Reliability Mechanisms for Very Large Storage Systems

Qin Xin[*], Ethan Miller[*], Thomas Schwarz[+], Darrell Long[*], Scott Brandt[*], Witold Litwin[++]

[*]Storage Systems Research Center, University of California, Santa Cruz
[+]Santa Clara University
[++]University Paris 9 Dauphine

UC Santa Cruz

Outline

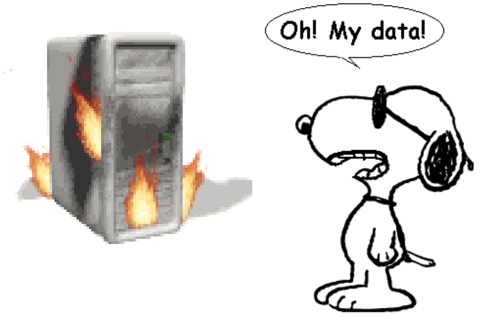
- Motivations and Goals
- Reliability Mechanisms
  - Signature scheme
  - Fast recovery schemes
- System Reliability Analysis
  - Size of a redundancy set
  - Mean-Time-To-Data-Loss of the system
- Conclusions & Future Work

Concerns for System Reliability

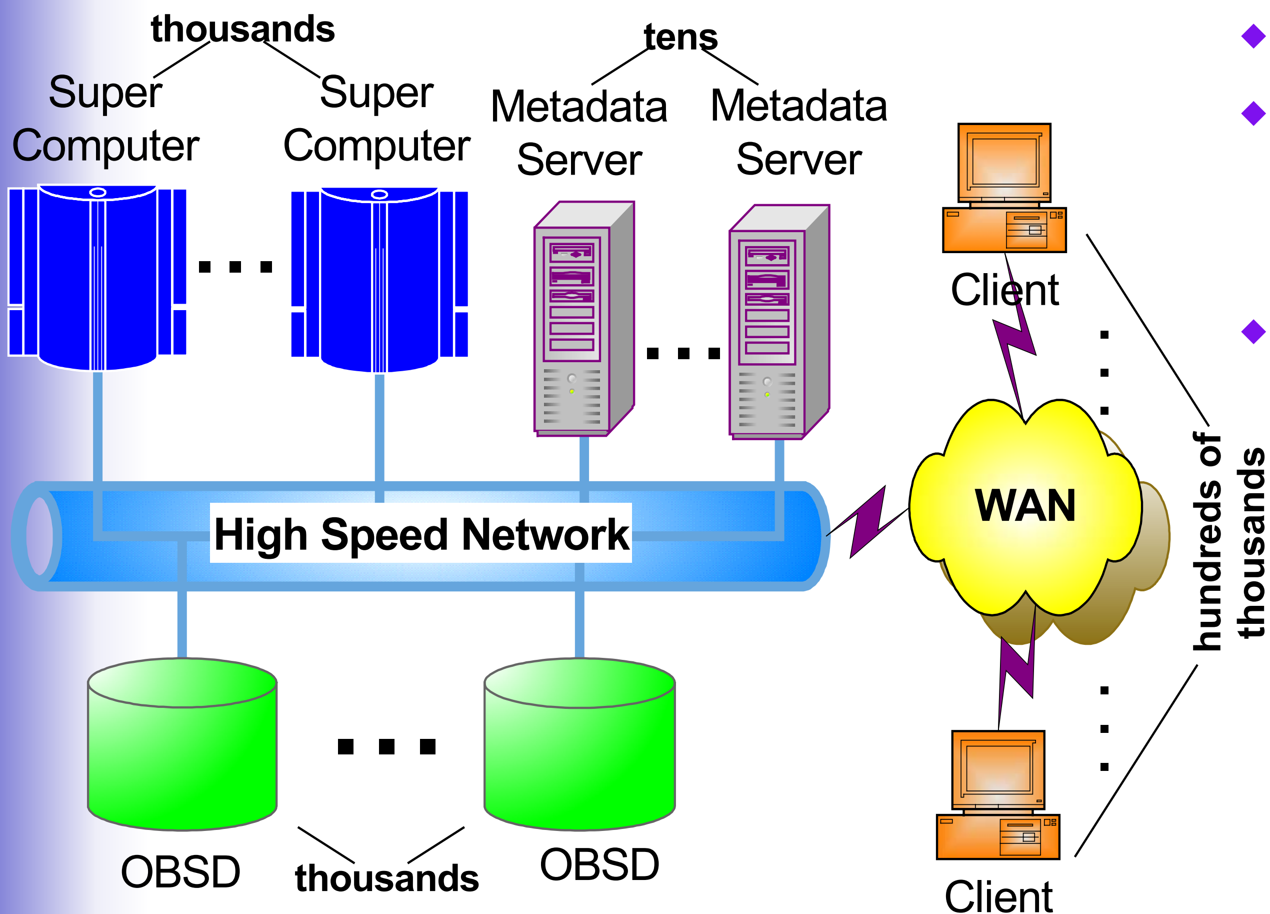
- Why are systems getting less reliable?
  - Complex computer components
  - Human errors
  - More components in large computer systems
- Impacts of system unreliability
  - Long down time
  - Increasing repair costs and Total Cost of Ownership
  - Frequent data loss

