TECHNISCHE UNIVERSITÄT DRESDEN

Department of Computer Science, Institute for System Architecture, Operating Systems Group

An Analysis of Data Corruption in the Storage Stack

Lakshmi N. Bairavasundaram, Garth Goodson, Bianca Schroeder, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau

Presented by Carsten Weinhold

Paper Reading Group, 2008-06-24

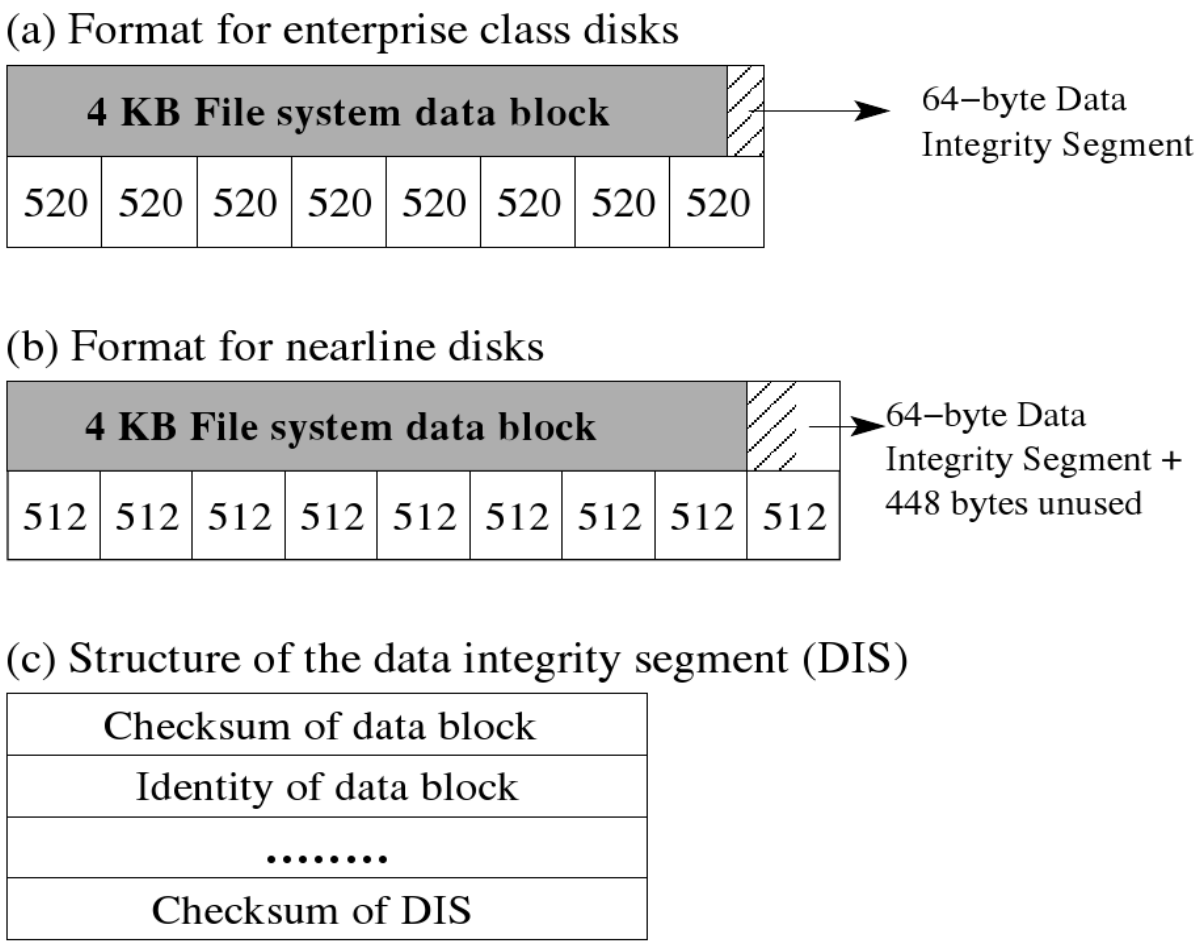
About the Study

- Large scale study:
  - Tens of thousands of production systems
  - 41 months
  - 1.53 million disks
  - 400,000+ checksum mismatches
- Both “nearline” and enterprise class disks
- Focus on silent data corruption
  (e.g., not about latent sector errors)

