



Education

RETAINING INFORMATION FOR 100 YEARS

Mary Baker, HP Labs
Roger Cummings, Symantec

- The material contained in this tutorial is copyrighted by the SNIA.
 - Member companies and individuals may use this material in presentations and literature under the following conditions:
 - ◆ Any slide or slides used must be reproduced without modification
 - ◆ The SNIA must be acknowledged as source of any material used in the body of any document containing material from these presentations.
 - This presentation is a project of the SNIA Education Committee.
 - Neither the Author nor the Presenter is an attorney and nothing in this presentation is intended to be nor should be construed as legal advice or opinion. If you need legal advice or legal opinion please contact an attorney.
 - The information presented herein represents the Author's personal opinion and current understanding of the issues involved. The Author, the Presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.
- NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.**

➤ RETAINING INFORMATION FOR 100 YEARS

- ◆ Many organizations now have a requirement to preserve large volumes of digital content indefinitely into the future, and to maintain access for reasons such as medical treatment decisions, retention of intellectual property, and appreciation of cultural and scientific history. Frequent news stories cover organizations' failures to be able to do this, such as the near loss of original video/data of the first Moon landing, eventually recovered from a set of 14-inch tape reels found in a dusty Australian basement.
- ◆ This session will focus on the most important questions in long-term digital preservation and will demonstrate why it is still so difficult. We will propose how the storage industry can help its customers preserve and use their digital content over the lifetimes that they expect from past experience with physical and analog assets, lifetimes that can greatly exceed those of any single digital storage device or storage technology.

➤ Introduction

- ◆ Why we need digital preservation
- ◆ The goals of digital preservation

➤ Threats

- ◆ Threats to long-term digital content
- ◆ How long-term and short-term threats differ
- ◆ Why it is hard to address these threats

➤ Current status

- ◆ Best practices
- ◆ Some current projects
- ◆ Open problems and opportunities
- ◆ Please give us feedback

Why we need digital preservation

- Regulatory compliance and legal issues
 - ◆ Sarbanes-Oxley, HIPAA, FRCP, intellectual property litigation
 - Emerging web services
 - ◆ Email, photo sharing, web site archives, IM, blogs
 - Many other fixed-content repositories
 - ◆ Scientific data, intelligence, libraries, movies, music
-
- Responses to 100 Year Archive Requirements Survey
 - ◆ 68% of organizations had requirements for over 100 years
 - ◆ 83% of organizations had requirements for over 50 years



Goals of digital preservation

- Digital assets stored now should remain
 - ◆ accessible
 - ◆ usable
 - ◆ undamaged

- for as long as desired – beyond the lifetime of
 - ◆ any particular storage system
 - ◆ any particular storage technology

- and at an affordable cost

➤ Introduction

- ◆ Why we need digital preservation
- ◆ The goals of digital preservation

➤ **Threats**

- ◆ Threats to long-term digital content
- ◆ How they differ from short-term threats
- ◆ Why it is hard to address these threats

➤ Current status

- ◆ Best practices
- ◆ Some current projects
- ◆ Open problems and opportunities
- ◆ Please give us feedback

Threats to long-term digital assets

- Large-scale disaster
- Human error
- Media faults

- Component faults
- Economic faults
- Attack
- Organizational faults

Long-term content suffers from more threats than short-term content

- Media/hardware obsolescence
- Software/format obsolescence
- Lost context/metadata

Threats to long-term digital assets

- Large-scale disaster
- Human error
- Media faults ←
- Component faults
- Economic faults
- Attack
- Organizational faults



- Media/hardware obsolescence
- Software/format obsolescence
- Lost context/metadata

Threats to long-term digital assets

- Large-scale disaster
- Human error
- Media faults

- Component faults
- Economic faults
- Attack ←
- Organizational faults



- Media/hardware obsolescence
- Software/format obsolescence
- Lost context/metadata

Threats to long-term digital assets

- Large-scale disaster
- Human error
- Media faults

- Component faults
- Economic faults
- Attack
- Organizational faults



- Media/hardware obsolescence ←
- Software/format obsolescence
- Lost context/metadata

Even bit preservation is hard

- Large scale & long time periods are a problem
- IPB, 50 years, 50% probability of no damage
 - ◆ Sounds reasonable, doesn't it?
- That's a bit half-life of 10^{17} years
 - ◆ A million times the age of the universe
 - ◆ Even improbable events will have an effect
- Now try to keep
 - ◆ The bits usable
 - ◆ The information reusable
- Preserve just the bits?
 - ◆ Can't interpret the content
- Focus only on the logical aspects?
 - ◆ The bits have been trashed