Reliability Challenges in Large Storage Systems

- More storage devices
  - High Performance vs. Low Reliability
- Larger disk capacity
  - Increase in disk capacity outpaces that of bandwidth
  - Disk rebuild time gets longer
- Goal: reduce the risk of data loss
- Main causes of data loss
  - Nonrecoverable Read Errors
  - Disk Failures



Object Based Storage System



- Petabyte scale
- Object-based storage management
- Deep concerns for data loss
  - Thousands of disks
  - Huge capacity

Cause I: Nonrecoverable Read Errors

- What is it?
  - Sector corruptions on disks and data cannot be read correctly.
  - Error rate: 1 in 10^{13} to 10^{15} bits
- Why do we care?
  - Increase in total data capacity and total system bandwidth
  - Once per year for a typical disk
  - Once per hour for the OBSD system
- Data corruption is not tolerable for storage systems

Solution: Signature Scheme

- A signature associated with each data block
  - Fixed-length: 8 or 16 bits
  - If ($Signature_{new}$!= $Signature_{prev}$), then flag an error.
  - Sources of errors
    - Data block error
    - Corrupted signature
- Data reconstruction
  - Replication
  - Parity
  - Erasure coding …

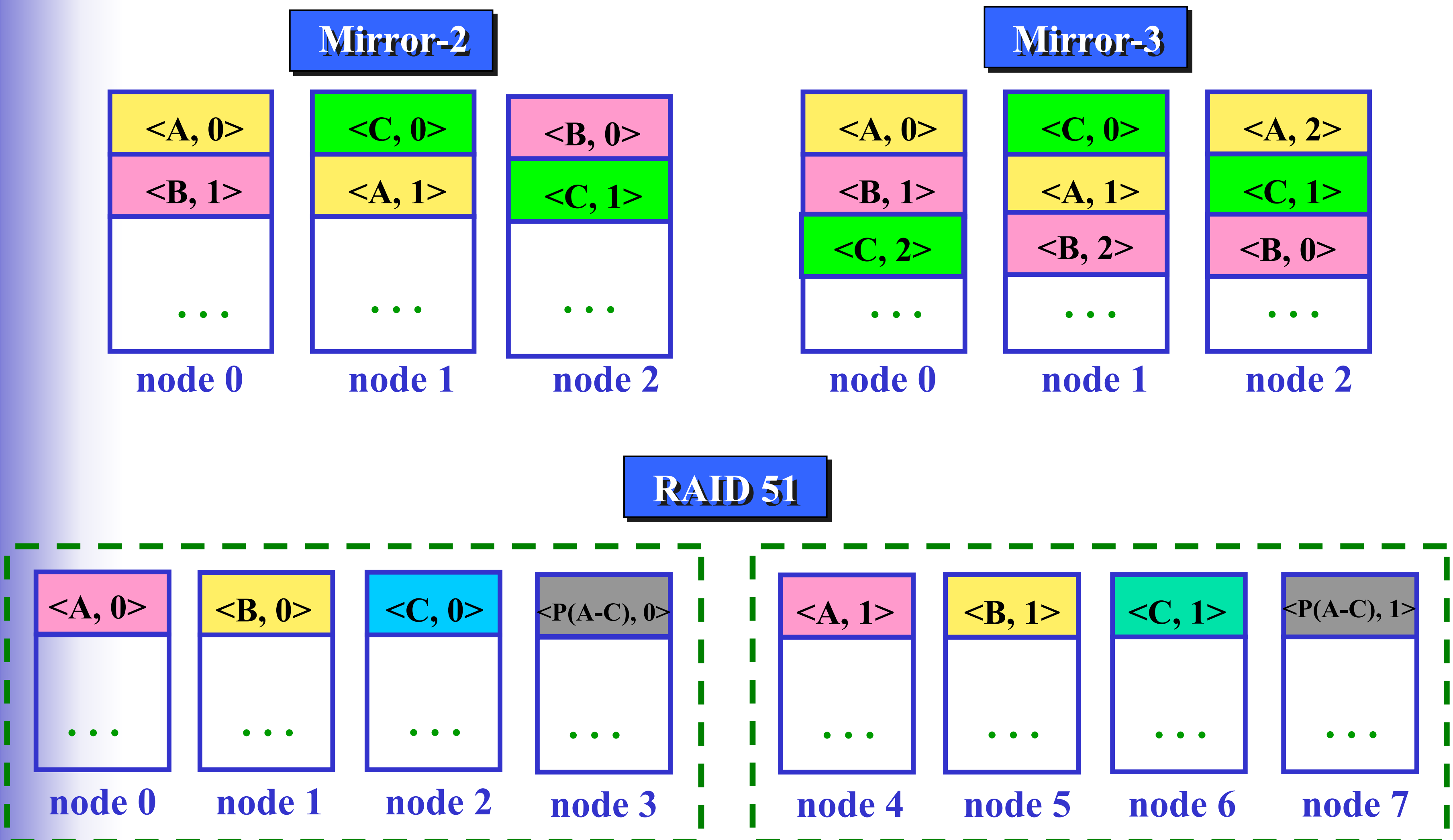
Cause II: Disk Failures

- Why we care? -- More frequent
  - 1 per 10^5 hours (11.4 years) for a single disk
  - For a system with thousands of disks, we might experience one disk failure per day.
- Why not just RAID?
  - Long disk rebuild time
  - The window of vulnerability gets wider.
  - To rebuild a 500 GB disk requires one day assuming rebuild rate is 5MB/sec.
  - MTTDL (Mean Time To Data Loss) = 3 years for a 2-Petabyte storage system.