Background: NetApp Storage Systems

- All storage systems by Network Appliance™
- Dedicated network filers:
  - WAFL file system
  - RAID with parity
  - SCSI layer
  - Fibre Channel (FC) loops
  - Fibre Channel disks / SATA disks with adapter
- Data collected using "Autosupport"
- Sent to central database
- *Note: not all disks were in use for the full duration of 41 months*

Background: Data Integrity Segments



Corruption & Detection

Corruption Class	Possible Causes	Detection Mechanism	Detection Operation
Checksum mismatch	Bit-level corruption; torn write; misdirected write	RAID block checksum	Any disk read
Identity discrepancy	Lost or misdirected write	File system-level block identity	File system read
Parity inconsistency	Memory corruption; lost write; bad parity calculation	RAID parity mismatch	Data scrub

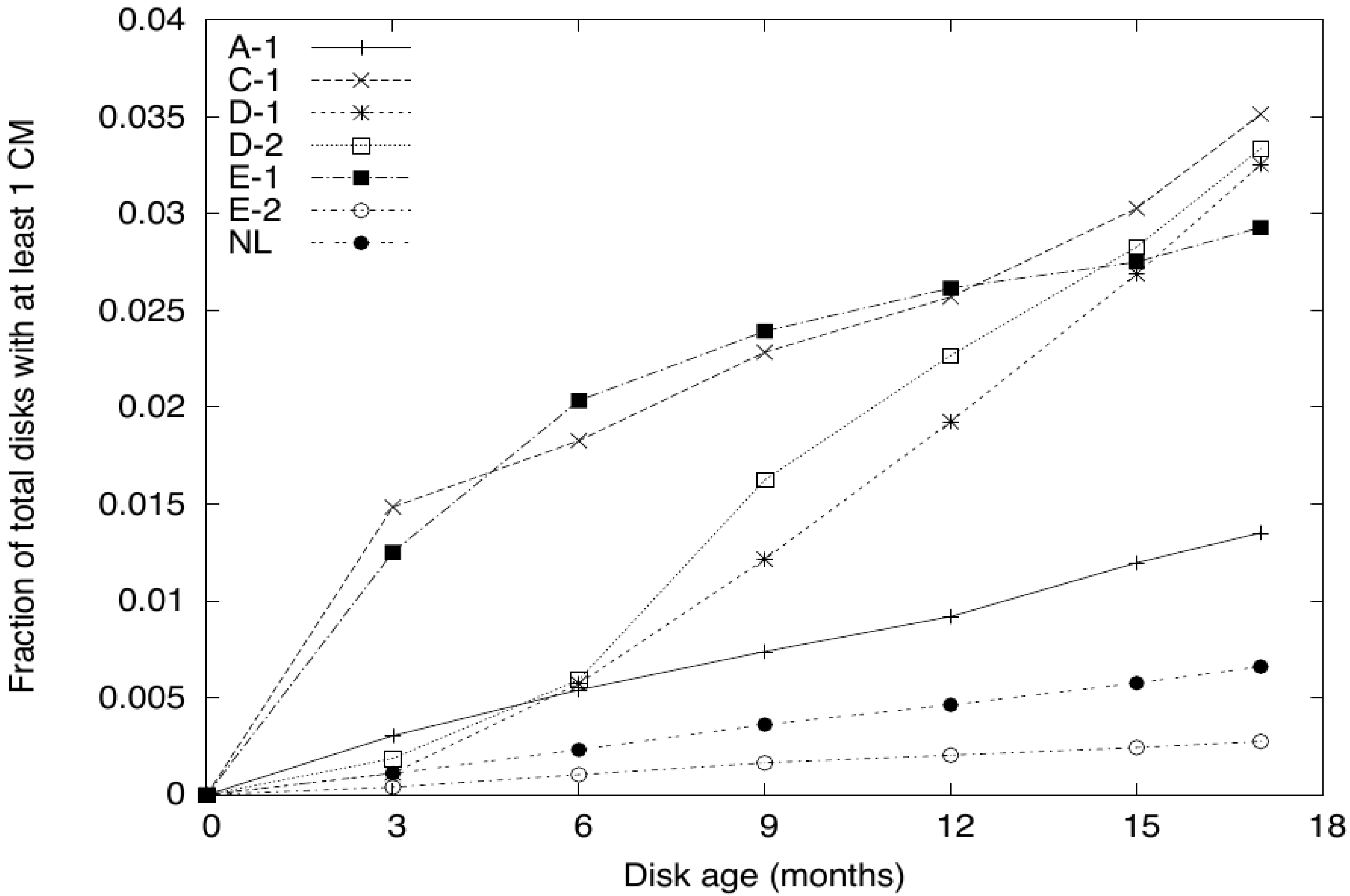
TECHNISCHE UNIVERSITÄT DRESDEN

Summary Statistics

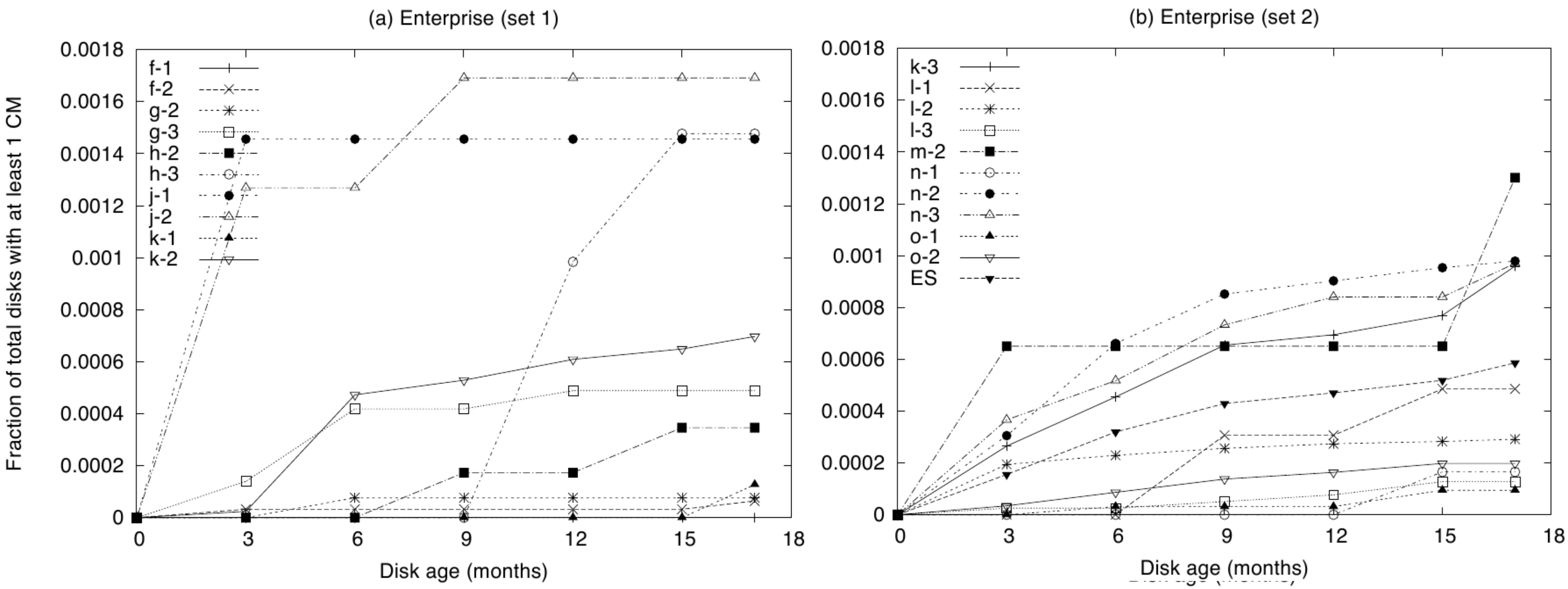
- Total of 1.53 million disks
- Total of 400,000+ checksum mismatches
- Percentage of corrupt disks varies:
  - **0.86%** of 358,000 *nearline* disks
  - **0.065%** of 1,170,000 *enterprise class* disks

Observation 1: the probability of developing checksum mismatches is an order of magnitude higher for nearline disks (+SATA/FC adapter) than for enterprise class disks