➤ Introduction

- ◆ Why we need digital preservation
- ◆ The goals of digital preservation

➤ **Threats**

- ◆ Threats to long-term digital content
- ◆ How they differ from short-term threats
- ◆ Why it is hard to address these threats

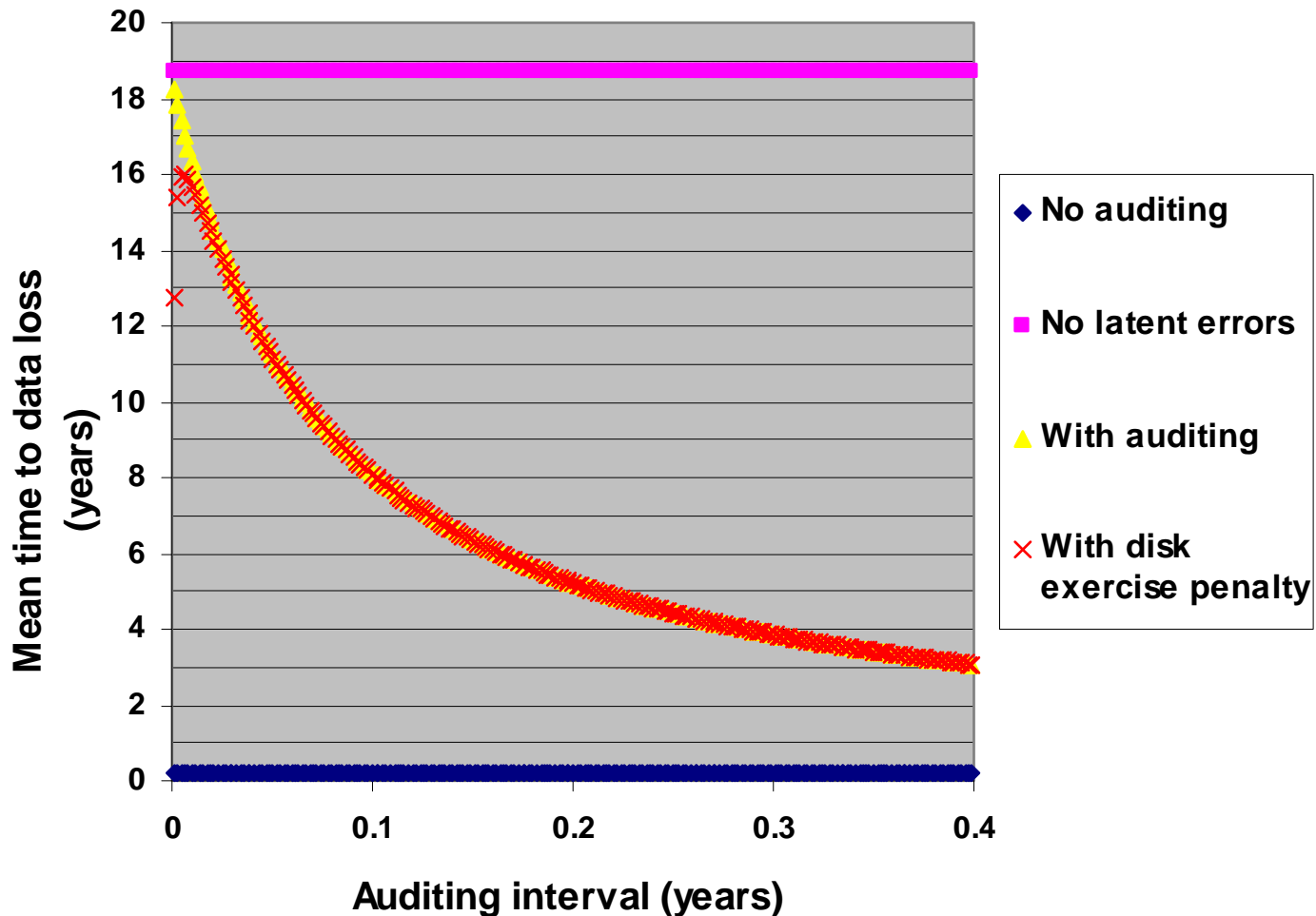
➤ Current status

- ◆ Best practices
- ◆ Some current projects
- ◆ Open problems and opportunities
- ◆ Please give us feedback