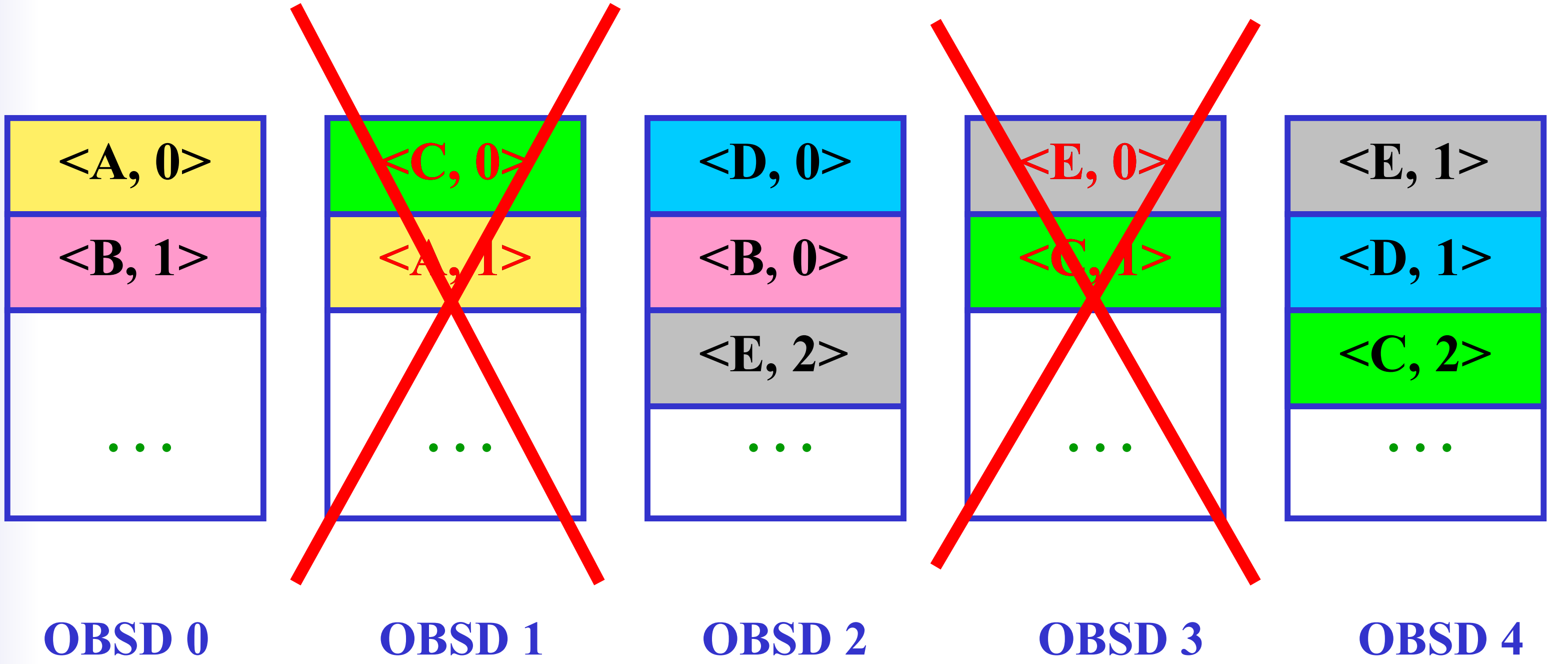
Solution: Reliability Mechanisms

- Redundancy set
  - A block group composed of data blocks and their associated replicas or parity blocks
- Configurations
  - 2-way mirroring (Mirror-2)
  - 3-way mirroring (Mirror-3)
  - RAID5+mirroring (RAID51)
- Fast Recovery Schemes
  - Fast Mirroring Copy
  - Lazy Parity Backup