Factor Disk Age: Nearline Disks



Factor Disk Age: Enterprise Class Disks

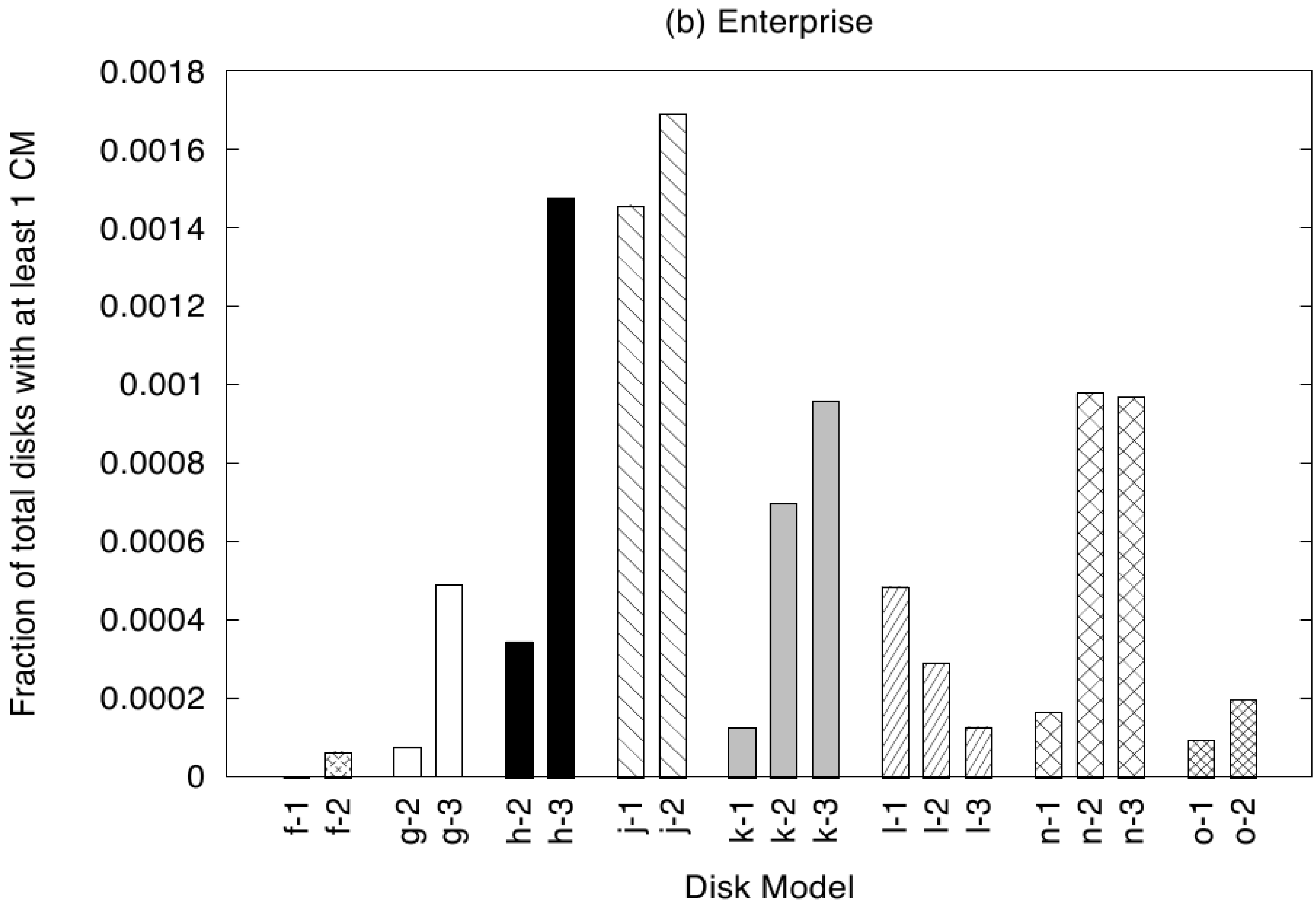


Observations

Observation 2: probability of developing checksum mismatches varies significantly across disk models in the same class of disks

Observation 3: age affects disk models differently with respect to the probability of developing checksum mismatches

Factor Disk Size ??