Key ideas for best practices

- Replicate content
 - ◆ If one copy is damaged, can repair from another
 - ◆ It's not enough to make a single “super reliable” copy
- Avoid failure correlations
 - ◆ Not just geographic, but administrative, platform, etc...
 - ◆ Heterogeneity helps avoid correlations
 - ◆ Must balance this with potential cost and administrative hassle
- Find & fix (if possible) latent faults before damage grows
 - ◆ Latent faults occur all the way up the “stack”
 - ◆ These faults can be physical or logical
 - ◆ Must look for silent damage and problems proactively: “audit”
 - ◆ This means content must be accessible!
- Use widely understood standards
 - ◆ Help customers avoid metadata and format traps
 - ◆ Help customers migrate content to (your!) new technologies

Reliability vs. Auditing



Best practices will vary over time

- We can't predict what will change – we only know it will
 - ◆ Ability to evolve is most important aspect of digital preservation
- This means processes are key
 - ◆ Must ensure our preservation processes are evolvable
 - ◆ Current processes are the first step in an iterative solution
 - > They get us to the next step
 - > At that point we will likely need new processes to take over
 - ◆ Physical preservationists ensure all transformations are undoable
 - ◆ Workable processes vary across organizations/domains
 - ◆ Widely understood standards make process evolution easier
- A good archive is almost always in motion
 - ◆ Migrating, auditing, re-keying, etc.
 - ◆ **Digital preservation is not a static activity!**
 - ◆ You can't just “do it and be done with it”

Best practices will vary by context

- What do we preserve?
 - ◆ Bits? Applications? Logical connections? Context? Etc.?
- Whatever the customer in that domain wants
 - ◆ Different domains/industries /organizations need different things
 - > Static versus dynamic content
 - > Self-contained content versus many external dependencies
 - > Different levels of fidelity and context
 - ◆ Example: digital copy of old book
 - > Just copy the words?
 - > Reproduce wear and tear on the paper?
 - > What about the political context in which it was read?
- Can't predict the eventual use of the material
- Affordability may force some decisions

Which methods are best?

- Do we use
 - ◆ Virtual machines?
 - ◆ Emulation?
 - ◆ Canonical formats?
 - ◆ Self-describing formats?
 - ◆ Standardized data models?
 - ◆ Loss-tolerant formats?
 - ◆ Format migration?
 - ◆ Preservation of ancient equipment?
- Yes: all could play a role for different domains
 - ◆ Some can be very expensive

Two current projects

- Both projects oriented toward solving migration
 - ◆ Logical & physical migration are important problems
 - ◆ A means to interpret content into the future
 - ◆ We need tested, affordable, scalable solutions
- Self-contained Information Retention Format
 - ◆ (SIRF)
 - ◆ From the Long-term Retention TWG (LTR)
 - ◆ Logical & physical migration part of TWG 1st-phase efforts
 - ◆ A CASPAR collaboration
- Object-based Storage Devices
 - ◆ Using OSDs to help with physical migration
 - ◆ Related work from the OSD TWG