Redundancy Set Configurations

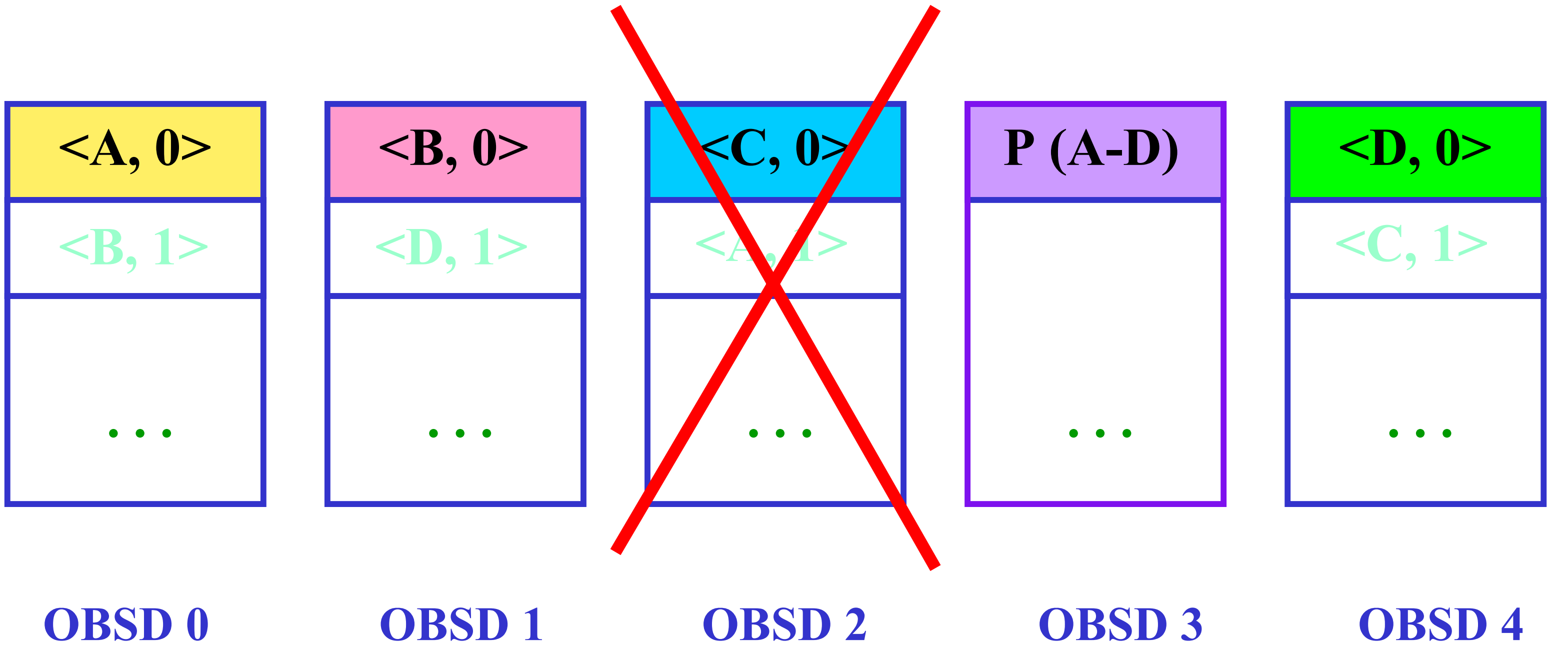


Fast Mirroring Copy (FMC)



No data loss