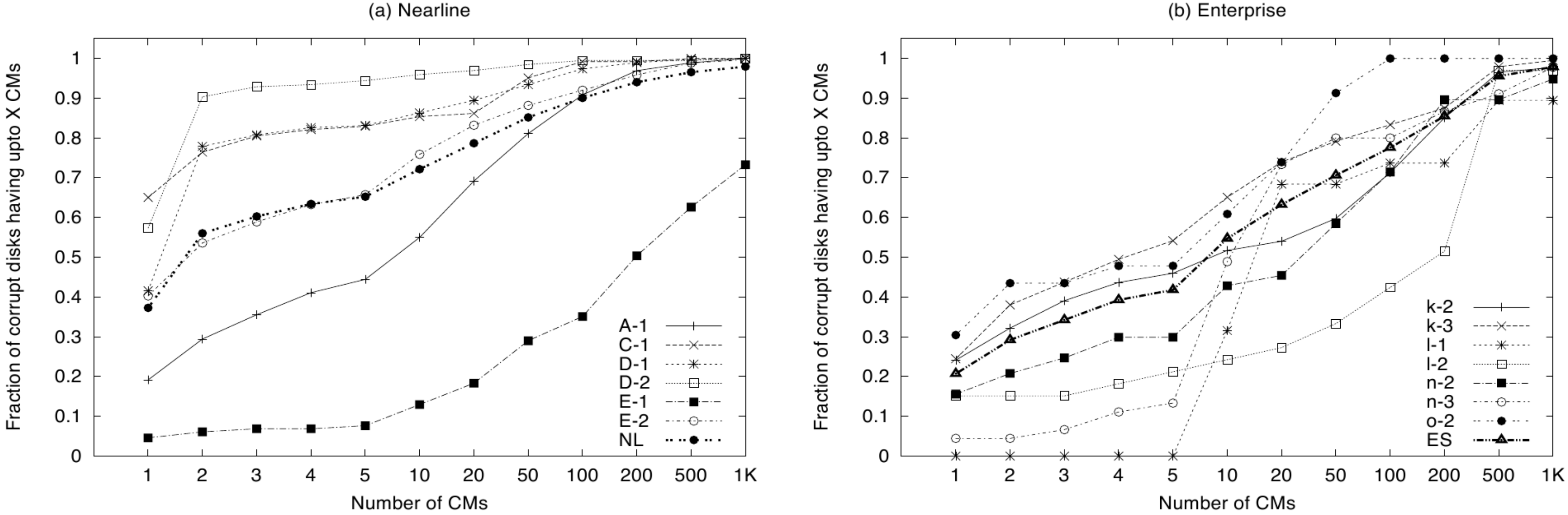
TECHNISCHE UNIVERSITÄT DRESDEN

(Non-)Factors ??

Observation 4: there is no clear indication that disk size affects the probability of developing checksum mismatches

Observation 5: there is no clear indication that workload affects the probability of developing checksum mismatches
... but: the collected data on access patterns was very coarse and likely to be insufficient

Characteristics: Models, Classes



Observation 6: the number of checksum mismatches varies greatly across disks

Observation 7: on average, corrupt enterprise class disks develop many more checksum mismatches than corrupt nearline disks