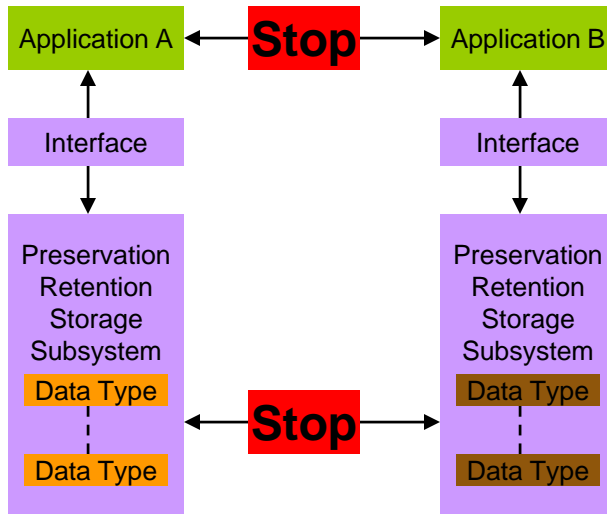


SIRF: logical container format

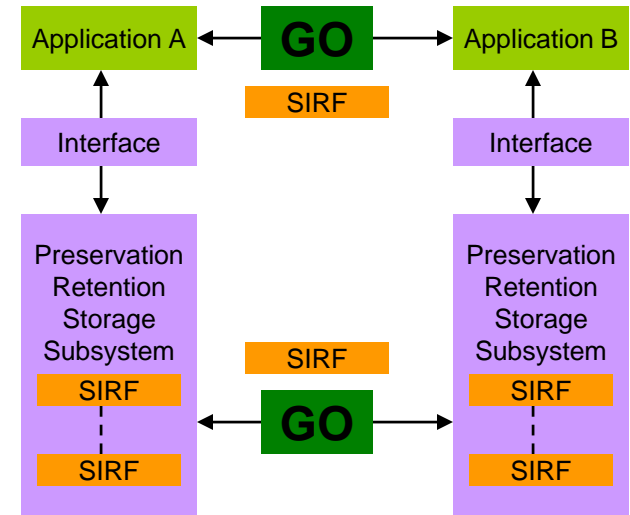
- Appropriate for long-term storage of digital information
- Logical data format of a mountable unit
 - ◆ File system, block device, stream device, object store, tape, etc.
- Includes a cluster of “interpretable” preservation objects
 - ◆ Self-describing – can be interpreted by different systems
 - ◆ Self-contained – all interpretation data contained in object cluster
- Facilitate transparent migration for long-term preservation
 - ◆ Logical
 - ◆ Physical
- Several implementations
 - ◆ Open Archival Information System (OAIS) ISO standard
 - ◆ Extensible Access Method (XAM)
 - ◆ Others

Problem SIRF addresses

Without SIRF

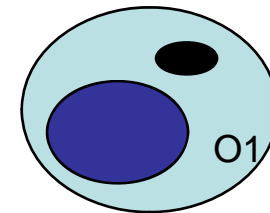


With SIRF



Cannot move cluster of preservation objects between systems without help	Can move cluster of preservation objects between systems by itself
Only original application that wrote the preservation objects can read and interpret them	Any SIRF compliant application can read and interpret preservation objects
Need export and import processes	No need for export and import processes
Preservation objects cannot be sustained for long-term	Preservation objects survive longer

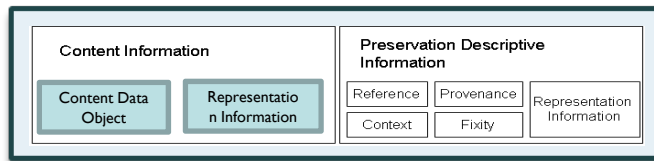
- Object-based Storage Devices help with archival storage
- Help to reduce scalability and cost problems
 - ◆ Logical migration – via SIRF and XAM
 - ◆ Physical migration – via SIRF and OSD
- Standards-based interfaces
 - ◆ Application gets information independent of storage platforms
 - ◆ *Encapsulated data and **metadata***
 - ◆ Applications write in standard archive-formats
 - ◆ Readers for long-term data access
- Supporting services
 - ◆ ILM-based practices
 - ◆ Discovery, security



Objects are the key enabler
Metadata goes with the object

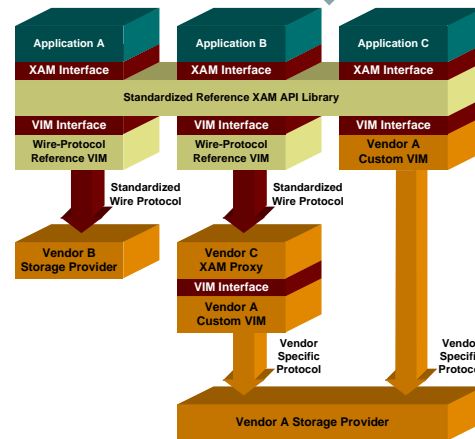
The ecosystem for OSD

- Standards-based container
- Data & metadata
- Logical migration



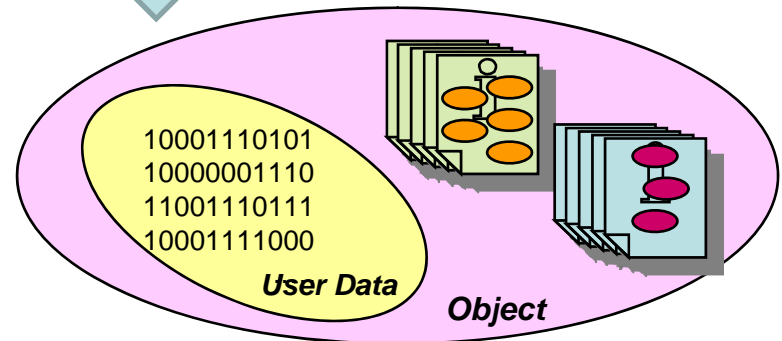
- Standards-based API
- Data & metadata
- Logical migration