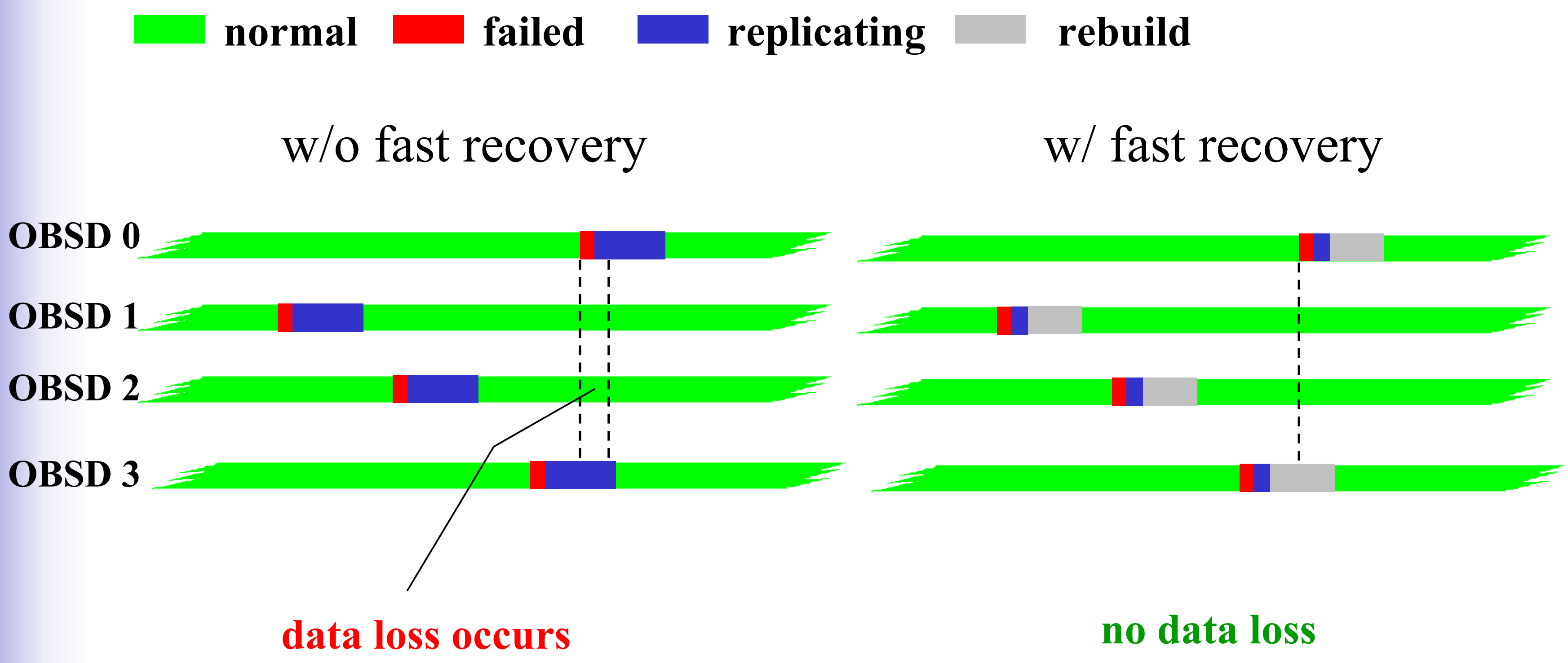
Lazy Parity Backup (LPB)



No data loss

Why does fast recovery work?