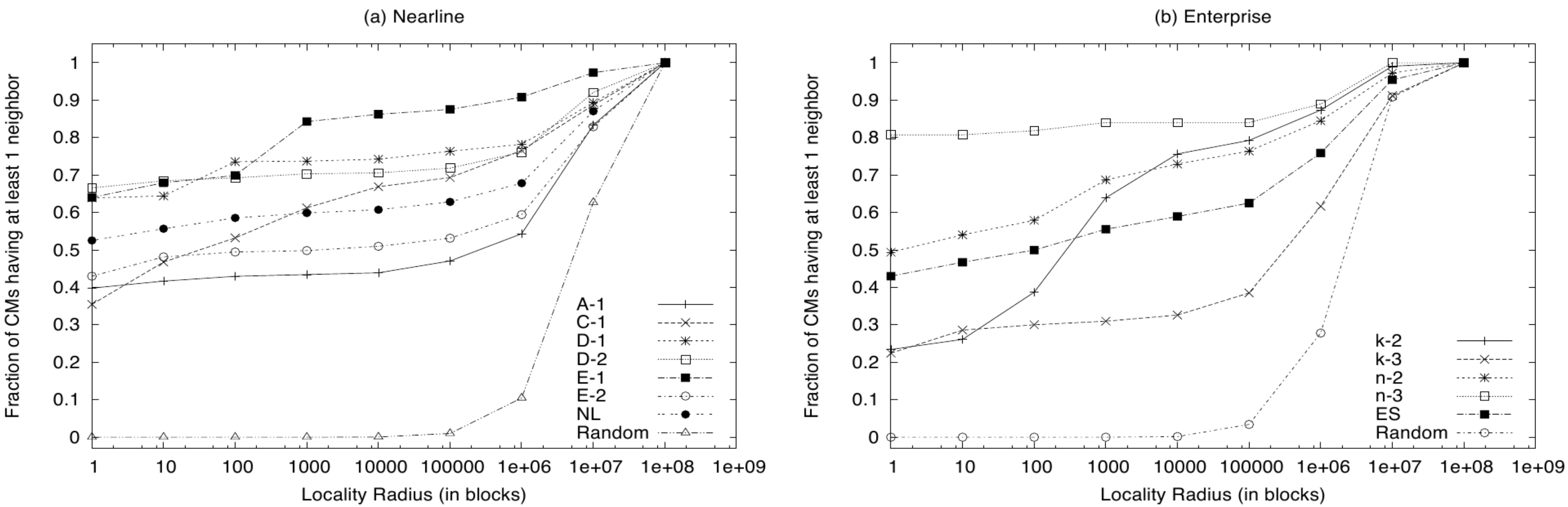
Characteristics: Disks and Disk Shelves

Observation 8: checksum mismatches within the same disk are not independent

Observation 9: the probability of developing a checksum mismatch is not independent of that of other disks in the same storage system

- Example:
  - One system had 92 disks develop errors
  - Caused by faulty storage controller