SIRF
(SNIA LTR TWG)
or
OAIS AIP



XAM
(SNIA FCAS TWG)
&
(SNIA XAM SDK TWG)

- Standards-based interface
- Data & metadata
- Physical migration



Objects
(SNIA OSD TWG)

Related tutorials



Check out **SNIA Tutorial:**

- ◆ Green Storage | Economics, Environment, Energy and Engineering



Check out **SNIA Tutorial:**

- ◆ Trends in Data Protection and Restoration Technologies



Check out **SNIA Tutorial:**

- ◆ Deduplication – Methods for Achieving Data Efficiency

For more information

- Survey of data retention problems and requirements
 - ◆ 100 Year Archive Task Force, SNIA Data Management Forum, “100 Year Archive Requirements Survey,” January 2007.
- Measurements, modeling of storage failure (still an immature area)
 - ◆ M. Baker, et al., “A Fresh Look at the Reliability of Long-term Digital Storage.” EuroSys’06.
 - ◆ E. Pinheiro et al., “Failure Trends in a Large Disk Drive Population.” Usenix FAST’07.
 - ◆ B. Schroeder et al., “Disk failures in the real world: What does an MTTF of 1,000,000 hours mean too you?” Usenix FAST’07.
 - ◆ L. Bairavasundaram, et al., “An analysis of data corruption in the storage stack.” Usenix FAST’08.
 - ◆ W. Jiang, et al., “Are Disks the Dominant Contributor for Storage Failures? A Comprehensive Study of Storage Subsystem Failure Characteristics.” Usenix FAST’08.
 - ◆ A. Krioukov et al., “Parity Lost and Parity Regained.” Usenix FAST’08.
- SNIA LTR TWG knowledge base for many other topics, including:
 - ◆ Standards such as OAIS
 - ◆ Pointers to work on related problems (authentication, power management, etc.)
 - ◆ Project and deployment descriptions (British Library, CASPAR, LOCKSS)

Some open problems & opportunities

- Logical and physical migration
 - ◆ Already activity in this area, but with lots more room for help
- Failure data
 - ◆ What are all the ways we really lose content?
- Reliability modeling
 - ◆ Holistic models to reason about probabilities of content loss
- Accelerated aging
 - ◆ How do we know if we've been successful?
- Dealing with secrets for long periods of time
 - ◆ Secrets can be the hardest things to preserve
- Long-term cost modeling
 - ◆ What is the cost to preserve this document for 100 years?
- 3rd-party validation of storage SLAs
 - ◆ Ways to tell that a preservation service is meeting its promises
- Choosing what to preserve
 - ◆ Can/should we save everything? If not, how do we choose?

Please get involved!

- Long-term Retention Technical Working Group
 - ◆ Both bit preservation and logical preservation
 - ◆ Contribute your experience and knowledge
 - ◆ Help set direction of technical efforts
 - ◆ What do we need to work on that we're not?
- Data Management Forum (DMF)
 - ◆ Particularly the Long Term Archive and Compliance Storage Initiative (LTACSI)
 - ◆ Requirements for long-term retention & compliance
 - ◆ Raise awareness in the industry
 - ◆ Standardize our terminology and understanding

- Please send any questions or comments on this presentation to SNIA: trackdatamgmt@snia.org

**Many thanks to the following individuals
for their contributions to this tutorial.**

- SNIA Education Committee

**Mary Baker
Simona Cohen
Roger Cummings
Sam Fineberg
Annie Foong
Sami Iren
Petros Maniatis
Rob Peglar**

**Michael Peterson
Jeff Porter
Bob Rogers
David Rosenthal
Ramin Samadani
Mehul Shah
Irwin Sobel
Gary Zasman**

- Digital preservation is important now
 - ◆ And is becoming more so
- Best practices center around
 - ◆ Replication
 - ◆ Avoiding correlated failures
 - ◆ Finding and fixing latent damage
 - ◆ Choosing formats/processes that are easy to evolve
- Preservation requires the ability to evolve
 - ◆ Current choices make future evolution harder or easier
- Both logical and bit preservation are important
 - ◆ And remain unsolved
 - ◆ At least in terms of scalability and affordability
 - ◆ Several interesting projects are underway
- There are many open, critical problems to work on
 - ◆ Please join us!