- Narrowing the windows of vulnerability



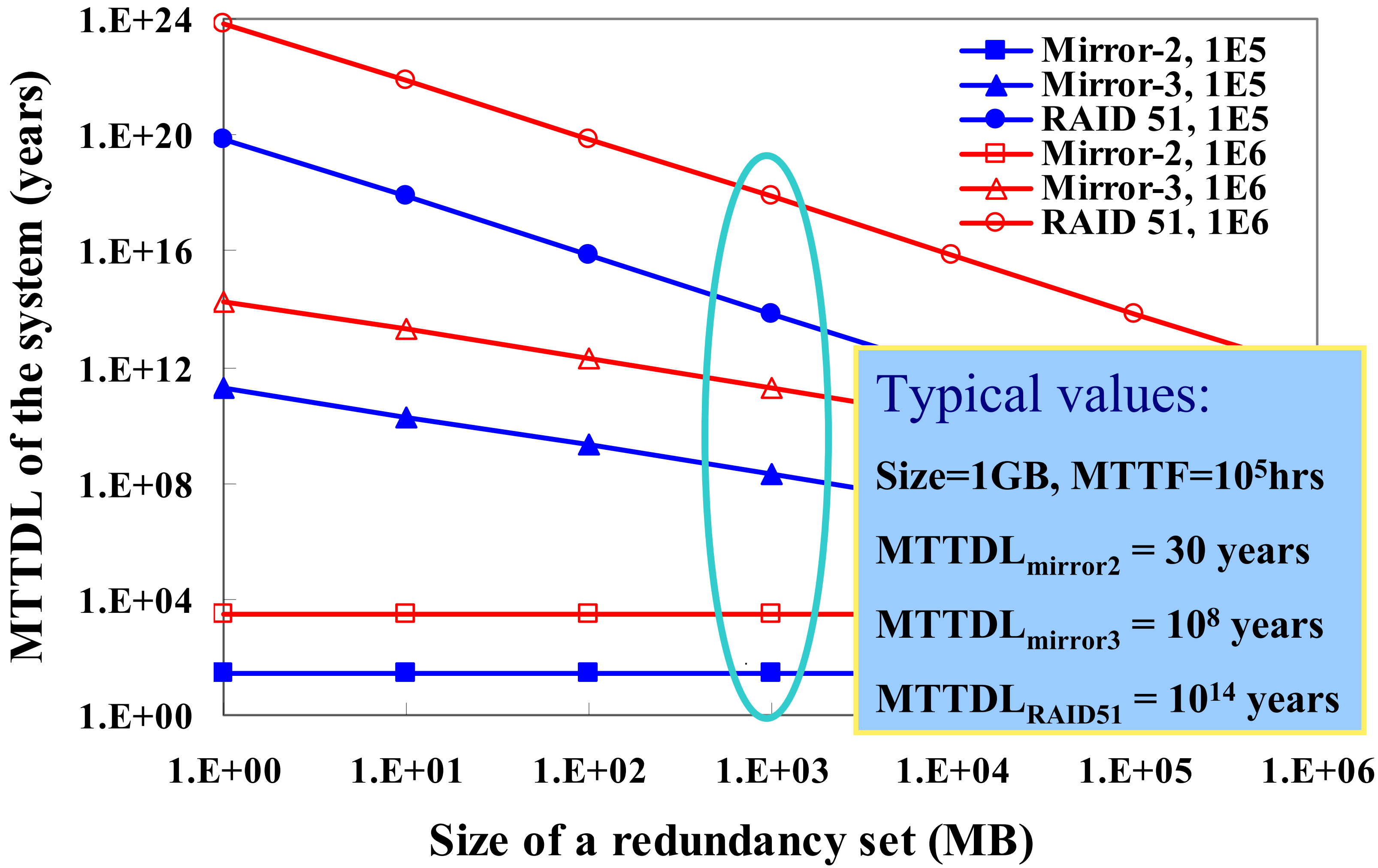
Reliability Analysis

- Assumptions
  - Total data capacity: Z = 2 Petabytes
  - $MTTF_{disk} = 10^5$ -- 10^6 hours
  - Failures of the disks are independent.
  - Recovery rate: γ = 100GB/hour
  - S: size of a redundancy set; D: # of disks in one RAID5
- Markov Models
- Mean-Time-To-Data-Loss (MTTDL)

$$MTTDL_{mirror2} = \frac{MTTF_{disk}^{2} \cdot \gamma}{2 \cdot Z} \qquad MTTDL_{mirror3} = \frac{MTTF_{disk}^{3} \cdot \gamma^{2}}{3 \cdot S \cdot Z}$$

$$MTTDL_{raid51} = \frac{MTTF_{disk}^{4} \cdot \gamma^{3}}{4 \cdot D \cdot (D-1) \cdot S^{2} \cdot Z}$$

Comparison of Reliability (log-log)



Typical values:

Size=1GB, MTTF=10^5hrs

$MTTDL_{mirror2}$ = 30 years

$MTTDL_{mirror3}$ = 10^8 years

$MTTDL_{RAID51}$ = 10^{14} years

Conclusions

- Two major sources of data loss in large storage systems
  - Nonrecoverable read errors
  - Disk failures
- Reliability mechanisms
  - Signature scheme
  - Fast recovery mechanisms
    - Fast Mirroring Copy
    - Lazy Parity Backup
- Reliability analysis
  - Mirror2 w/ fast recovery can provide 30-year MTTDL.
