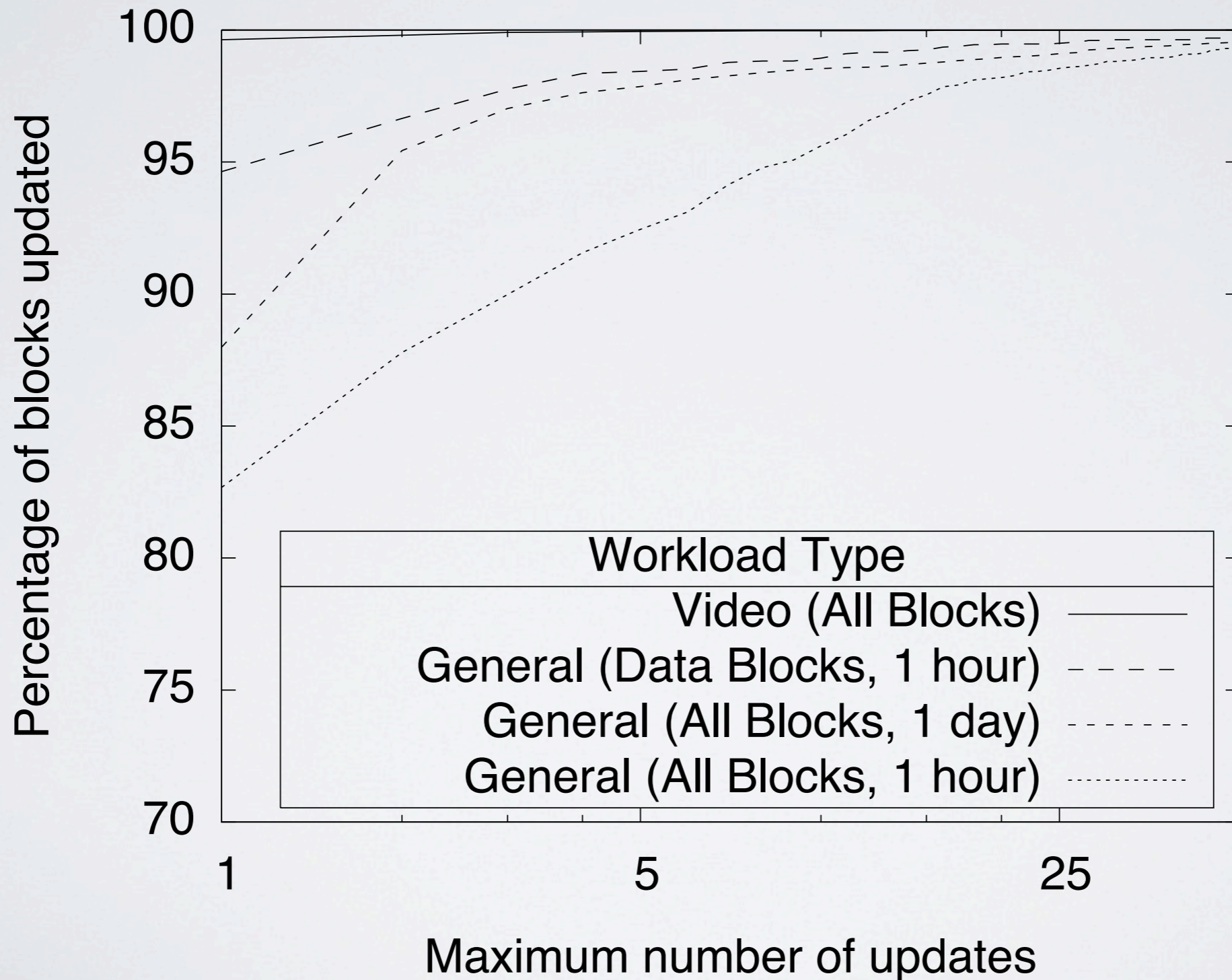
TYPE: FOUR DAY



FAVORABLE WORKLOAD

- Application: Video Editing
 - OSX 10.5.8, HFS+, active system in use for over a year
 - Interactive video editing
 - Simultaneous transcoding of a movie
 - Three hour observation period

FAVORABLE WORKLOADS



CONCLUSIONS

- Shingled Writing is a promising new recording technology
- Redirection and log-structuring of writes
 - Circular logs or Segment-based band division
 - In-band or intra-band cleaning
- Self-sufficient and hybrid usage options
 - Metadata separation results support object-based interfaces and hybrid usage
 - Self-sufficient usage feasible with appropriate workloads
- While application is dominant factor, general purpose usage appears promising

FUTURE WORK

- Evaluating further workloads
- Device simulation
- Prototype driver implementation
- Testing the automatic identification and classification of stable blocks

ACKNOWLEDGMENTS

- National Science Foundation Award# CNS-0917396
- National Science Foundation Award# IIP-093440
- The Industrial Sponsors of the Storage Systems Research Center (SSRC) and the Center for Research in Intelligent Storage (CRIS)
- The Stiles Family Foundation

QUESTIONS?