Characteristics: Locality



Observation 10: checksum mismatches have high spatial locality

Observation 11 & 12: there is temporal locality

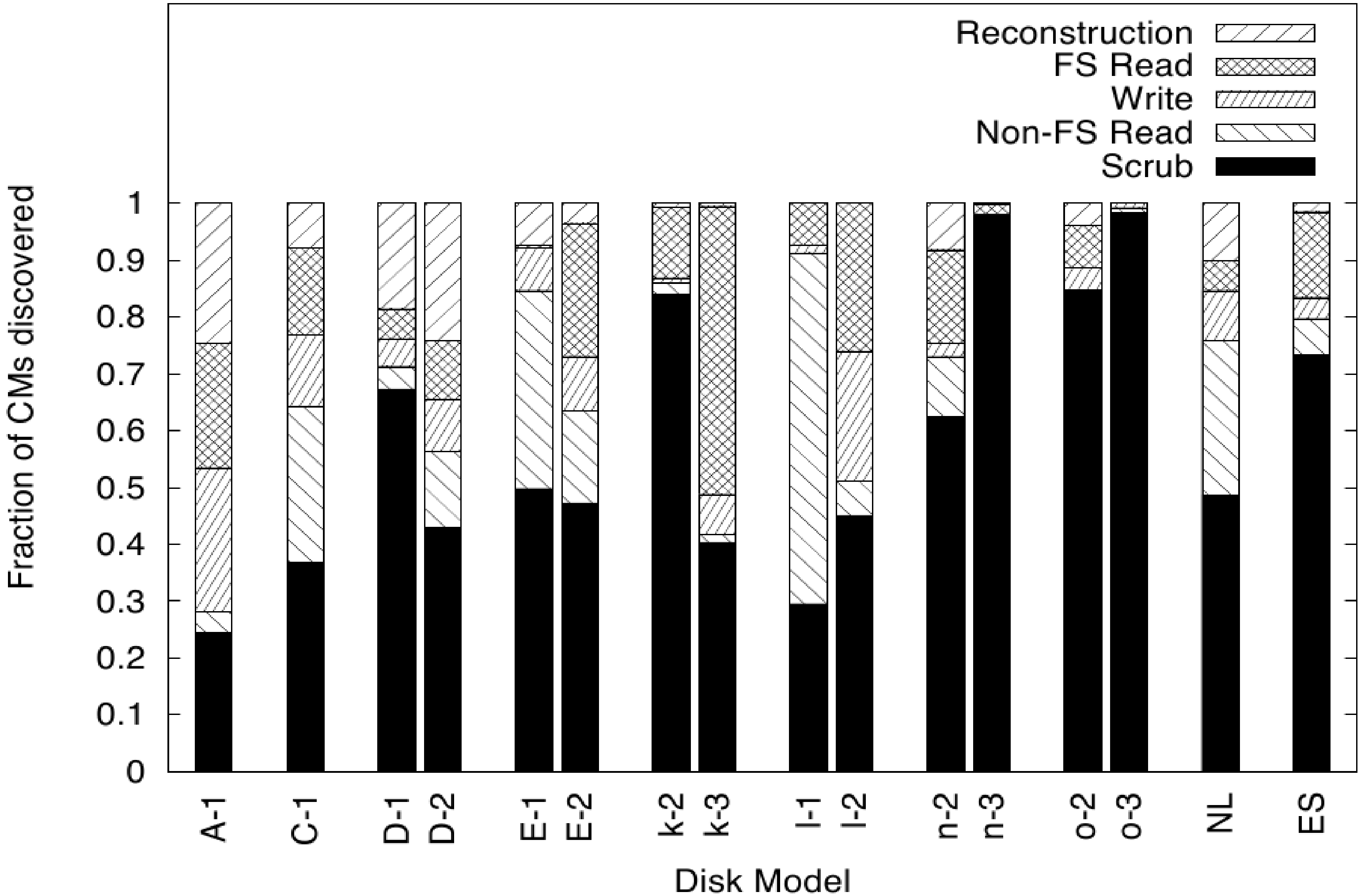
Characteristics: Error Type Correlation

Observations 12: checksum mismatches correlate with system resets

Observation 13: weak positive correlation between checksum mismatches and latent sector errors

- If latent sector errors detected, probability of developing checksum mismatches increases:
  - Nearline disks: 1.4 times
  - Enterprise class disks: 2.2 times

Request Type Analysis



TECHNISCHE UNIVERSITÄT DRESDEN

Comparison to Latent Sector Errors

Characteristic	Latent sector errors		Checksum mismatches	
	Nearline	**Enterprise**	**Nearline**	**Enterprise**
% disks affected per year (avg)	9.5%	1.4%	0.466%	0.042%
As disk age increases, P(1st error)	increases	remains constant	remains fairly constant	decreases, then stabilizes
As disk size increases, P(1st error)	increases	increases	unclear	unclear
No. of errors per disk with errors (80 percentile)	about 50	about 50	about 100	about 100
Are errors independent ?	No	No	No	No
Spatial locality	at 10 MB	at 10 MB	at 4KB	at 4KB
Temporal locality	very high	very high	very high	very high
Known Correlations	not-ready-conditions	recovered errors	system resets, not-ready-conditions	system resets, not-ready-conditions

TECHNISCHE UNIVERSITÄT DRESDEN

Lessons Learned

- Silent corruption does happen: **up to 4%** of drives developed errors in 17 months
- On average, **8%** of checksum mismatches detected during RAID reconstruction
  ➔ *Protection against double disk failure required*
- An enterprise class disk is likely to quickly develop more corruption after first occurrance
  ➔ *The faulty disk should be replaced soon*
- Some block numbers are more likely to be affected, possibly due to hardware/firmware bugs
  ➔ *Staggered striping for RAID should be used*

Lessons Learned (II)

- Corruptions have strong spatial locality
  - ➔ *Redundant data structures should stored distant from each other*
- Corruptions also have strong temporal locality
  - ➔ *Same write request? Use multiple write request for important / redundant data?*
  - ➔ *To be leveraged for smarter scrubbing?*
- Correlation of silent corruption and other errors could be used to improve failure prediction (e.g., latent sector errors)

TECHNISCHE UNIVERSITÄT DRESDEN

Discussion Points

- RAID does not (always) help and most file systems don't do checksumming! Is everything lost?
- Laptops have only one disk. ZFS supports redundancy on same disk. Any experiences?
- Can checksumming in the disk itself be improved? What would that mean with respect to firmware bugs?
- Why are enterprise class disks so much more reliable? Is there any hope that consumer disks catch up in the future?
- What about flash disks?

TECHNISCHE UNIVERSITÄT DRESDEN

References

- Lakshmi N. Bairavasundaram, Garth Goodson, Bianca Schroeder, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau, **“An Analysis of Data Corruption in the Storage Stack”**, FAST '08, San Jose