  - Mirror3 or RAID51 w/ fast recovery can provide very high reliability.

Future Work

- More details on failure distributions
- Impacts of data placement policies on system reliability
- Data consistency schemes
- Advanced erasure coding

Acknowledgements

- Jean-Jacques Bedet -- our shepherd
- Members of the Storage System Research Center (SSRC) at the University of California, Santa Cruz
- Project sponsors
  - Lawrence Livermore National Laboratory
  - Los Alamos National Laboratory
  - Sandia National Laboratory
  - IBM Research
  - European Commission
  - Microsoft Research

Questions or comments?

Galois Power Signatures

- Why not SHA1?
  - Need for consistency checking in large storage systems
- Galois Field (GF) : a finite set
- Galois power signatures for a block
  - A block $\boldsymbol{P}$ has $\boldsymbol{l}$ symbols, each symbol is $\boldsymbol{f}$ bits long.

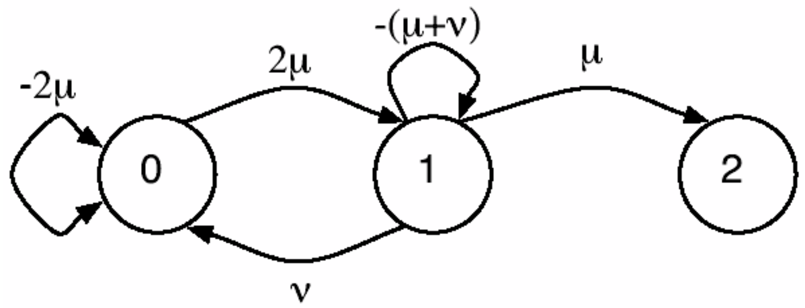
$$P = p_1 p_2 p_3 \ldots p_l$$

  - β : element of GF(2^f)
  - β signature of a block $sig_{\beta}(P) = \sum_{\mu=1}^{l} p_{\mu} \beta^{\mu-1}$
  - n-fold α-signature