Additional material

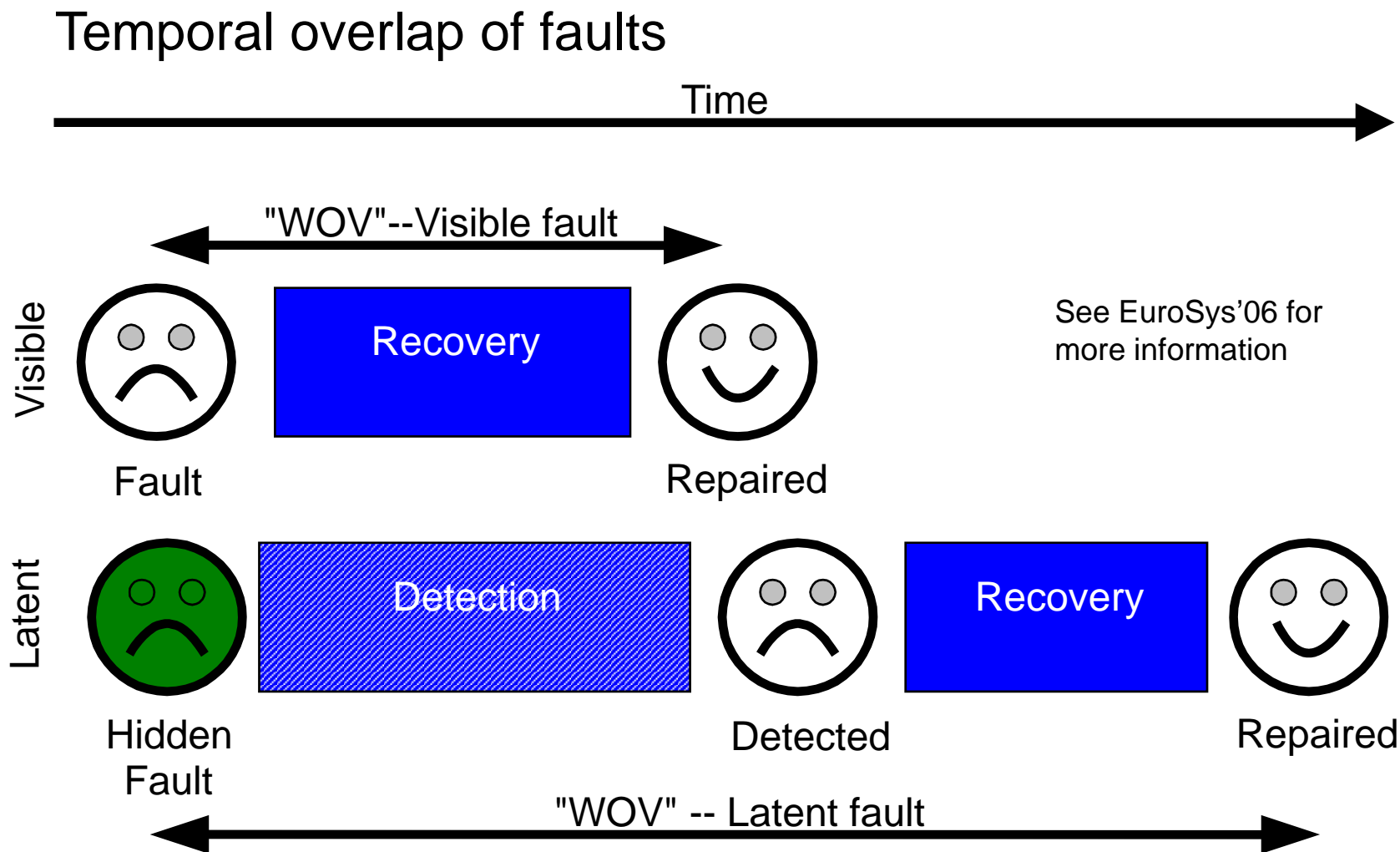
Can we model long-term reliability?

- Abstract reliability model for replicated data
 - ◆ Applies to all units of replication
 - ◆ Applies to many types of faults
- Extend RAID model
 - ◆ Account for latent as well as visible faults
 - ◆ Account for correlated faults: temporal and spatial
- Simple, coarse model
 - ◆ Suggest and compare strategies (choose trade-offs)
 - ◆ Point out areas where we need to gather data
- *Not for exact reliability numbers*

A current approach

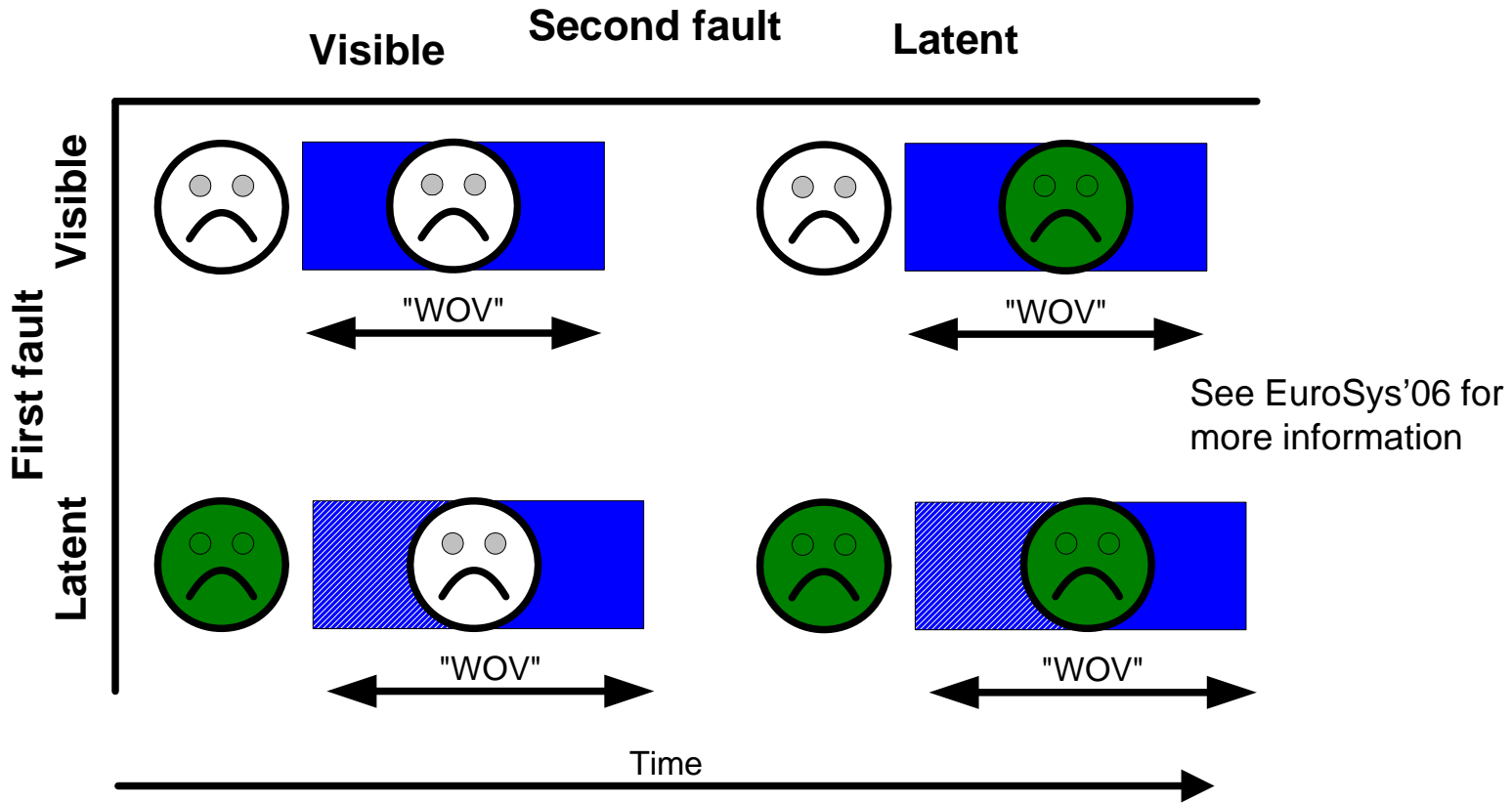
- Start with two replicas, then add more
- Derive MTDDL of mirrored data in the face of
 - ◆ Both immediately visible and latent faults
- Mirrored data is unrecoverable
 - ◆ If copy fails before initial fault can be repaired
- Time between fault and its repair is
 - ◆ *Window of Vulnerability (WOV)*