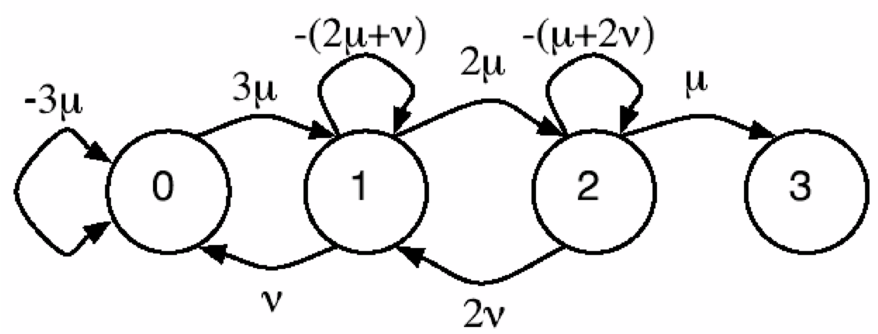
$$sig_{\alpha,n}(P) = (sig_{\alpha}(P), sig_{\alpha^2}(P), \ldots, sig_{\alpha^n}(P))$$

Markov Models

- 2 way mirroring

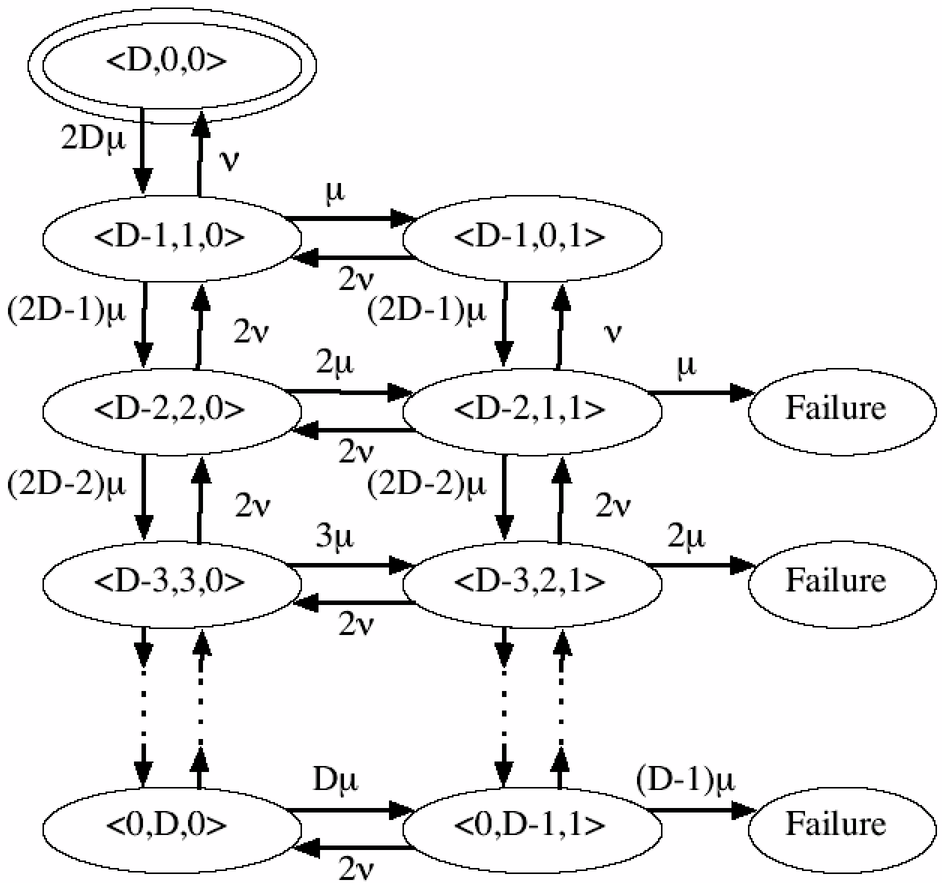


- 3 way mirroring



Markov Models (cont.)

- RAID 51



Related Work

- RAID: classic method for reliability and recovery
- OceanStore: designed to have a long MTTDL
- FARSITE: replica placement policies
- ROC: decrease TCO by reducing recovery time
- Muntz and Liu: disk array declustering
- Menon and Mattson: distributed sparing
- Long: consistency management for mirrored disks
- Castro and Liskov: secure replication to tolerate Byzantine faults
- Honicky and Miller: online data reorganization
- Litwin and Schwarz: a family of linear hashing models
- Schwarz: a Markov model to estimate system availability