Window of vulnerability



➤ Want detection time to be small

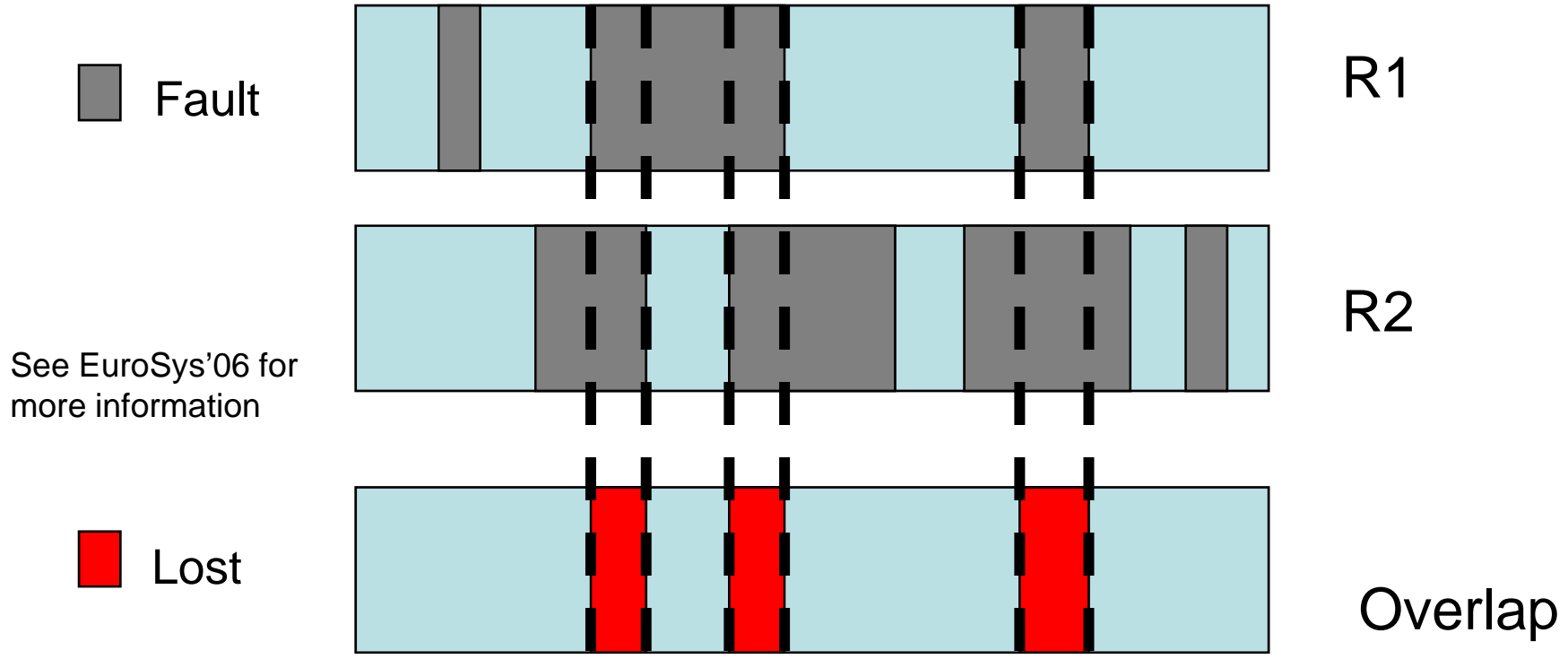
Data loss cases with 2 replicas



➤ Overall probability = sum of each case

Spatial overlap of faults

➤ Temporal overlap alone overstates likelihood of data loss



➤ Faults may be bits, sectors, files, disks, arrays, etc.

➤ If any two faults overlap, data is lost

➤ The smaller the faults, the less likelihood of overlap

Completing the model

- Multiply temporal and spatial probabilities
 - ◆ For each of the four loss cases
- Correlation: use multiplicative scaling factors for
 - ◆ Temporal correlation of faults
 - ◆ Spatial correlation of faults
- We also extend the model for further replication

Example using the model

- Shorter detection time helps how much?
- Portion of real archive (www.archive.org)
 - ◆ Monthly snapshots of web pages
 - ◆ 1.5 million immutable files
 - ◆ 1795 200GB ATA drives, “JBOD”
 - ◆ Mean time to visible (disk) failure: 20 hours
 - ◆ Almost 3 years of monthly file checksums
 - ◆ Mean time to latent fault 1531 hours
- See slide #15 for the results